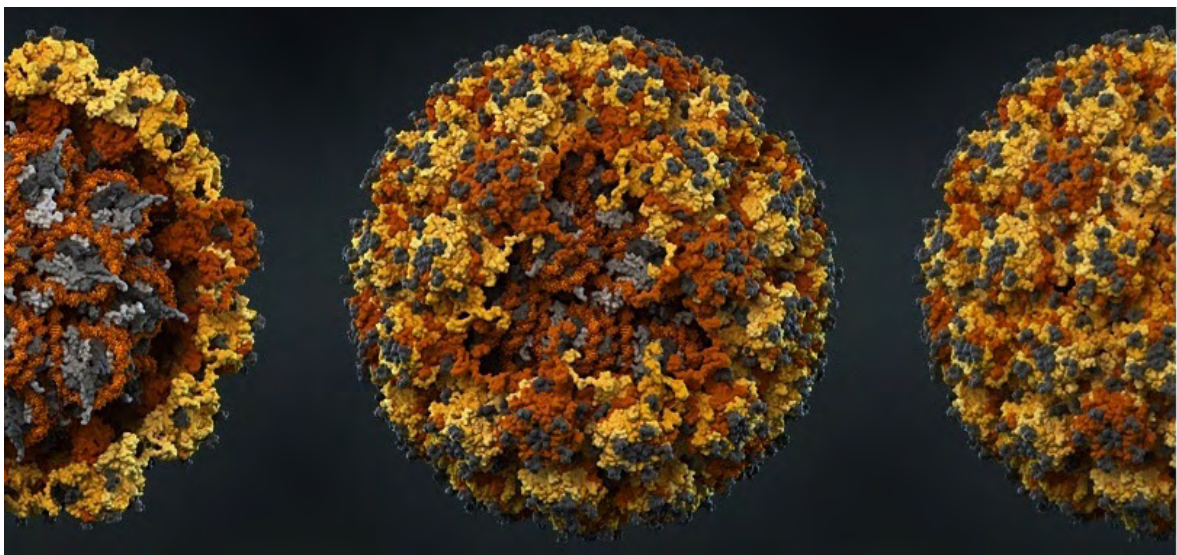


**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΔΗΜΗΤΡΗΣ ΤΣΑΚΟΓΙΑΝΝΗΣ

**ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΜΟΡΦΗΣ ΤΟΥ ΓΟΝΙΔΙΩΜΑΤΟΣ ΤΟΥ
ΣΤΕΛΕΧΟΥΣ HPV-16, ΣΥΣΧΕΤΙΣΗ ΜΕ ΤΗΝ ΚΛΙΝΙΚΗ
ΕΙΚΟΝΑ ΤΟΥ ΑΣΘΕΝΗ ΚΑΙ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΘΕΣΗΣ
ΕΝΣΩΜΑΤΩΣΗΣ ΤΟΥ ΙΟΥ ΣΤΟ ΚΥΤΤΑΡΙΚΟ ΓΟΝΙΔΙΩΜΑ**



2014

**«ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΜΟΡΦΗΣ ΤΟΥ ΓΟΝΙΔΙΩΜΑΤΟΣ ΤΟΥ
ΣΤΕΛΕΧΟΥΣ ΗΡV-16, ΣΥΣΧΕΤΙΣΗ ΜΕ ΤΗΝ ΚΛΙΝΙΚΗ ΕΙΚΟΝΑ ΤΟΥ
ΑΣΘΕΝΗ ΚΑΙ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΘΕΣΗΣ ΕΝΣΩΜΑΤΩΣΗΣ ΤΟΥ ΙΟΥ
ΣΤΟ ΚΥΤΤΑΡΙΚΟ ΓΟΝΙΔΙΩΜΑ»**

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ

Π. ΜΑΡΚΟΥΛΑΤΟΣ (ΕΠΙΒΛΕΠΩΝ)

Καθηγητής Εφαρμοσμένης Μικροβιολογίας με έμφαση στη Βιοτεχνολογία
Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Σ. ΛΕΒΙΔΙΩΤΟΥ- ΣΤΕΦΑΝΟΥ

Καθηγήτρια Μικροβιολογίας
Ιατρική Σχολή Πανεπιστημίου Ιωαννίνων

G. ORFANOUDAKIS

Καθηγητής – Δ/ντης De l' Ecole Superieure de Biotechnologie de Strasbourg
Universite de Strasbourg

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Π. ΜΑΡΚΟΥΛΑΤΟΣ

Καθηγητής Εφαρμοσμένης Μικροβιολογίας με έμφαση στη Βιοτεχνολογία
Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Σ. ΛΕΒΙΔΙΩΤΟΥ- ΣΤΕΦΑΝΟΥ

Καθηγήτρια Μικροβιολογίας
Ιατρική Σχολή Πανεπιστημίου Ιωαννίνων

G. ORFANOUDAKIS

Καθηγητής – Δ/ντης De l' Ecole Superieure de Biotechnologie de Strasbourg
Universite de Strasbourg

Δ. ΚΟΜΙΩΤΗΣ

Καθηγητής Οργανικής Χημείας με έμφαση στη σύνθεση βιοδραστικών μορίων
Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Κ. ΓΚΑΡΤΖΟΝΙΚΑ

Επίκουρη Καθηγήτρια Μικροβιολογίας
Ιατρική Σχολή Πανεπιστημίου Ιωαννίνων

Γ. ΑΜΟΥΤΖΙΑΣ

Επίκουρος Καθηγητής Βιοπληροφορικής στη Γενωμική
Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Δ. ΜΟΣΙΑΛΟΣ

Επίκουρος Καθηγητής Βιοτεχνολογίας Μικροβίων
Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Δημήτρης Τσακογιάννης

**ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΜΟΡΦΗΣ ΤΟΥ ΓΟΝΙΔΙΩΜΑΤΟΣ ΤΟΥ
ΣΤΕΛΕΧΟΥΣ ΗΡV-16, ΣΥΣΧΕΤΙΣΗ ΜΕ ΤΗΝ ΚΛΙΝΙΚΗ ΕΙΚΟΝΑ ΤΟΥ
ΑΣΘΕΝΗ ΚΑΙ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΘΕΣΗΣ ΕΝΣΩΜΑΤΩΣΗΣ ΤΟΥ ΙΟΥ
ΣΤΟ ΚΥΤΤΑΡΙΚΟ ΓΟΝΙΔΙΩΜΑ**

Δημήτρης Τσακογιάννης

2014

**ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΜΟΡΦΗΣ ΤΟΥ ΓΟΝΙΔΙΩΜΑΤΟΣ ΤΟΥ
ΣΤΕΛΕΧΟΥΣ ΗΡV-16, ΣΥΣΧΕΤΙΣΗ ΜΕ ΤΗΝ ΚΛΙΝΙΚΗ
ΕΙΚΟΝΑ ΤΟΥ ΑΣΘΕΝΗ ΚΑΙ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΘΕΣΗΣ
ΕΝΣΩΜΑΤΩΣΗΣ ΤΟΥ ΙΟΥ ΣΤΟ ΚΥΤΤΑΡΙΚΟ ΓΟΝΙΔΙΩΜΑ**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ**

Αριθμός προκαταρτικών σελίδων: 14

Συνολικός αριθμός σελίδων :188

Αριθμός πινάκων: 25

Αριθμός εικόνων: 33

Αριθμός βιβλιογραφικών παραπομπών: 280

Περίληψη

Οι Ιοί των Θηλωμάτων του ανθρώπου (HPV) είναι μικροί, δίκλωνοι κυκλικοί DNA ιοί, χωρίς εξωτερικό περίβλημα και προσβάλλουν το επιθήλιο του δέρματος ή το επιθήλιο του βλεννογόνου. Περισσότεροι από 150 HPV τύποι έχουν ανιχνευτεί, ενώ πάνω από 40 τύποι συνδέονται με αλλοιώσεις του ουρογεννητικού συστήματος και διαχωρίζονται στους τύπους χαμηλού (low-risk) και υψηλού κινδύνου (high-risk). Επιδημιολογικές μελέτες έχουν προτείνει ότι "επίμονη" μόλυνση με HPV τύπους υψηλού κινδύνου, είναι σημαντικός παράγοντας για την ανάπτυξη προκαρκινικών και καρκινικών αλλοιώσεων με τους HPV16 και HPV18 να εντοπίζονται σε περισσότερο από το 72% των περιπτώσεων καρκίνου του τραχήλου της μήτρας παγκοσμίως. Η εμφάνιση καρκίνου συνδέεται με μόλυνση με HPV τύπους υψηλού κινδύνου και με την απορύθμιση της έκφρασης των ιικών ογκογονιδίων E6 και E7. Η ανεξέλεγκτη έκφραση των ογκογονιδίων οδηγεί στον εκτεταμένο πολλαπλασιασμό των μολυσμένων κυττάρων, στην ατελή επιδιόρθωση του DNA και στη συσσώρευση μεταλλάξεων στα μολυσμένα κύτταρα.

Η "επίμονη" υψηλού βαθμού αλλοίωση συνδέεται με αυξημένο κίνδυνο ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα και την πρόοδο της αλλοίωσης σε καρκίνο. Η ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα αποσταθεροποιεί την έκφραση των γονιδίων E6 και E7. Το γεγονός αυτό δίνει την δυνατότητα στον ιό να διεγείρει έντονα τον κυτταρικό κύκλο και την διαρκή είσοδο του κυττάρου στη φάση της μίτωσης, δίνοντας στα μολυσμένα κύτταρα ισχυρό αναπτυξιακό πλεονέκτημα. Σε δεύτερη φάση μετά τον συνεχή πολλαπλασιασμό των μολυσμένων κυττάρων προκαλείται συσσώρευση γενετικών λαθών και αποσταθεροποίηση του γονιδιώματος του ξενιστή με αποτέλεσμα την δημιουργία καρκινικών αλλοιώσεων.

Σκοπός της παρούσας διατριβής ήταν η μελέτη του γονιδιώματος του HPV16, ενώ παράλληλα επιχειρήθηκε ο προσδιορισμός της μορφής του γονιδιώματος του ιού και ο προσδιορισμός των θέσεων ενσωμάτωσης του HPV16 στο κυτταρικό χρωμόσωμα. Η μελέτη πραγματοποιήθηκε σε κλινικά δείγματα τραχήλου της μήτρας που απομονώθηκαν από τον Ελληνικό πληθυσμό, θετικά για μόλυνση με τον ιό. Η νουκλεοτιδική και φυλογενετική ανάλυση των αποτελεσμάτων έδειξε ότι το πρότυπο στέλεχος του HPV16 καθώς και η Ευρωπαϊκή παραλλαγή του γονιδιώματος του ιού επικρατούν στον Ελληνικό πληθυσμό, ενώ παράλληλα παρατηρήθηκε ότι μη Ευρωπαϊκές παραλλαγές του HPV16 DNA κυκλοφορούν στον Ελληνικό πληθυσμό.

Ακόμη εντοπίστηκαν καινούργιες μεταλλάξεις εντός των γονιδίων E1, E2, E4, E6 και E7 που δεν έχουν αναφερθεί στη διεθνή βιβλιογραφία καθώς επίσης και φαινόμενα διπλασιασμού και ενδοτυπικού ανασυνδυασμού μεταξύ διαφορετικών παραλλαγών του HPV16 DNA. Τέλος, οι διαφορετικές περιοχές της πρώιμης περιοχής του γονιδιώματος του HPV16 DNA εξελίσσονται με διαφορετικό τρόπο και αυτό ίσως να σχετίζεται με τις διαφορετικές λειτουργίες που επιτελούν οι ιικές πρωτεΐνες στο κύκλο ζωής του ιού.

Παράλληλα, προσδιορίστηκε η μορφή του HPV16 DNA, χαρτογραφώντας τις πιο συχνές θέσεις ρήξης του ιικού DNA. Η ανάλυση των αποτελεσμάτων έδειξε, ότι η ενσωμάτωση του HPV16 είναι ένα φαινόμενο το οποίο συμβαίνει νωρίς στο κύκλο ζωής του ιού. Τέλος, από τον προσδιορισμό των θέσεων ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα εντοπίστηκαν φαινόμενα αναδιάταξης του HPV16 DNA προτείνοντας ένα μοντέλο, σύμφωνα με το οποίο ενσωματωμένα μόρια DNA σε μορφή αλυσομερών HPV16 DNA - κυτταρικό DNA έρχονται σε αντιπαράλληλο προσανατολισμό και ανασυνδυάζονται μεταξύ τους μέσω μιας δομής θηλιάς. Το φαινόμενο αυτό προκαλεί διαρκώς αποσταθεροποίηση του κυτταρικού χρωμοσώματος, γεγονός που δηλώνει, ότι ακόμη και μετά την ενσωμάτωση του ιού στο κυτταρικό χρωμόσωμα ο ιός συνεχίζει να προκαλεί διαρκώς φαινόμενα αποσταθεροποίησης του κυτταρικού DNA. Συμπερασματικά, η χαρτογράφηση των συνηθέστερων θέσεων ρήξης των E1 και E2 γονιδίων σε συνδυασμό με τη νουκλεοτιδική ανάλυση των γονιδίων E1, E2 και E6 του HPV16 και τον προσδιορισμό γεγονότων αναδιάταξης του HPV16 DNA θα πρέπει να λαμβάνεται σοβαρά υπόψη για τον προσδιορισμό της μορφής του γονιδιώματος των διαφορετικών HPV16 στελεχών και της περαιτέρω κατανόησης των γεγονότων τα οποία εμπλέκουν τον HPV16 στην ογκογένεση.

Πρόλογος - Ευχαριστίες

Η Διδακτορική Διατριβή πραγματοποιήθηκε στο Εργαστήριο Μοριακής Ιολογίας του τμήματος Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας με επιβλέποντα καθηγητή τον κ. Π. Μαρκουλάτο. Το θέμα της διατριβής είναι "Προσδιορισμός της μορφής του γονιδιώματος του στελέχους HPV-16, συσχέτιση με την κλινική εικόνα του ασθενή και προσδιορισμός της θέσης ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα", δίνοντας μου την δυνατότητα να μελετήσω την βιολογία του ιού και να εντυπώσω στο χώρο της Μοριακής Ιολογίας και της Έρευνας. Για το λόγο αυτό θα ήθελα να εκφράσω ένα μεγάλο ευχαριστώ στον καθηγητή μου κ. Π. Μαρκουλάτο. Η συνεργασία μας κράτησε για επτά ολόκληρα χρόνια, μέσα από τα οποία μου δίδαξε με ιδιαίτερη ευγένεια τον χώρο της επιστήμης και τον τρόπο της διαρκούς αναζήτησης για την διερεύνηση και την επίλυση καιρίων επιστημονικών ερωτημάτων. Οφείλω ένα μεγάλο ευχαριστώ και ιδιαίτερη ευγνωμοσύνη για τις γνώσεις που μου προσέφερε, την πολύτιμη και διαρκή καθοδήγηση του και καθοριστική του συμβολή στην επίλυση προβλημάτων και την επίτευξη των στόχων μου.

Επίσης θα ήθελα να εκφράσω ένα μεγάλο ευχαριστώ στην καθηγήτρια Λεβιδιώτου - Στεφάνου Σταματίνα και στον καθηγητή Ορφανουδάκη Γιώργο για τη συμμετοχή και συμβολή τους στη τριμελή συμβουλευτική επιτροπή. Ένα μεγάλο ευχαριστώ στα υπόλοιπα μέλη της επταμελούς εξεταστικής επιτροπής τον καθηγητή Κομιώτη Δημήτρη και τους επίκουρους καθηγητές Αμούτζια Γρηγόρη, Μόσιαλο Δημήτρη και Γκαρτσονίκα Κωνσταντίνα για την συμμετοχή τους στην επταμελή εξεταστική επιτροπή και για το χρόνο που αφιέρωσαν στη διόρθωση και αξιολόγηση της διατριβής.

Επιπλέον ευχαριστώ θερμά τη κ. Νέγκα Χαριτίνη από το εργαστήριο Κυτταροπαθολογίας του Πανεπιστημιακού Νοσοκομείου Λάρισας, την κ. Πανοτοπούλου Ευσταθία υπεύθυνη τμήματος Ιολογίας ΚΟΕΠΧ «Γ. Παπανικολάου» ΑΟΝΑ Αγ. Σάββας, τον κ. Σκορδά Βασίλη και την κ. Ζαρκαδούλα Ιωάννα για την παραχώρηση των κλινικών δειγμάτων που χρησιμοποιήθηκαν στη παρούσα διατριβή.

Ένα μεγάλο ευχαριστώ οφείλω στους φίλους μου και συναδέλφους Κυριακοπούλου Ζαχαρούλα, Ρούτερ Ειρήνη και Δημητρίου Τηλέμαχο για τη συνεργασία και για την ουσιαστική φιλία που αναπτύχθηκε όλο το διάστημα της παρουσίας μου στην ερευνητική ομάδα. Τους ευχαριστώ θερμά για τις δημιουργικές

στιγμές και για την αμέριστη συμπαράσταση τους σε δύσκολες καταστάσεις. Επιπλέον θα ήθελα να εκφράσω ένα μεγάλο ευχαριστώ σε όλους του προπτυχιακούς και μεταπτυχιακούς φοιτητές με τους οποίους συνεργάστηκα και κυρίως στη μεταπτυχιακή φοιτήτρια Μπλέτσα Μαγδαληνή, για την πολύτιμη βοήθεια της στο τελευταίο μέρος της διατριβής.

Τέλος ένα θερμό ευχαριστώ στους γονείς μου Κωνσταντίνο και Ιωάννα για όλα τα χρόνια που με στηρίζουν, δίνοντας μου τα μέσα και τη δυνατότητα να πετύχω τους στόχους μου. Τους οφείλω ιδιαίτερη ευγνωμοσύνη και τους χρωστάω πολλά ευχαριστώ για τη συμβολή τους σε όλη την πορεία των σπουδών μου και που είναι διαρκώς δίπλα μου σε κάθε βήμα και επιλογή μου. Ακόμη ένα μεγάλο ευχαριστώ στην αδελφή μου Ευαγγελία για τη συμπαράσταση και την αισιοδοξία που μου μεταφέρει σε όλη τη πορεία μου μέχρι σήμερα

Τσακογιάννης Δημήτρης
Βιοχημικός - Βιοτεχνολόγος, MSc

*Στους γονείς μου,
Κωνσταντίνο και Ιωάννα
και την αδελφή μου,
Ευαγγελία*

Περιεχόμενα

Κεφάλαιο 1	1
1. Εισαγωγή	1
1.1 Ιστορική αναδρομή	1
1.2 Γενικά χαρακτηριστικά των Papilloma ιών του ανθρώπου	2
1.3 Ταξινόμηση των Papilloma ιών	3
1.3.1 Άλφα-Papilloma ιοί.....	6
1.3.2 Βήτα, Γάμμα, Μι και Νι- Papilloma ιοί.....	8
1.4 Εξέλιξη των Ανθρώπινων Papilloma ιών	9
1.5. Δομή των Papilloma ιών του ανθρώπου	11
1.5.1 Δομή καψιδίου	11
1.6 Οργάνωση του γονιδιώματος.....	13
1.6.1 Η ιική πρωτεΐνη E1	14
1.6.2 Η ιική πρωτεΐνη E2	16
1.6.3 Η ιική πρωτεΐνη E4	18
1.6.4 Η ιική πρωτεΐνη E5	19
1.6.5 Η ιική πρωτεΐνη E6	21
1.6.6 Η ιική πρωτεΐνη E7	22
1.7 Μορφή του γονιδιώματος	25
1.8 Θέσεις ενσωμάτωσης.....	29
1.9 Ο κύκλος ζωής των Papilloma ιών του ανθρώπου	31
1.9.1 Είσοδος του ιού στο κύτταρο.....	31
1.9.2 Αντιγραφή του ιικού γονιδιώματος	32
1.9.3 Μεταγραφή του ιικού γονιδιώματος.....	33
1.9.4 Απελευθέρωση του ιού	35
1.10 Ιστολογικές αλλοιώσεις	36

1.11 Επιδημιολογία.....	38
Κεφάλαιο 2.....	40
2. Υλικά και Μέθοδοι	40
2.1 Κλινικά δείγματα	40
2.2 Εκχύλιση DNA	43
2.2.1 Εκχύλιση DNA (ThinPrep).....	43
2.2.2 Εκχύλιση DNA (Τραχηλικά επιχρίσματα)	44
2.2.3 Εκχύλιση DNA (Ιστός μονιμοποιημένος σε παραφίνη)	44
2.3 Έλεγχος της εκχύλισης	45
2.4 Μοριακή ανάλυση του γονιδιώματος του HPV16.....	45
2.4.1 Σχεδιασμός εκκνητικών μορίων	46
2.4.2 PCR.....	46
2.4.3 Auto Nested PCR.....	46
2.4.4 Ηλεκτοφόρηση προϊόντων της PCR.....	49
2.4.5 Απομόνωση των προϊόντων από το πήκτωμα αγαρόζης.....	49
2.4.6 Μοριακή κλωνοποίηση των ενισχυμένων γονιδίων και αλληλούχιση.....	49
2.4.7 Έλεγχος των αλληλουχιών.....	50
2.4.8 Ομοπαράθεση νουκλεοτιδικών και αμινοξικών αλληλουχιών.....	51
2.4.9 Κατασκευή φυλογενετικών δέντρων	51
2.4.10 Ανάλυση επιλεκτικής πίεσης	53
2.4.11 Προσδιορισμός του μέσου ρυθμού υποκατάστασης των γονιδίων	54
2.5 Προσδιορισμός της μορφής του γονιδιώματος του HPV16	55
2.5.1 Χαρτογράφηση των E1 και E2 γονιδίων	55
2.5.2 Έλεγχος ακεραιότητας του HPV16 γονιδιώματος.....	57
2.5.3 Προσδιορισμός ευαισθησίας της μεθόδου.....	57
2.6 Προσδιορισμός της μορφής του γονιδιώματος του HPV16 με Ποσοτική PCR Πραγματικού Χρόνου (quantitative - Real Time PCR).....	59

2.6.1 Κατασκευή πλασμιδίων	59
2.6.2 Ποσοτική PCR Πραγματικού Χρόνου.....	60
2.6.3 Καθορισμός του εύρους των τιμών E1/E6 και E2/E6	62
2.6.4 Στατιστική μελέτη των αποτελεσμάτων	62
2.7 Προσδιορισμός της θέσης ενσωμάτωσης του HPV16 στο κυτταρικό χρωμόσωμα	64
2.7.1 RS-PCR.....	65
2.7.2 Nested RS-PCR.....	65
2.7.3 Ανάλυση των αλληλουχιών	67
2.7.4 Επιβεβαίωση αναδιατάξεων του HPV16 DNA	67
Κεφάλαιο 3.....	69
3. Αποτελέσματα	69
3.1 Μοριακή ανάλυση του γονιδιώματος του HPV16.....	69
3.1.1 Νουκλεοτιδική ανάλυση των E6 και E7 γονιδίων	69
3.1.2 Ανίχνευση νέων αλληλουχιών των E6 και E7 γονιδίων.....	77
3.1.3 Νουκλεοτιδική ανάλυση του E2 γονιδίου	82
3.1.4 Νουκλεοτιδική ανάλυση του E4 γονιδίου	87
3.1.5 Νουκλεοτιδική ανάλυση του E1 γονιδίου	94
3.1.6 Ανάλυση διπλασιασμού στο γονίδιο E1.....	103
3.2 Προσδιορισμός της μορφής του γονιδιώματος του HPV16	107
3.2.1 Χαρτογράφηση των γονιδίων E1 και E2	107
3.2.2 Προσδιορισμός της μορφής του γονιδιώματος του HPV16 με Ποσοτική PCR Πραγματικού Χρόνου (quantitative - Real Time PCR).....	112
3.3 Προσδιορισμός θέσεων ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα.....	120
Κεφάλαιο 4.....	126
4. Συζήτηση	126
4.1 Νουκλεοτιδική ανάλυση των γονιδίων E6 και E7.....	127
4.2 Νουκλεοτιδική ανάλυση του E2 γονιδίου	131

4.2 Νουκλεοτιδική ανάλυση του E4 γονιδίου	133
4.3 Νουκλεοτιδική ανάλυση του E1 γονιδίου	135
4.4 Ανάλυση του διπλασιασμού E1-1374 ^{63nt}	136
4.5 Εξελικτική μελέτη της πρώιμης περιοχής του HPV16 DNA	138
4.6 Προσδιορισμός της μορφής του γονιδιώματος του HPV16	140
4.7 Φαινόμενα αναδιάταξης του HPV16 DNA	146
Βιβλιογραφία	150
Abstract	171
ΔΗΜΟΣΙΕΥΣΕΙΣ	173

Κατάλογος Πινάκων

Πίνακας 2.1. Τριάντα τρία κλινικά δείγματα υψηλού βαθμού αλλοίωσης.....	41
Πίνακας 2.2. Σαράντα κλινικά δείγματα χαμηλού βαθμού αλλοίωσης.....	42
Πίνακας 2.3. Έντεκα δείγματα καρκίνου.....	43
Πίνακας 2.4. Εκκινητές ενίσχυσης πρωίμης περιοχής HPV16 DNA.....	48
Πίνακας 2.5. Εκκινητές ενίσχυσης E1 και E2 γονιδίων.....	56
Πίνακας 2.6. Εκκινητές για την ποσοτικοποίηση των E1, E2, E6 και GAPDH γονιδίων.....	61
Πίνακας 2.7. Περιγραφή των διαλυμάτων των πλασμιδίων.....	62
Πίνακας 2.8. Η κωδικοποίηση των εκκινητών RS-PCR.....	66
Πίνακας 3.1. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις των E6 και E7 γονιδίων.....	75
Πίνακας 3.2. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις των E6 και E7 γονιδίων.....	76
Πίνακας 3.3. Εξελικτική σχέση των νέων αλληλουχιών E6-E7.....	81
Πίνακας 3.4. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις του E2 γονιδίου.....	84
Πίνακας 3.5. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις του E2 γονιδίου.....	85
Πίνακας 3.6. Νουκλεοτιδική και αμινοξική ανάλυση του E4 γονιδίου.....	91
Πίνακας 3.7. Νουκλεοτιδική και αμινοξική ανάλυση του E4 γονιδίου.....	92
Πίνακας 3.8. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις μεταξύ των νουκλεοτιδίων 878 έως 1523.....	98
Πίνακας 3.9. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις μεταξύ των νουκλεοτιδίων 1503 έως 1991.....	99
Πίνακας 3.10. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις μεταξύ των νουκλεοτιδίων 1952 έως 2485.....	100
Πίνακας 3.11. Νουκλεοτιδικές και αμινοξικές μεταλλάξεις μεταξύ των νουκλεοτιδίων 2423 έως 2848.....	101
Πίνακας 3.12. Η κατανομή του E1-1374 ⁶³ nt διπλασιασμού.....	106
Πίνακας 3.13. Θέσεις ρήξης των γονιδίων E1 και E2.....	111
Πίνακας 3.14. Τιμές E1/E6, E2/E6 στα διαλύματα πλασμιδίων.....	114
Πίνακας 3.15. Δειγμάτα που φέρουν την επισωμική μορφή του HPV16.....	117
Πίνακας 3.16. Δειγμάτα που φέρουν την ενσωματωμένη μορφή του HPV16.....	118
Πίνακας 3.17. E1/E6, E2/E6 και της αναδιάταξης του γονιδιώματος του HPV16.....	125

Κατάλογος Εικόνων

Εικόνα 1.1. Φυλογενετικό δέντρο 118 PVs με βάση την αλληλουχία του γονιδίου L1.....	4
Εικόνα 1.2. Φυλογενετικό δέντρο με βάση την αλληλουχία του γονιδίου L1 189 PVs.....	5
Εικόνα 1.3. Φυλογενετικό δέντρο με βάση τη περιοχή LCR και του E6 γονιδίου.....	7
Εικόνα 1.4. Εξελικτική σχέση μεταξύ PVs.....	8
Εικόνα 1.5. T=7 HPV16, αλληλεπίδραση ιικού DNA με ιστόνες.....	11
Εικόνα 1.6. Δομικά στοιχεία της L1.....	12
Εικόνα 1.7. Οργάνωση του γονιδιώματος του HPV16	13
Εικόνα 1.8. Σχηματική αναπαράσταση της πρωτεΐνης E1 και E2	16
Εικόνα 1.9. Ρυθμιστική περιοχή LCR.....	18
Εικόνα 1.10. Μοντέλο εξαμερούς καναλιού ιόντων της E5 πρωτεΐνης	20
Εικόνα 1.11. Σχηματική αναπαράσταση της E6 πρωτεΐνης.....	22
Εικόνα 1.12. Σχηματική αναπαράσταση E7 πρωτεΐνης	24
Εικόνα 1.13. Σχηματική αναπαράσταση των στόχων των E6 και E7 γονιδίων	24
Εικόνα 1.14. Σχέση ενσωμάτωσης του ιού και τραχηλικής ενδοεπιθηλιακής αλλοίωσης.....	28
Εικόνα 1.15. Μοντέλο αλληλεπίδρασης του ιικού καψιδίου με την εξωκυττάρια μήτρα.....	31
Εικόνα 1.16. Πρόσδεση και της ενδοκυττάρωσης του HPV16.....	32
Εικόνα 1.17. Ενίσχυση του ιικού γονιδιώματος.....	33
Εικόνα 1.18. Έκφραση των γονιδίων του HPV16 σε ενδοεπιθηλιακές αλλοιώσεις	35
Εικόνα 1.19. Παραδείγματα τραχηλικής ενδοεπιθηλιακής νεοπλασίας.....	37
Εικόνα 1.20. Παγκόσμια κατανομή HPV τύπων.....	39
Εικόνα 2.1. Απεικόνιση του γονιδιώματος του HPV16	45
Εικόνα 2.2. Θέση πρόσδεσης των εκκινητών της μεθόδου RS-PCR	64
Εικόνα 3.1. Φυλογενετικό δέντρο με βάση την αλληλουχία του E6 γονιδίου	74
Εικόνα 3.2. Bootscan ανάλυση του δείγματος A15	79
Εικόνα 3.3. Φυλογενετικό δέντρο ανασυνδυασμών.....	81
Εικόνα 3.4. Φυλογενετικό δέντρο με βάση την αλληλουχία του E4 γονιδίου	90
Εικόνα 3.5. Φυλογενετικό δέντρο με βάση την αλληλουχία του E1 γονιδίου	97
Εικόνα 3.6. Διπλασιασμός.....	104
Εικόνα 3.7. Παρουσίαση των ποσοστών ρήξης των γονιδίων E1 και E2	110
Εικόνα 3.8. Παρουσίαση των θέσεων ρήξης των E1 και E2 γονιδίων.....	110
Εικόνα 3.9. Κατανομή της μορφής του γονιδιώματος του HPV16.....	116
Εικόνα 3.10. Αναδιάταξη του γονιδιώματος του HPV16.....	123
Εικόνα 4.1. Πιθανό μοντέλο σχηματισμού τμημάτων αναδιάταξης HPV16 DNA.....	148

Κεφάλαιο 1

1. Εισαγωγή

1.1 Ιστορική αναδρομή

Οι Ιοί των Θηλωμάτων - Papillomaviruses (PVs) συγκροτούν την μεγάλη οικογένεια των Papillomaviridae. Η οικογένεια των Papillomaviridae είναι μια μεγάλη ετερογενής ομάδα ιών που προσβάλλει τα περισσότερα θηλαστικά, πτηνά και ερπετά, ενώ περισσότεροι από 150 τύποι έχουν ανιχνευτεί στον άνθρωπο (Human Papillomavirus - HPV), (Bernard et al., 2010). Ανακαλύφθηκαν πρώτη φορά στις αρχές του 20^{ου} αιώνα όταν αποδείχθηκε ότι τα δερματικά θηλώματα μπορούσαν να μεταδοθούν μεταξύ διαφορετικών ασθενών. Η ική φύση των ανθρώπινων θηλωμάτων προσδιορίστηκε το 1907 και είναι ο πρώτος ιός που ανακαλύφθηκε ότι προσβάλλει τον άνθρωπο. Εκείνο το διάστημα οι ερευνητές παρατήρησαν, ότι τα δερματικά θηλώματα μπορούσαν να μεταδοθούν από έναν ασθενή σε έναν άλλο, υποβάλλοντας τα σε φιλτράρισμα χρησιμοποιώντας φίλτρα με τα οποία απομόνωναν βακτήρια. Οι PVs απομονώθηκαν για πρώτη φορά από κουνέλια από τον Richard Shope το 1933. Ο πρώτος Ιός των Θηλωμάτων που απομονώθηκε ονομάστηκε Shope Papillomavirus ή CRPV (cottontail rabbit papillomavirus). Ο Shope παρατήρησε ότι μολυσμένα κουνέλια του γένους *sylvilagus* εμφάνιζαν θηλώματα τα οποία μετά από σύντομο χρονικό διάστημα μετατρέπονταν σε κακοήθεις όγκους. Ο Shope απομόνωσε ιικά σωματίδια από όγκους από μολυσμένους λαγούς και στην συνέχεια τα εμβολίασε σε οικόσιτους υγιείς λαγούς, οι οποίοι ανέπτυξαν τις ίδιες αλλοιώσεις (Shope, 1933). Ο Shope Papillomavirus αλληλουχήθηκε το 1984 και έδειξε σημαντική ομοιότητα με τον ιό HPV1 (Giri et al., 1985). Το 1935 ο Francis Peyton Rous που είχε αποδείξει την πρόκληση σαρκώματος στα κοτόπουλα από ιό, συσχέτισε τους Ιούς των Θηλωμάτων με καρκίνο του δέρματος στα μολυσμένα αυτά κουνέλια και για 30 χρόνια συνεχόμενων ερευνών χρησιμοποιούσε καλοήθη ιικά θηλώματα ως πρόδρομο των σημερινών χημικών καρκινογόνων για την πρόκληση όγκων σε ερευνητικό επίπεδο (Rous, 1935; Rous, 1940; Rous et al., 1952). Αυτή ήταν η πρώτη απόδειξη που ενοχοποιούσε τους ιούς για την πρόκληση καρκίνου στα θηλαστικά. Από τότε

διάφοροι τύποι των Ιών των Θηλωμάτων εντοπίστηκαν και συνδέθηκαν με την δημιουργία όγκων σε πολλά είδη μεταξύ των οποίων και ο άνθρωπος.

Το 1976, ο Γερμανός Ιολόγος Harald zur Hausen δημοσίευσε ότι οι Ιοί των Ανθρώπινων Θηλωμάτων (Human Papillomaviruses, HPV) παίζουν σημαντικό ρόλο στην ανάπτυξη καρκίνου του τραχήλου της μήτρας και τα έτη 1983 και 1984 ο Harald zur Hausen με τους συνεργάτες του απομόνωσε τα πρώτα στελέχη HPV16 και HPV18 σε γυναίκα που έπασχε από καρκίνο του τραχήλου της μήτρας (Gissmann et al., 1984, Zur Hausen, 1996). Το 2008, του απενεμήθη το βραβείο Nobel Ιατρικής για την έρευνα του. Αργότερα ανακαλύφθηκε, ότι συγκεκριμένοι τύποι HPV συσχετίζονται με τον καρκίνο του τραχήλου της μήτρας. Σήμερα, επιδημιολογικές μελέτες αποδεικνύουν, ότι ο καρκίνος του τραχήλου της μήτρας αποτελεί την τρίτη πιο συχνή ένδειξη κακοήθειας στο γυναικείο πληθυσμό παγκοσμίως (Jemal, 2011).

1.2 Γενικά χαρακτηριστικά των Papilloma ιών του ανθρώπου

Οι Ιοί των Ανθρώπινων Θηλωμάτων (HPV) είναι δίκλωνοι κυκλικοί DNA ιοί, μεγέθους περίπου 8000 ζευγών βάσεων, χωρίς εξωτερικό περίβλημα και διαμέτρου της τάξης 50 έως 60 νανομέτρων. Οι HPV προκαλούν ενδοεπιθηλιακές αλλοιώσεις στον ξενιστή τους (κονδυλώματα, θηλώματα), οι οποίες εκτείνονται από κοινά θηλώματα μέχρι καρκίνο (Zur Hausen, 1996). Μέχρι σήμερα έχουν ταυτοποιηθεί και έχουν πλήρως αλληλουχηθεί περισσότεροι από 150 διαφορετικοί τύποι HPV, ενώ πάνω από 40 τύποι συνδέονται με αλλοιώσεις του ουρογεννητικού συστήματος (Bernard et al., 2010). Επιπλέον έχουν αλληλουχηθεί πάνω από 60 τύποι που προσβάλλουν τα ζώα. Η παρουσία PVs στα θηλαστικά καθώς επίσης και σε πολλούς διαφορετικούς ξενιστές όπως τα πουλιά, χελώνες και φίδια αποδεικνύει την ευρεία παρουσία τους μεταξύ αμνιωτών (Bravo et al., 2010)

Όσον αφορά τους HPV χωρίζονται σε δύο κατηγορίες με βάση τον επιθηλιακό τροπισμό τους. Έτσι έχουμε τους HPV που προσβάλλουν το επιθήλιο του δέρματος και εκείνους που προσβάλλουν το επιθήλιο του βλεννογόνου. Οι δερματικοί HPV τύποι προκαλούν καλοήγη δερματικά θηλώματα, ενώ οι HPV που προσβάλλουν το επιθήλιο του βλεννογόνου αποτελούν τον κύριο αιτιολογικό παράγοντα για την πρόκληση ενδοεπιθηλιακών αλλοιώσεων και ανάπτυξη καρκίνου. Επιπλέον, οι HPV της τελευταίας αυτής τάξης ανάλογα με την έκταση και την σοβαρότητα της βλάβης που

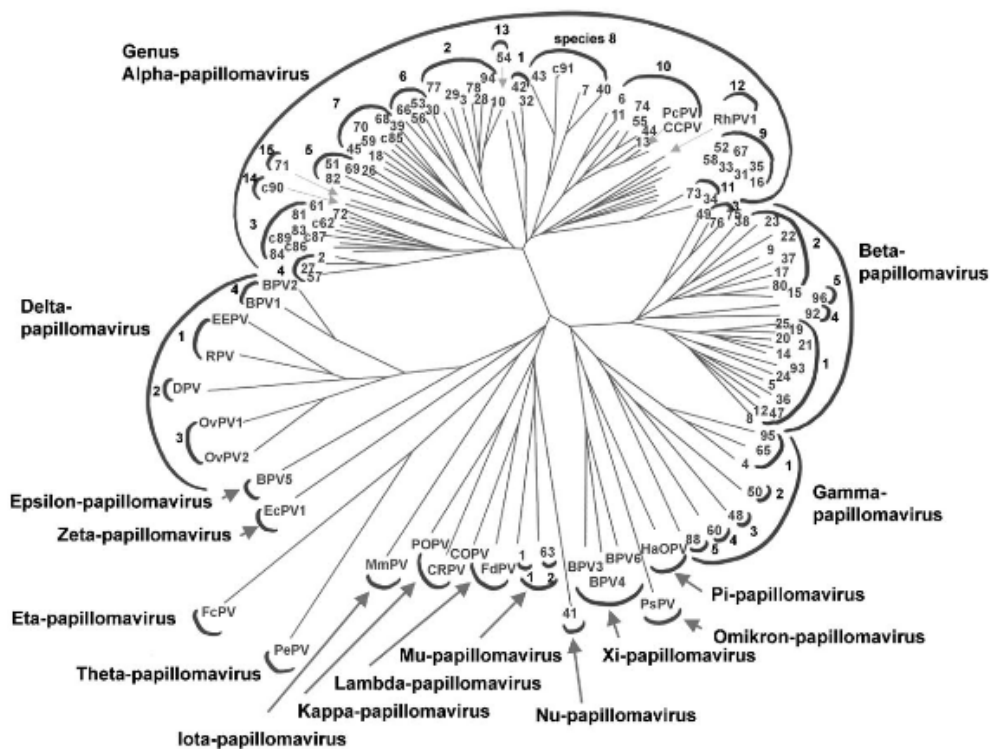
προκαλούν, μπορούν να κατηγοριοποιηθούν σε δύο χαρακτηριστικές ομάδες. Στην ομάδα υψηλού κινδύνου (*high-risk*), τα μέλη της οποίας είναι συνδεδεμένα με την πρόκληση διηθητικού καρκίνου συνήθως του ουρογεννητικού συστήματος και στην ομάδα χαμηλού κινδύνου (*low-risk*), τα μέλη της οποίας εντοπίζονται σπανίως σε καρκίνο άλλα είναι στενά συνδεδεμένα με την πρόκληση θηλωμάτων (Bernard et al., 2010).

1.3 Ταξινόμηση των Papilloma ιών

Οι Ιοί των Θηλωμάτων (Papillomaviruses) αρχικά ομαδοποιήθηκαν μαζί με τους Polyoma ιούς και τον SV40 στην μεγάλη οικογένεια των Papovaviridae. Η ομαδοποίηση αυτή βασίστηκε στην ομοιότητα του καψιδίου των παραπάνω ιών, στην έλλειψη περιβλήματος καθώς επίσης και στην ύπαρξη δίκλωνου κυκλικού μορίου DNA (Danos, 1982). Μετά από την έβδομη αναφορά της Διεθνούς Επιτροπής Ταξινόμησης των Ιών (International Committee on Taxonomy of Viruses, ICTV) έγινε αποδεκτό ότι οι Ιοί των Θηλωμάτων αποτελούν μια ξεχωριστή ταξινομική οικογένεια ιών (Papillomaviridae), η οποία δεν σχετίζεται καθόλου με τους Polyomavirus και τον SV40 (van Regenmortel et al., 2002). Ο διαχωρισμός αυτός έγινε με βάση τη διαφορά στην οργάνωση, το μέγεθος και τη νουκλεοτιδική αλληλουχία μεταξύ των PVs και των μελών της οικογένειας των Papovaviridae. Η de Villiers και οι συνεργάτες της περιέγραψαν τη διαμόρφωση των φυλογενετικών δέντρων των PVs, με βάση τη νουκλεοτιδική αλληλουχία, τα βιολογικά χαρακτηριστικά (είδη ξενιστών, ιστός, παθογένεια, οργάνωση γενώματος) και θέσπισαν την ταξινόμηση των PVs σε γένη (de Villiers et al., 2004). Η ονοματολογία αυτών των γενών βασίζεται στο ελληνικό αλφάβητο και έγινε αμέσως αποδεκτό από την ICTV και την επιστημονική κοινότητα. Το 2004 δεκαέξι ομάδες PVs οργανώθηκαν σε γένη χρησιμοποιώντας την ελληνική αλφάβητο από το γράμμα Άλφα μέχρι το γράμμα Πι με βάση τα κριτήρια που θεσπίστηκαν.

Οι τύποι των Ιών των Ανθρώπινων Θηλωμάτων (HPVs) διαχωρίζονται σε πέντε γένη με βάση την ανάλυση του γονιδιώματος του ιού, με τους διαφορετικούς τύπους να παρουσιάζουν διαφορετικά χαρακτηριστικά στο κύκλο ζωής τους και να συνδέονται με διαφορετικό βαθμό με την πρόκληση ενδοεπιθηλιακής αλλοίωσης (Doorbar et al., 2006, Bosch et al., 2008, Bernard et al., 2010, Ekström et al., 2011).

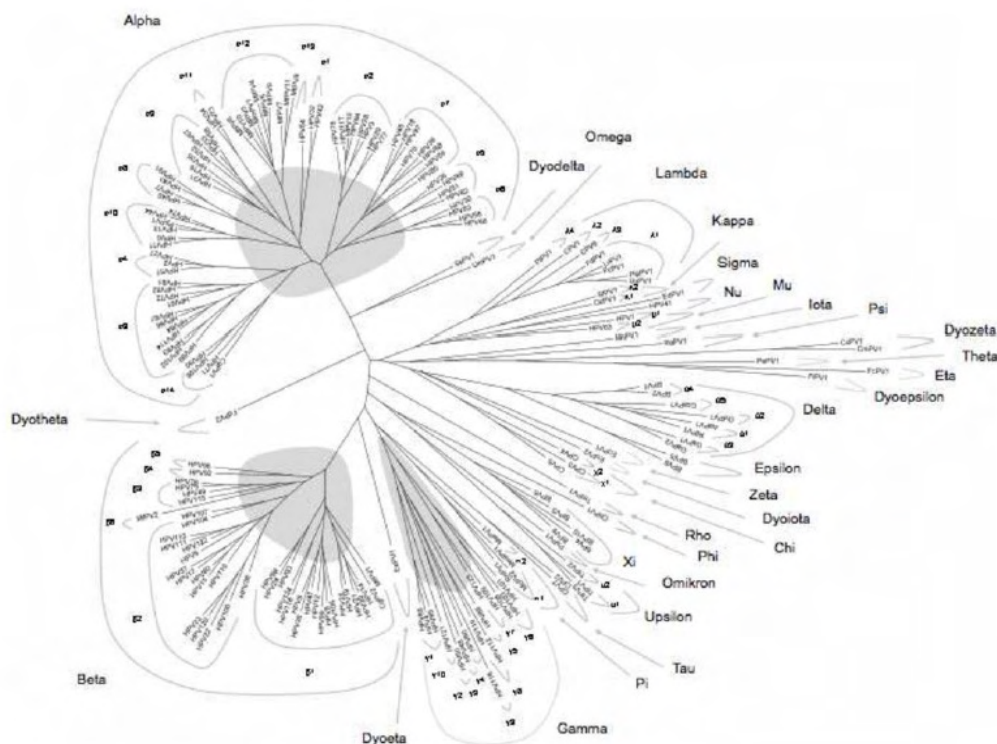
Έτσι δημιουργήθηκαν τα γένη Άλφα-, Βήτα-, Γάμμα-, Μι- και Νι-PVs (Εικόνα 1.1). Κάθε γένος διαχωρίζεται περαιτέρω σε τύπους, υπότυπους και παραλλαγές (variants), βάσει της ομοιότητας της αλληλουχίας του καψιδιακού γονιδίου L1. Η αλληλουχία του γονιδίου L1 διαφέρει μεταξύ των διαφορετικών τύπων τουλάχιστον κατά 10%, ενώ μεταξύ υπότυπων και παραλλαγών 2% -10% και 2% αντίστοιχα (Bernard et al., 2006, 2010). Εκτός από τα γένη Άλφα-, Βήτα-, Γάμμα-, Μι- και Νι, που περιέχουν τύπους των PVs που έχουν απομονωθεί από τον άνθρωπο, δύο γένη (Ηττα- and Θήτα-PVs) έχουν απομονωθεί από πουλιά, ενώ τα υπόλοιπα εννέα γένη έχουν απομονωθεί από θηλαστικά. (de Villiers et al., 2004).



Εικόνα 1.1. Φυλογενετικό δέντρο 118 PVs που βασίζεται στην αλληλουχία του καψιδιακού γονιδίου L1. Το δέντρο κατασκευάστηκε με το πρόγραμμα Treeview και με την μέθοδο neighbor-joining. Στο δέντρο αυτό παρουσιάζονται όλα τα γένη και οι αντιπροσωπευτικοί τύποι των Ιών των Θηλωμάτων (de Villiers et al., 2004).

Η τελευταία επίσημη ταξινόμηση των γενών των Ιών των Θηλωμάτων τελείωσε με το γένος Πι (pi-PVs). Η περιγραφή όμως νέων PV γενών θα είχε ως συνέπεια να εξαντληθεί το ελληνικό αλφάβητο. Ο Bernard et al. (2010) πρότεινε να συνεχιστεί η χρήση του ελληνικού αλφάβητου για δεύτερη φορά χρησιμοποιώντας

την πρόθεση "δύο". Μελέτη που πραγματοποιήθηκε από την συγκεκριμένη ομάδα με αλληλουχίες όλων των PVs από την βιβλιογραφία και την τράπεζα δεδομένων GenBank οδήγησε στην δημιουργία νέων γενών από το Ρο (Rho-PVs) έως το Ωμέγα (Omega-PVs) και συνέχισαν με τους όρους "Δύο Δέλτα" (Dyodelta-PVs), "Δύο Έψιλον" (Dyoepsilon-PVs). Η συγκεκριμένη ομάδα απέφυγε να χρησιμοποιήσει την πρόθεση "δύο" στα γένη Άλφα-, Βήτα-, Γάμμα-, Μι- και Νι -PVs. Σήμερα, η οικογένεια των Papillomaviridae χωρίζεται σε είκοσι εννέα γένη, με πέντε γένη να έχουν απομονωθεί από τον άνθρωπο (Άλφα-, Βήτα-, Γάμμα-, Μι- και Νι -PVs), είκοσι γένη από άλλα θηλαστικά, τρία γένη από πτηνά και ένα γένος από ερπετά. Στην Εικόνα 1.2. παρουσιάζεται το φυλογενετικό δέντρο, το οποίο κατασκευάστηκε με βάση την νουκλεοτιδική αλληλουχία του L1 γονιδίου 189 PVs. Στο δέντρο αυτό παρατηρείται η σχέση μεταξύ των προηγούμενων και των νέων γενών, ενώ έχουν επισημανθεί τα τρία κύρια γένη που συμπεριλαμβάνουν την πλειοψηφία των HPV τύπων.



Εικόνα 1.2. Φυλογενετικό δέντρο που προήρθε από πολλαπλή ομοπαράθεση L1 αλληλουχιών 189 PV τύπων. Για την κατασκευή του δέντρου χρησιμοποιήθηκε το πρόγραμμα MrBayes v3.1.2 με 10.000.000 κύκλους για τον αλγόριθμο Markov chain Monte Carlo (MCMC). Το πρόγραμμα ModelTest v3.7 προσδιόρισε το καλύτερο εξελικτικό μοντέλο (Bernard et al., 2010).

1.3.1 Άλφα-Papilloma ιοί

Το γένος Άλφα (Alpha-Papillomaviruses) είναι το μεγαλύτερο γένος μεταξύ των HPV και διαχωρίζεται σε τύπους που προσβάλλουν το επιθήλιο του βλεννογόνου και σε τύπους που προσβάλλουν το επιθήλιο του δέρματος. Οι HPV τύποι που προσβάλλουν το επιθήλιο του βλεννογόνου διαχωρίζονται περεταίρω σε υψηλού κινδύνου (*high-risk*) και χαμηλού κινδύνου (*low-risk*) (Bernard et al., 2010). Οι δερματικοί Άλφα- HPV τύποι είναι επίσης χαμηλού κινδύνου και η κατηγορία αυτή συμπεριλαμβάνει τους HPV2 και HPV57, οι οποίοι προκαλούν κοινά θηλώματα και τους τύπους HPV3 και HPV10, οι οποίοι προκαλούν επίπεδα κονδυλώματα και σπάνια εμπλέκονται στην πρόκληση καρκίνου (Hsueh et al., 2009, Bernard et al., 2010). Οι HPV τύποι χαμηλού κινδύνου που προσβάλλουν το επιθήλιο του βλεννογόνου μπορούν επίσης να προκαλέσουν δερματικές αλλοιώσεις στο γεννητικό σύστημα, αλλά δεν προκαλούν νεοπλασία (Middleton et al., 2003).

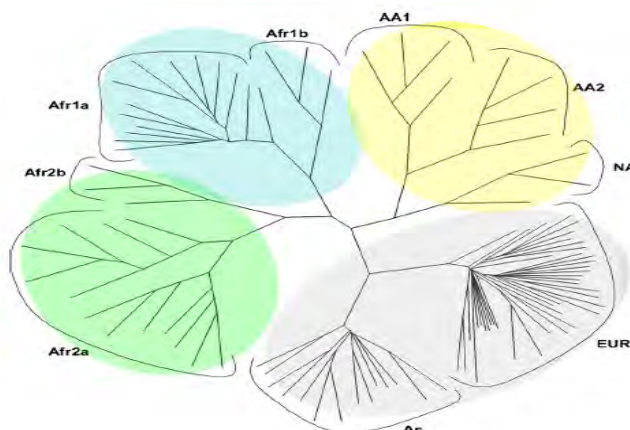
Οι καρκινικές αλλοιώσεις έχουν συσχετιστεί με HPV τύπους υψηλού κινδύνου (Bosch et al., 2008, zur Hausen, 2009). Δώδεκα HPV τύποι (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, και 59) έχουν καθοριστεί από την Παγκόσμια Οργάνωση Υγείας ως τύποι υψηλού κινδύνου που προκαλούν καρκίνο, με τους επιπλέον τύπους (68, 73) να αναγνωρίζονται ως "δυνητικοί" παράγοντες πρόκλησης καρκίνου (Munoz et al., 2006, Schiffman et al., 2009). Ορισμένοι επιπλέον τύποι συμπεριλαμβάνονται στην ομάδα υψηλού κινδύνου με βάση την εξελικτική τους ομοιότητα με τύπους που προκαλούν καρκίνο (Munoz et al., 2003, Schiffman et al., 2005).

1.3.1.1 Ταξινόμηση του HPV16

Ο HPV16 ανήκει στο γένος των Άλφα PVs και ευθύνεται για περισσότερο από το 50% των περιπτώσεων καρκίνου του τραχήλου της μήτρας (de Sanjose et al., 2010, Li et al., 2011). Εκτεταμένη μελέτη της αλληλουχίας του γονιδιώματος του HPV16 και σε μικρότερο βαθμό του HPV18 οδήγησε στο συμπέρασμα, ότι συγκεκριμένες ενδοτυπικές παραλλαγές του ιού (variants) σχετίζονται με την γεωγραφική προέλευση και την εθνικότητα του πληθυσμού όπου επικρατεί (Chan et al., 1992, Yamada et al., 1995). Πιο συγκεκριμένα το γονιδίωμα του HPV16 ταξινομείται σε πέντε μεγάλες διαφορετικές ενδοτυπικές γενιές παραλλαγών γνωστές ως HPV16 African type I, African type II, Asian-American, European και Asian. Οι

γενιές αυτές ομαδοποιούνται σε τέσσερα διαφορετικά φυλογενετικά κλαδιά (Yamada et al., 1995, 1997, Bernard et al., 2006). Η ανάλυση της αλληλουχίας του γονιδιώματος του HPV16, οδήγησε στον εντοπισμό ενδοτυπικών νουκλεοτιδικών πολυμορφισμών σε συγκεκριμένες περιοχές (L1, L2 E6, E7, E2 LCR), γεγονός που επέκτεινε την φυλογενετική του ταξινόμηση (Yamada et al., 1995, 1997). Πρόσφατες μελέτες κατάφεραν να ταξινομήσουν τις γενιές των ενδοτυπικών παραλλαγών του HPV16 σε μεγαλύτερο βαθμό και να δημιουργήσουν τις "υπογενιές" (sublineages). Έτσι έχει προκύψει: (i) η European-Asian γενιά η οποία περιλαμβάνει τις υπογενιές European και Asian, (ii) η African 1 (AFR1) γενιά που περιλαμβάνει δυο υπογενιές AFR1a και AFR1b, (iii) η African 2 (AFR2) γενιά που περιλαμβάνει τις υπογενιές AFR2a και AFR2b, και (iv) η Asian American/North American γενιά, η οποία περιλαμβάνει τις υπογενιές North American (NA), Asian-American-1 (AA1) και Asian-American-2 (AA2). Οι "υπογενιές" των ενδοτυπικών παραλλαγών του HPV16 παρουσιάζουν μεταξύ τους διαφορές της τάξης 0.5% έως 1% σε επίπεδο γονιδιώματος (Smith et al., 2011, Chen et al., 2011, Huertas-Salgado et al., 2011, Cornet et al., 2012).

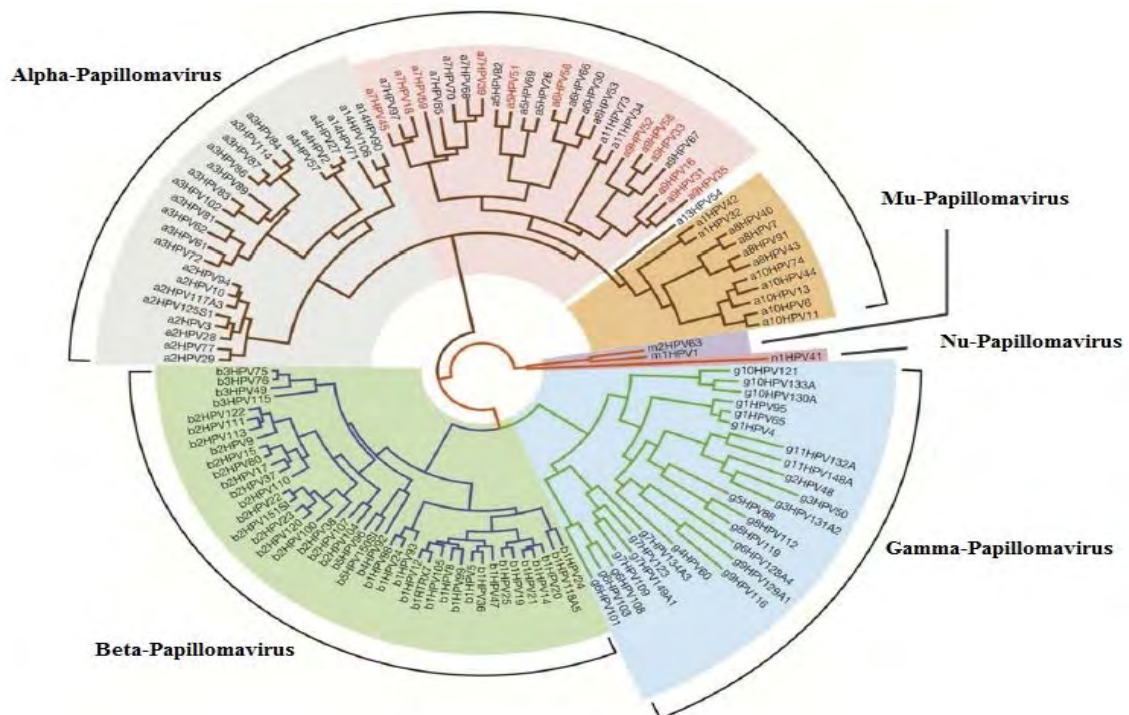
Η μελέτη των ενδοτυπικών παραλλαγών του ιού παρουσιάζει μεγάλο ερευνητικό ενδιαφέρον καθώς μελέτες έχουν προτείνει, ότι οι μη Ευρωπαϊκές ενδοτυπικές παραλλαγές του HPV16 συνδέονται με μεγαλύτερη συχνότητα με υψηλότερου βαθμού δυσπλασίες (Tornesello et al., 2011, Sichero et al., 2007).



Εικόνα 1.3. Φυλογενετικό δέντρο (UMPGA, με 100 bootstrap values) που κατασκευάστηκε με την χρήση 99 αλληλουχιών της ρυθμιστικής περιοχής LCR και του E6 γονιδίου. Το δέντρο αυτό περιγράφει τον διαχωρισμό των HPV16 παραλλαγών σε τέσσερα φυλογενετικά κλαδιά με τις αντίστοιχες υπογενιές τους (Cornet et al., 2012).

1.3.2 Βήτα, Γάμμα, Μι και Νι- Papilloma ιοί

Οι HPV τύποι που προσβάλλουν το δέρμα χωρίζονται σε τέσσερα διαφορετικά γένη (Βήτα, Γάμμα, Μι και Νι -papillomaviruses). Οι ιοί αυτοί συχνά εντοπίζονται σε δείγματα υγιούς δέρματος, αποδεικνύοντας ότι προκαλούν ασυμπτωματικές μολύνσεις. Τα μέλη του γένους Beta προσβάλλουν το επιθήλιο του δέρματος χωρίς να προκαλούν αλλοιώσεις. Παρόλα αυτά έχει παρατηρηθεί, ότι η μόλυνση από τα στελέχη του συγκεκριμένου γένους σε ανοσοκατεσταλμένους ασθενείς και ασθενείς που πάσχουν από επιδερμοδυσπλασία (epidermodysplasia verruciformis), μπορούν να προκαλέσουν μη μελανωματικό καρκίνο του δέρματος (Pfister, 2003). Οι HPV τύποι που προέρχονται από τα γένη (Γάμμα, Μι και Νι), προκαλούν δερματικά θηλώματα τα οποία δεν εξελίσσονται ποτέ σε καρκίνο. Στο παρακάτω φυλογενετικό δέντρο (Εικόνα 1.4) παρουσιάζεται η φυλογενετική σχέση μεταξύ των HPV γενών (Doorbar, 2012).



Εικόνα 1.4. Εξελικτική σχέση μεταξύ των HPV. Οι HPV αποτελούνται από πέντε γένη με διαφορετικό τροπισμό επιθηλίου και διαφορετική συσχέτιση με την ανάπτυξη ενδοεπιθηλιακών αλλοιώσεων. Οι Άλφα PVs περιλαμβάνουν τους τύπους χαμηλού κινδύνου (πορτοκαλί πλαίσιο) και τους τύπους υψηλής κινδύνου (ροζ πλαίσιο). Οι δερματικοί HPV τύποι γκρι (Alpha), πράσινο (Beta), μπλε (Gamma) δεν σχετίζονται με την ανάπτυξη καρκίνου. Η εικόνα αυτή δείχνει ένα φυλογενετικό δέντρο το οποίο κατασκευάστηκε με την μέθοδο maximum likelihood με την χρήση των γονιδίων E1, E2, L1, L2 συνολικά 132 HPV. Η ομοπαράθεση των αλληλουχιών πραγματοποιήθηκε με τον αλγόριθμο MUSCLE. Το φυλογενετικό δέντρο κατασκευάστηκε με το πρόγραμμα RaxML. (Doorbar et al., 2012).

1.4 Εξέλιξη των Ανθρώπινων Papilloma ιών

Κατά την διάρκεια της εξέλιξης οι HPV έχουν αναπτύξει συγκεκριμένους επιθηλιακούς τροπισμούς, με διαφορετικούς τύπους να παρουσιάζουν διαφορετική συσχέτιση με την πρόκληση αλλοιώσεων (Gottschling et al., 2007, Shah et al., 2010, Chen et al., 2011). Μελέτες έχουν δείξει ότι πολλοί HPV τύποι που ανήκουν στα γένη Βήτα και Γάμμα προκαλούν μόνο ασυμπτωματικές μολύνσεις σε ανοσολογικά υγιείς ασθενείς (Nindl et al., 2007, Gottschling et al., 2009, Ekstrom et al., 2010). Τέτοιοι ιοί είναι καλά προσαρμοσμένοι στους ξενιστές τους και σε πολλές περιπτώσεις μπορούν να ολοκληρώσουν το κύκλο ζωής τους και να παραμείνουν στον πληθυσμό χωρίς να προκαλούν κάποια ασθένεια (Forslund et al., 2007, Ekström et al., 2011). Αυτά τα χαρακτηριστικά δείχνουν, ότι η αλληλεπίδραση των PVs με τους ξενιστές τους είναι πολύ παλιά και έχει δημιουργηθεί μια ισορροπία μεταξύ του κύκλου ζωής του ιού και της ανοσολογικής απόκρισης του ξενιστή (Woolhouse, 2007). Η εξελικτική καταγωγή των PVs εντοπίζεται από την αρχή της εμφάνισης των αμινωτών, περίπου 350 εκατομμύρια χρόνια πριν. Πολλοί εξελικτικοί μηχανισμοί φαίνεται να έχουν συμβάλει στην ποικιλότητα των PVs. Οι μηχανισμοί αυτοί περιλαμβάνουν, τη συν-εξέλιξη ξενιστή και ιού, τον ανασυνδυασμό και την επιλογή νέου ξενιστή. Η επιλογή νέου ξενιστή περιλαμβάνει την επιλογή επιπρόσθετων ξενιστών από έναν ιό ή ακόμα και την εξαφάνιση ολόκληρης γενιάς ιών από τους ξενιστές τους (Gottschling et al., 2007, 2011).

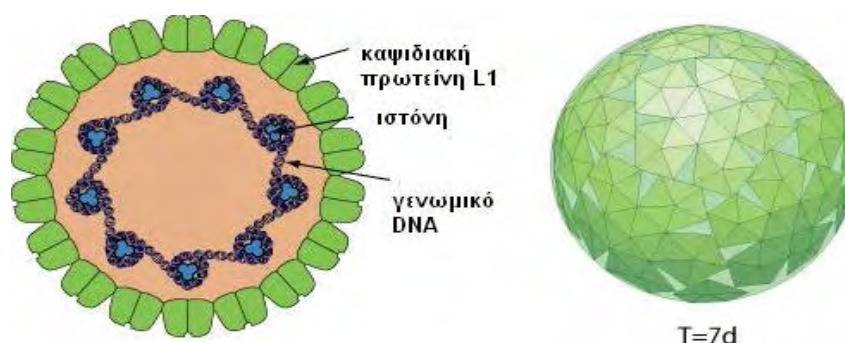
Πιο συγκεκριμένα, φυλογενετικές διαφορές μεταξύ πρώιμων και όψιμων γονιδίων έχουν βρεθεί σε τύπους που εμπλέκονται στην καρκινογένεση και ανήκουν στο γένος Άλφα (García-Vallné et al., 2005). Μελέτες έχουν δείξει ότι η εξελικτική αυτή ανακολουθία μεταξύ των γονιδίων μπορεί να προέρχεται από μεμονωμένα περιστατικά που συνέβησαν στο παρελθόν, όπως ανασυνδυασμός, επιλογή νέων ξενιστών, ή/και από έντονη εξελικτική πίεση που ασκείται στα γονίδια των ιών (Narechania et al., 2005, Varsani et al., 2006). Προηγούμενες μελέτες οι οποίες χρησιμοποίησαν αλληλουχίες ιών που ανήκουν στο γένος Άλφα από την βάση δεδομένων "Los Alamos National Laboratory" ανέφεραν την ύπαρξη ένδειξης σήματος ανασυνδυασμού στα γονίδια E6, E7, L2 και L1 (Angulo et al., 2007, Carvajal-Rodríguez et al., 2008). Παρόλα αυτά φυλογενετικές διαφορές μεταξύ πρώιμων και όψιμων γονιδίων δεν έχουν εντοπιστεί στους Βήτα-PVs (Gottschling et al., 2007), η οποία αποτελεί σημαντική και ετερογενή ομάδα ιών.

Τα τελευταία χρόνια φυλογενετικές και στατιστικές μελέτες που έχουν ως στόχο την μελέτη της εξέλιξης του γονιδιώματος των HPV, έχουν επικεντρωθεί στην ύπαρξη και την ανάλυση ενδοτυπικών ανασυνδυασμών του HPV16. Πιο συγκεκριμένα ο Carvajal-Rodriguez και συνεργάτες, χρησιμοποιώντας αλληλουχίες από την βάση δεδομένων GenBank απέδειξε ότι υπάρχει ένδειξη ενδοτυπικού ανασυνδυασμού εντός του γονιδίου E7 του HPV16 (Carvajal-Rodríguez et al., 2008). Παρόλα αυτά, ο πρώτος ενδοτυπικός ανασυνδυασμός σε φυσική μόλυνση περιγράφηκε το 2009 στο γονιδίωμα του HPV16, μεταξύ της Ευρωπαϊκής ενδοτυπικής παραλλαγής και της Αφρικανικής τύπου II παραλλαγής (Jiang et al. 2009). Επομένως ο ανασυνδυασμός μεταξύ διαφορετικών παραλλαγών του ίδιου τύπου ίσως να μπορεί να δώσει χρήσιμες πληροφορίες για τον μηχανισμό εξέλιξης του γονιδιώματος των HPV.

1.5. Δομή των Papilloma ιών του ανθρώπου

1.5.1 Δομή καψιδίου

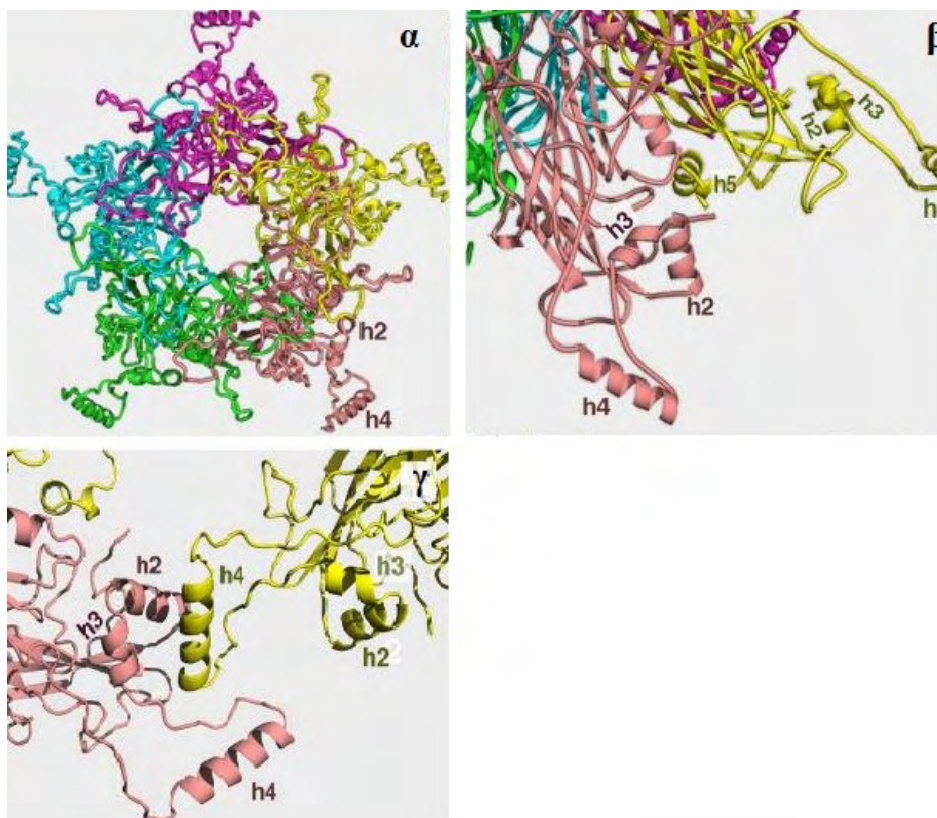
Το ικό σωματίδιο των HPV στην εξωτερική του επιφάνεια αποτελείται από 72 πενταμερή της καψιδιακής πρωτεΐνης L1 και σχηματίζει T=7 εικοσαεδρικά καψίδια (Kimbauer et al., et al 1992). Η καψιδιακή πρωτεΐνη L1 είναι ο παράγοντας, ο οποίος συμβάλει στην αρχική αλληλεπίδραση του ιού με τον ξενιστή του και στην αναγνώριση των ικών σωματιδίων από το ανοσολογικό σύστημα του προσβεβλημένου ατόμου (Yang R et al., 2004, Johnson et al., 2009). Η καψιδιακή πρωτεΐνη L2 τοποθετείται εσωτερικά του σκελετού της L1 και εξέρχεται του καψιδίου κατά την διάρκεια της εισόδου του ιού στο κύτταρο (Day et al., 2004, 2008). Μελέτες έχουν δείξει ότι η πρωτεΐνη L2 με την βοήθεια μιας υδρόφοβης περιοχής κοντά στο C-τελικό άκρο της, συνδέεται με την L1. Μέσω αυτής της περιοχής η L2 εισέρχεται στο κέντρο του L1 πενταμερούς (Finnen., et al 2003). Βασική της λειτουργία είναι η συμβολή της μαζί με την ική πρωτεΐνη E2 στο πακετάρισμα του ικού DNA στα νεοσχηματιζόμενα ικά καψίδια ενώ έχει προταθεί ότι συμμετέχει σε αλλαγές στην διαμόρφωση των ικών σωματιδίων κατά την πρόσδεση τους στον ξενιστή και στην καταστροφή των ενδοσωμικών μεμβρανών με σκοπό την απελευθέρωση των ικών σωματιδίων (Day et al., 2004, Buck et al 2004, Bossis et al., 2005, Kämper et al., 2006, Richards et al., 2006). Οι κυτταρικές ιστόνες συμπυκνώνουν το γονιδίωμα του ιού σε ένα δίκλωνο κυκλικό DNA "μίνι-χρωμόσωμα" εντός του καψιδίου L1-L2. (Howley 1996).



Εικόνα 1.5. T=7 HPV16, αλληλεπίδραση ικού DNA με ιστόνες (Swiss institute of Bioinformatics).

Η κρυσταλλική δομή της μικρότερης δομικής μονάδας, T=1 περιέχει δώδεκα L1 πενταμερή (Modis et al., 2002). Η υψηλή ανάλυση της κρυσταλλικής δομής της T=1 δομής

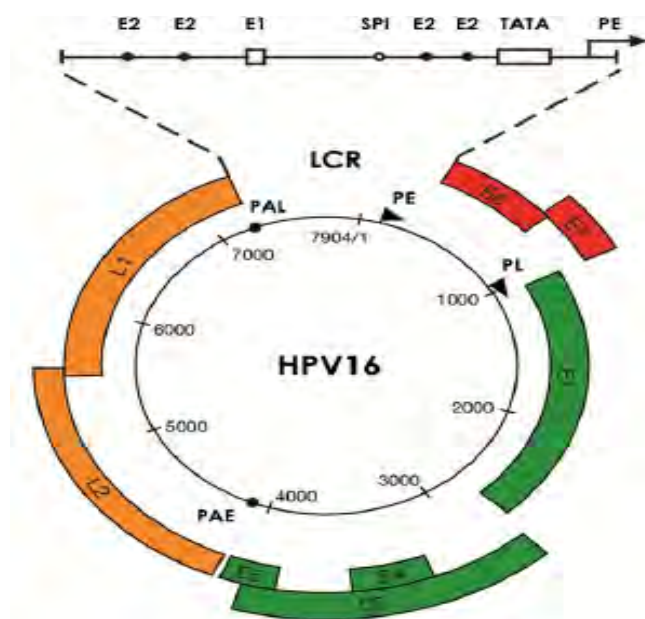
αποκαλύπτει την L1 πενταμερή διαμόρφωση καθώς επίσης και την αλληλεπίδραση μεταξύ των πενταμερών, που αποτελεί βασική αρχή για την δημιουργία της T=1 δομής. Το L1 πενταμερές περιέχει πέντε θηλιές που μοιάζουν με πλαϊνές προεξοχές, οι οποίες αποτελούνται από 100 C-τελικά άκρα. (Εικόνα 1.6). Κάθε προεξέχουσα θηλιά αποτελείται από μια α -έλικα, η οποία είναι 'αγκιστρωμένη' στον δομικό πυρήνα του L1 πενταμερούς δια μέσου δύο ελίκων, την έλικα 2 και την έλικα 3, (Εικόνα 1.6). Τα πενταμερή είναι συνδεδεμένα μεταξύ τους με ισχυρούς υδρόφοβους δεσμούς μεταξύ της έλικας 4 του ενός πενταμερούς με την έλικα 2 και 3 του γειτονικού πενταμερούς. Τα υπόλοιπα C-τελικά άκρα επιστρέφουν στον πυρήνα της L1 για να σχηματίσουν την έλικα 5 (Εικόνα 1.6) (Bishop et al., 2007).



Εικόνα 1.6. Δομικά στοιχεία της L1 για τον σχηματισμό του καψιδίου. α) Όψη ενός L1 πενταμερούς. Απεικονίζονται οι πλευρικές προεξοχές από κάθε ένα από τα πέντε μονομερή (το κάθε ένα με διαφορετικό χρώμα). β) Πιο κοντινή όψη των πλευρικών προεξοχών δύο μονομερών δείχνοντας την ακριβή θέση των ελίκων- 2, 3, 4, 5, και του C-τελικού άκρου της L. γ) Η δια-πενταμερής έλικα-έλικα αλληλεπίδραση μεταξύ δύο πενταμερών στο T=1. Η έλικα 4 του ενός πενταμερούς (κίτρινο) αλληλεπιδρά με την έλικα 2 και 4 άλλου πενταμερούς. (Bishop et al., 2007).

1.6 Οργάνωση του γονιδιώματος

Η οργάνωση του γονιδιώματος είναι καλά συντηρημένη μεταξύ των ιών των Θηλωμάτων. Το γονιδίωμα του HPV16 ακολουθεί την τυπική οργάνωση των υψηλού κινδύνου Άλφα PVs. Όπως φαίνεται στην Εικόνα 1.7 περιλαμβάνει 6 μη δομικά ικά γονίδια E1, E2, E4, E5, E6, E7 και δύο δομικά γονίδια L1, L2, τα οποία εκφράζονται σε διαφορετικά στάδια του κύκλου ζωής του ιού. Τα γονίδια αυτά εκφράζουν ένα μεγαλύτερο αριθμό προϊόντων ως αποτέλεσμα του ματίσματος του ιικού mRNA. (Zheng, 2006) Το γονιδίωμα του HPV16 περιέχει οχτώ ανοιχτά πλαίσια ανάγνωσης (ORFs) και διαχωρίζεται σε τρεις τομείς, με βάση τις λειτουργικές του δραστηριότητες. Έτσι έχουμε την πρόιμη περιοχή (E= Early region), την όψιμη περιοχή (L=late region) και την ρυθμιστική περιοχή LCR. Η πρόιμη περιοχή κωδικοποιεί για τις πρόιμες ικές ρυθμιστικές πρωτεΐνες E1, E2, E4, E5, E6 και E7 που επηρεάζουν την αντιγραφή και την έκφραση του DNA του ιού. Η όψιμη περιοχή κωδικοποιεί για τις όψιμες δομικές πρωτεΐνες L1 και L2 που σχηματίζουν το ικό καψίδιο (Danos et al., 1982).



Εικόνα 1.7. Οργάνωση του γονιδιώματος του HPV16

Η έκφραση του γονιδιώματος του HPV16 ρυθμίζεται από δύο βασικούς υποκινητές. Τον PE υποκινητή (αποκαλούμενος επίσης και ως p97 υποκινητής) που βρίσκεται στην LCR και είναι υπεύθυνος για την έκφραση των πρόιμων γονιδίων και τον PL υποκινητή (γνωστός και ως p670), ο οποίος βρίσκεται στο E7 γονίδιο και

ρυθμίζει την έκφραση των όψιμων γονιδίων και του E1 γονιδίου (Grassmann et al., 1996, Hummel et al., 1992). Όπως φαίνεται στην Εικόνα 1.7 οι περιοχές που επισημαίνονται ως PAE και PAL δηλώνουν τις πρώιμες και όψιμες θέσεις πολυαδενιλίωσης εντός του γονιδιώματος του ιού αντίστοιχα. Οι θέσεις 97 και 670 αποτελούν την θέση της 5' καλύπτρας των ικών μεταγράφων (Doorbar et al., 2012).

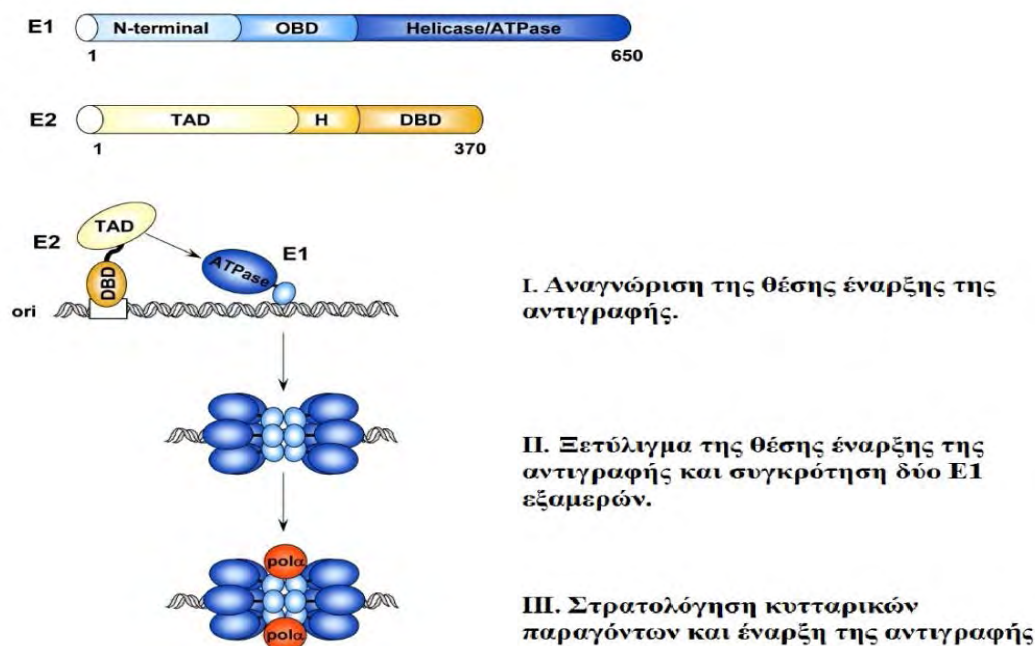
Τέλος, η ρυθμιστική περιοχή LCR έχει μέγεθος 850 bp και δεν περιέχει στην αλληλουχία της ανοιχτά πλαίσια ανάγνωσης. Η περιοχή αυτή εντοπίζεται μεταξύ του τέλους του γονιδίου L1 και στην αρχή του γονιδίου E6. Ακόμη η LCR περιέχει την θέση έναρξης της αντιγραφής και θέσεις πρόσδεσης μεταγραφικών παραγόντων (π.χ. SP1, AP1, Oct1) για την έκφραση των ικών γονιδίων (Hebner, 2006).

1.6.1 Η ική πρωτεΐνη E1

Το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου αποτελεί μια από τις πιο καλά συντηρημένες περιοχές του γονιδιώματος των PVs (García-Vallné, 2005). Η πρωτεΐνη E1 είναι μια πολυλειτουργική πυρηνική φωσφοπρωτεΐνη με χαμηλό επίπεδο έκφρασης. Έχει μέγεθος 70-80 kDa, αποτελείται περίπου από 650 αμινοξέα και η παρουσία της είναι απαραίτητη για την αντιγραφή του ικού γονιδιώματος. Η E1 πρωτεΐνη εκφράζεται στην πρώιμη φάση του κύκλου ζωής του ιού με στόχο να διατηρήσει το ικό DNA την επισωμική του μορφή.

Δομικές και λειτουργικές μελέτες έχουν επικεντρωθεί στην ανάλυση της E1 πρωτεΐνης του HPV11 και έχουν δείξει ότι αποτελείται από τρεις λειτουργικές περιοχές (Amin et al., 2000, White et al., 2001, Titolo et al., 2000, 2003). α) Το C-τελικό άκρο, το οποίο έχει την δραστικότητα ATPασης και 3'-5' ελικάσης, συμβάλει στον ολιγομερισμό της πρωτεΐνης σε ένα εξαμερές και βοηθά στην αλληλεπίδραση της E1 πρωτεΐνης με την DNA πολυμεράση α του ξενιστή. β) Μια κεντρική περιοχή η οποία εμπλέκεται στην πρόσδεση της πρωτεΐνης με το DNA του ιού. γ) Το N-τελικό άκρο το οποίο αποτελεί ρυθμιστική περιοχή της πρωτεΐνης καθώς διαθέτει συντηρημένες αλληλουχίες για πυρηνικό εντοπισμό (NLS), πυρηνική εξαγωγή (NES), ένα συντηρημένο μοτίβο πρόσδεσης των κυκλινών A, E (CBM) που αλληλεπιδρούν με την Cdk2, καθώς επίσης πολλές θέσεις φωσφορυλίωσης κινασών (Εικόνα 1.8) (Ma et al., 1999, Deng et al., 2004, Fradet-Turcotte et al., 2010).

Σαν πρωτεΐνη έναρξης της αντιγραφής η E1 έχει δύο βασικούς ρόλους. Πρώτον δρα σαν πρωτεΐνη πρόσδεσης, αναγνωρίζοντας την θέση έναρξης της αντιγραφής του DNA και δεύτερον δρα σαν ελικάση με αποτέλεσμα να ξεδιπλώνει την θέση έναρξης της αντιγραφής και να προχωρά η αντιγραφική διχάλα (Liu , 1995). Πιο συγκεκριμένα παρουσία ATP η E1 προσδέεται στην LCR σε μία περιοχή πλούσια σε A-T εντός της θέσης έναρξης της αντιγραφής του ιού με την βοήθεια της πρωτεΐνης E2 (Titolo S. et al 2003). Το σύμπλοκο E1-E2 αποτελείται από ένα E2 διμερές και δύο μόρια E1. Το σύμπλοκο αυτό δρα σαν δομική μονάδα για την συγκρότηση ενός μεγαλύτερου E1 εξαμερούς συμπλόκου (Εικόνα 1.8). Με την βοήθεια της E2 πρωτεΐνης στρατολογούνται επιπρόσθετα μόρια E1 και σχηματίζεται ένα διπλό E1 εξαμερές. (Sedman et al., 1998, D’Abramo and Archambaul, 2011). Το σύμπλοκο αυτό προσδέεται στην θέση έναρξης της αντιγραφής με υψηλή συγγένεια, ξεκινάει το ξετύλιγμα της διπλής έλικας του DNA και προάγει την αντιγραφή του ιού με την βοήθεια της ATP-εξαρτώμενης δραστηριότητας ελικάσης της E1 πρωτεΐνης (Εικόνα 1.8). Επιπρόσθετα, το C-τελικό άκρο της πρωτεΐνης αλληλεπιδρά με την DNA πολυμεράση α-πριμάση, την τοποισομεράση I και την πρωτεΐνη RPA με στόχο να στρατολογηθούν οι κυτταρικοί αντιγραφικοί παράγοντες στην θέση έναρξης της αντιγραφής του ιού για να δημιουργηθεί ένα ενεργό αντιγραφικό σύμπλοκο (Εικόνα 1.8). (Masterson et al., 1998, Conger et al., 1999, Parker et al., 2000, Loo et al., 2004, Clower et al., 2006).



Εικόνα 1.8. Σχηματική αναπαράσταση της πρωτεΐνης E1 και E2 που απαιτείται για την αντιγραφή του HPV γονιδιώματος. Η διαδικασία έναρξης της αντιγραφής του ιικού DNA περιλαμβάνει τα εξής βήματα. I) Η αντιγραφή ξεκινά με την στρατολόγηση των E1 (μπλε) και E2 (κίτρινο) πρωτεϊνών στην θέση έναρξης της αντιγραφής. Η στρατολόγηση απαιτεί την πρωτεϊνική αλληλεπίδραση μεταξύ της περιοχής TAD της E2 πρωτεΐνης και της περιοχής της ελικάσης της E1 πρωτεΐνης. II) Η E2 πρωτεΐνη στρατολογεί επιπλέον E1 μόρια και προάγει την συναρμολόγηση τους σε δυο E1 εξαμερή καθιστώντας το σύμπλοκο αυτό μια λειτουργική ελικάση. Το ATP επίσης διεγείρει τον ολιγομερισμό της E1 πρωτεΐνης και θα χρησιμοποιηθεί για την ενεργότητα της ελικάσης. III) Τέλος το E1 σύμπλοκο αλληλεπιδρά με κυτταρικούς αντιγραφικούς παράγοντες όπως η DNA πολυμεράση α-πριμάση (πορτοκαλί) με στόχο να προάγει την αντιγραφή του ιικού γονιδιώματος (D' Abramo and Archambaul, 2011).

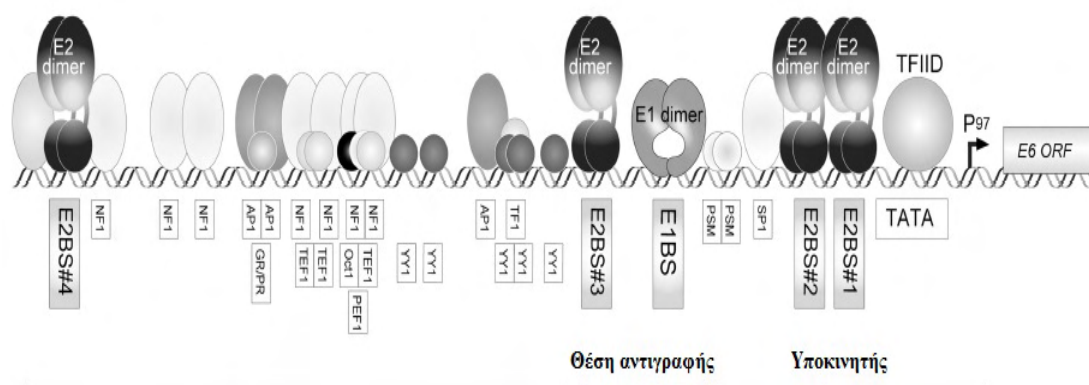
1.6.2 Η ιική πρωτεΐνη E2

Η E2 πρωτεΐνη είναι μια πολυλειτουργική πρωτεΐνη μεγέθους 50 kDa και μήκους περίπου 370 αμινοξέων. Προσδένεται ως διμερές σε ειδικές θέσεις στην LCR και προάγει την αντιγραφή του ιού, ρυθμίζει την μεταγραφή των ιικών γονιδίων και κατευθύνει την κατανομή του επισωματικού γονιδιώματος του ιού στα θυγατρικά κύτταρα στην φάση της μίτωσης (Hedge, 2002, Blachon, 2003, McBride et al., 2004, Pves et al., 2006). Η E2 πρωτεΐνη είναι οργανωμένη σε δύο λειτουργικές περιοχές: α) Τη N-τελική περιοχή διενεργοποίησης (Transactivation domain, TAD), η οποία εμπλέκεται στην ρύθμιση της μεταγραφής των ιικών γονιδίων και στην αλληλεπίδραση με την E1 πρωτεΐνη, β) μία C-τελική περιοχή, η οποία συμβάλει στην πρόσδεση της E2 πρωτεΐνης στο DNA του ιού και βοηθά στον διμερισμό της E1

πρωτεΐνης (DBD, DNA-binding/dimerization domain). Αυτές οι δύο περιοχές διαχωρίζονται μεταξύ τους από την περιοχή που ονομάζεται "αρμός" (hinge region), η οποία πιστεύεται ότι είναι εύκαμπτη περιοχή της πρωτεΐνης, αλλά η λειτουργία της ακόμα δεν έχει αποσαφηνιστεί πλήρως (Εικόνα 1.8) (Hedge, 2002).

Η E2 πρωτεΐνη εκτός από τον ρόλο της στην έναρξη της αντιγραφής του ιού, επιδρά στην ρύθμιση της μεταγραφής δρώντας ως ενεργοποιητής ή καταστολέας. Η E2 προσδένεται σε τέσσερις θέσεις της συντηρημένης αλληλουχίας ACCGN₄CGGT στην ρυθμιστική περιοχή LCR (Bouvard. et al., 1994). Οι περιοχές αυτές ρυθμίζουν την μεταγραφή των γονιδίων E6 και E7 (Bouvard. et al., 1994). Όσον αφορά τον HPV16, δύο από αυτές τις περιοχές βρίσκονται στο TATA Box ακριβώς ανοδικά του υποκινητή p97 του γονιδίου E6 και συνορεύουν στο 5' άκρο με μια θέση πρόσδεσης του μεταγραφικού παράγοντα Sp1 και στο 3' άκρο με την πρωτεΐνη πρόσδεσης στο TATA box (TBP) (Εικόνα 1.9). Η πρόσδεση της E2 σε αυτές τις περιοχές κοντά στον υποκινητή εμποδίζει την πρόσδεση του Sp1 και του TBP στις αντίστοιχες θέσεις τους και επάγει την καταστολή της μεταγραφής (Dostatni et al 1991, Tan et al., 2002). Οι άλλες δύο θέσεις εντοπίζονται πιο ανοδικά του υποκινητή (Εικόνα 1.9). Η πρόσδεση της E2 στις περιοχές αυτές οδηγεί στην ενεργοποίηση της μεταγραφής (Ham et al., 1994).

Η E2 παρουσιάζει διαφορετική συγγένεια για τις θέσεις πρόσδεσης και εξαρτάται από τη συγκέντρωση της E2 για τον αν θα καλυφθούν όλες οι θέσεις. Σε χαμηλές συγκεντρώσεις η E2 συνδέεται σε θέσεις που τοποθετούνται μακριά από τον υποκινητή του γονιδίου E6. Στην περίπτωση αυτή οι κυτταρικοί μεταγραφικοί παράγοντες μπορούν να προσδεθούν και να ξεκινήσει η έκφραση των γονιδίων E6 και E7. Όταν η συγκέντρωση της E2 είναι υψηλή καταλαμβάνονται και οι τέσσερις θέσεις πρόσδεσης της E2 και οι κυτταρικοί μεταγραφικοί παράγοντες, όπως ο Sp1 και TFIIID/B εκτοπίζονται από το TATA box. Σε αυτή την περίπτωση η E2 δρα ως μεταγραφικός καταστολέας και σταματά την έκφραση των ογκογονιδίων E6 και E7. (Steger and Corbach, 1997).



Εικόνα 1.9. Ρυθμιστική περιοχή LCR. Αναπαράσταση της θέσης πρόσδεσης μεταγραφικών παραγόντων και των θέσεων πρόσδεσης της E2 πρωτεΐνης (Kajitani, 2012).

Η E2 πρωτεΐνη έχει αποδειχτεί ότι συνδέεται με διάφορους κυτταρικούς μεταγραφικούς παράγοντες όπως ο TFIIIB, TBP, TFIIID, AMF-1/GPS, p300/CBP, NAP-1, P/CAF, και TopBP1 (Lee et al., 2000, Boner et al., 2002, Rehtanz et al., 2004). Πιο πρόσφατες μελέτες έχουν δείξει ότι η αλληλεπίδραση της E2 πρωτεΐνης με τον μεταγραφικό παράγοντα Brd4, επηρεάζει την δράση της E2 πρωτεΐνης ως μεταγραφικός ενεργοποιητής ή καταστολέας (Schweiger et al., 2006). Επιπλέον, βιοχημικές μελέτες απέδειξαν ότι η αλληλεπίδραση της πρωτεΐνης E2 με τον μεταγραφικό παράγοντα Brd4 παίζει σημαντικό ρόλο στην δράση της E2 πρωτεΐνης στην πρόσδεση του ιικού επισώματος στην μιτωτική χρωματίνη και πιθανώς στην μιτωτική άτρακτο του κυττάρου ξενιστή (Ives et al., 2006).

Λόγω του καθοριστικού ρόλου που διαδραματίζουν στην αντιγραφή του ιού και την ιική παθογένεια οι E1 και E2 πρωτεΐνες αποτελούν ελκυστικούς στόχους για την ανάπτυξη αντί-HPV φαρμάκων (Archambault, 2013).

1.6.3 Η ιική πρωτεΐνη E4

Το ανοιχτό πλαίσιο ανάγνωσης του E4 γονιδίου βρίσκεται εντός του ανοιχτού πλαισίου ανάγνωσης του E2 γονιδίου και διαφέρει σημαντικά σε μέγεθος μεταξύ διαφορετικών τύπων HPV. Η πρωτεΐνη E4 συντίθενται στην όψιμη φάση του κύκλου ζωής του HPV στις ανώτερες επιθηλιακές στοιβάδες και εκφράζεται σε μεγαλύτερες ποσότητες από όλες τις πρωτεΐνες του ιού. Στους HPV το προϊόν του πρώιμου γονιδίου E4 εκφράζεται από ένα ματισμένο μόριο mRNA του οποίου τα πέντε πρώτα αμινοξέα προέρχονται από το E1 γονιδίο (Middleton et al., 2003, Doorbar, 2013).

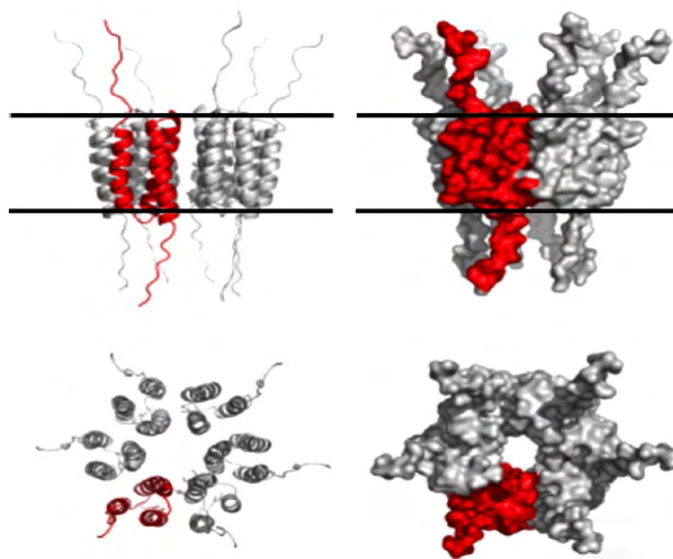
Η πρωτεΐνη E1^{E4} αλληλεπιδρά με το δίκτυο ινιδίων κερατίνης στα διαφοροποιημένα κερατινοκύτταρα επάγοντας την καταστροφή του δικτύου και την απελευθέρωση των νεοσχηματιζόμενων ικών σωματιδίων (McIntosh et al., 2010). Η πρωτεΐνη E1^{E4} του HPV16 διαχωρίζεται σε τρεις λειτουργικές περιοχές: α) Το N-τελικό άκρο με το συντηρημένο μοτίβο πλούσιο σε λευκίνη (LLXLL), είναι απαραίτητο για την αλληλεπίδραση με το δίκτυο ινιδίων κερατίνης, β) την κεντρική περιοχή πλούσια σε προλίνη, γ) την C-τελική περιοχή η οποία συμβάλει στην καταστροφή του δικτύου κερατίνης (Roberts et al., 1997).

Μια ακόμη λειτουργία της E1^{E4} πρωτεΐνης που έχει παρατηρηθεί στους τύπους HPV11 και HPV16 είναι η παύση του κυτταρικού κύκλου στη φάση G2 (G2 arrest) σε μία ποικιλία κυτταρικών σειρών (Davy et al., 2002, 2005, Roberts et al., 2008). Επιπλέον, έχει παρατηρηθεί ότι η E4 πρωτεΐνη αλληλεπιδρά και σταθεροποιεί την E2 πρωτεΐνη, ενώ η υπερέκφρασή της επηρεάζει την δραστικότητα της E2 πρωτεΐνης (Davy et al., 2009). Τέλος, η E1^{E4} πρωτεΐνη του HPV16 φαίνεται να έχει την δυνατότητα να επάγει την απόπτωση καθώς έχει μελετηθεί η σύνδεση και η αλληλεπίδραση της με τα μιτοχόνδρια (Raj et al., 2004).

1.6.4 Η ιική πρωτεΐνη E5

Η E5 πρωτεΐνη είναι μια μικρή πρωτεΐνη πλούσια σε υδρόφοβα αμινοξέα και εκφράζεται από ένα ματισμένο μόριο mRNA το οποίο ξεκινά από το E2 γονίδιο. Δεν έχει ενζυμική δράση αλλά δρα ρυθμίζοντας την δραστικότητα κυτταρικών πρωτεϊνών. Η E5 πρωτεΐνη κωδικοποιείται από πολλούς αλλά όχι από όλους τους τύπους HPV (Bernard et al., 2010). Οι Άλφα HPV τύποι υψηλού κινδύνου μεταξύ των οποίων και ο HPV16 κωδικοποιούν μια συντηρημένη E5 πρωτεΐνη μεγέθους περίπου 80 αμινοξέων (Bravo and Allonso, 2004, Di Mao and Petti, 2013). Αποτελεί μια διαμεμβρανική πρωτεΐνη με κυτταροπλασματικό C-τελικό άκρο, η οποία διέρχεται της μεμβράνης τρεις φορές (Εικόνα 1.10) (Nath et al., 2006, Wetherill et al., 2012). Τα διαμεμβρανικά τμήματα της πρωτεΐνης φαίνεται να αλληλεπιδρούν μεταξύ τους μέσω δισουλφιδικών δεσμών. Τέλος, η πρωτεΐνη ολιγομερίζεται σε ένα εξαμερές σχηματίζοντας μια δομή "δαχτυλιδιού" (Εικόνα 1.10) (Wetherill et al., 2012). Τα E5 εξαμερή παρουσιάζουν δραστικότητα pH-εξαρτώμενων καναλιών ιόντων στα λιποσώματα (Wetherill et al., 2012). Η πρωτεΐνη αυτή εντοπίζεται στις ενδοσωμικές μεμβράνες και στην συσκευή

Golgi και περιστασιακά βρίσκεται στις κυτταρικές μεμβράνες (Hu and Ceresa, 2009; Hu et al., 2009; Krawczyk et al., 2010).



Εικόνα 1.10. Μοντέλο ενός εξαμερούς καναλιού ιόντων που σχηματίζεται από την E5 πρωτεΐνη *in vitro*. Α. Παρουσιάζεται ένα μονομερές της E5 πρωτεΐνης με τρία διαμεμβρανικά τμήματα (κόκκινο χρώμα). Η θέση της μεμβράνης ορίζεται με τις οριζόντιες γραμμές. Η κάτω εικόνα δείχνει την κάτοψη της ίδιας αναπαράστασης (Wetherill et al., 2012).

Έχει παρατηρηθεί ότι η E5 πρωτεΐνη ενισχύει την ικανότητα των ογκογονιδίων E6 και E7 να προκαλούν αθανатоποίηση των πρώιμων ανθρώπινων κερατινοκυττάρων (Stopler et al., 1996) και αυξάνουν την κινητικότητα και την ικανότητα διήθησης των πρώιμων ανθρώπινων κερατινοκυττάρων (Kivi et al., 2008, Barbaresi et al., 2010). Έχει προταθεί ότι η E5 πρωτεΐνη συνδέεται με τον υποδοχέα του παράγοντα EGF (EGFR). Πολλές μελέτες έχουν δείξει ότι η HPV16 E5 πρωτεΐνη αυξάνει τα επίπεδα του EGFR στην επιφάνεια των κυττάρων ευαισθητοποιώντας τα κύτταρα στον EGF (Straight et al 1993). Επιπλέον διεγείρει τα μονοπάτια που ενεργοποιούνται από τον EGFR. Πιο συγκεκριμένα, προάγει την εκτεταμένη ενεργοποίηση ERK1/2 και AKT (Zhang et al., 2002). Διεγείρει την έκφραση των c-jun και c-fos ογκογονιδίων δρώντας στον ρυθμιστικό παράγοντα NF1 (Chen et al., 1996) και αυξάνει την έκφραση του παράγοντα COX-2 και VEGF (Kim et al., 2009). Παρόλα αυτά ο μοριακός μηχανισμός με τον οποίο η E5 επιδρά στα επίπεδα του EGFR παραμένει άγνωστος.

Τέλος, έχει προταθεί ότι η E5 μπορεί να εμπλέκεται στην ικανότητα του ιού να ξεφεύγει από την ανοσολογική απάντηση κρατώντας τον MHC τύπου I, στην συσκευή Golgi, εμποδίζοντας με αυτό τον τρόπο την μεταφορά του στην κυτταρική επιφάνεια (Ashrafi et al., 2005, 2006).

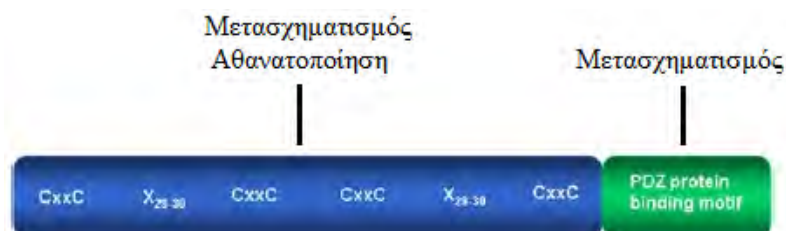
1.6.5 Η ιική πρωτεΐνη E6

Η E6 πρωτεΐνη των ιών υψηλού και χαμηλού κινδύνου είναι μία από τις πρώτες πρωτεΐνες που εκφράζονται κατά την διάρκεια HPV λοίμωξης (Kajitani, 2012). Αποτελείται περίπου από 150 αμινοξέα με μοριακό βάρος 16-18 kDa και περιλαμβάνει τέσσερις περιοχές σύνδεσης του ψευδαργύρου με το μοτίβο Cys-X-X-Cys (Εικόνα 1.11). Αυτό το χαρακτηριστικό είναι συντηρημένο μεταξύ των HPV και αποτελεί σημαντική λειτουργία για την δράση της πρωτεΐνης όπως, μετασχηματισμό των κυττάρων και σύνδεση με κυτταρικές πρωτεΐνες. Στην περίπτωση των HPV τύπων υψηλού κινδύνου το C-τελικό άκρο περιέχει ένα τομέα πρόσδεσης PSD-95/Dlg/ZO1 (PDZ), ο οποίος εμπλέκεται στην αλληλεπίδραση της E6 πρωτεΐνης με κυτταρικές πρωτεΐνες που διαθέτουν PDZ περιοχές (Εικόνα 1.11) (Nomine, 2006, Boulet, 2007).

Η πιο σημαντική λειτουργία της E6 πρωτεΐνης είναι η αλληλεπίδραση της με την ογκοκατασταλτική πρωτεΐνη p53. Η πρωτεΐνη p53 ρυθμίζει την έκφραση πρωτεϊνών που εμπλέκονται στην ρύθμιση του κυτταρικού κύκλου. Μια από αυτές τις πρωτεΐνες είναι ο αναστολέας της κινάσης της κυκλίνης p21. Όταν το κύτταρο βρεθεί σε συνθήκες στρες, όπως βλάβη στο DNA, η p53 επάγει την έκφραση της p21, η οποία με την σειρά της αναστέλλει τον κυτταρικό κύκλο και δίνει χρόνο στο κύτταρο να επιδιορθώσει την βλάβη. Εάν η βλάβη δεν επιδιορθωθεί τότε η p53 οδηγεί το κύτταρο σε απόπτωση (Artandi et al., 2005). Όσον αφορά την σύνδεση του μονοπατιού αυτού με τους HPVs έχει διαπιστωθεί ότι η E6 προσδέεται στην p53 μαζί με την E6 ουβικουτινική λιγάση E6AP. Ο σχηματισμός αυτός του συμπλόκου οδηγεί στην ουβικουτινίωση της p53 και την αποδόμησή της στο 26S πρωτεάσωμα. Έχει παρατηρηθεί ότι οι HPV υψηλού και χαμηλού κινδύνου επηρεάζουν την λειτουργία της p53, αλλά μόνο οι τύποι υψηλού κινδύνου διεγείρουν το μονοπάτι της ουβικουτινίωσης και την αποδόμησης της πρωτεΐνης (Pim et al., 2010, Fu et al., 2010, Zanier et al., 2012). Ακόμη η E6 μπορεί να αναστείλει την δράση της p53 έμμεσα μέσω σύνδεσης της με την p300/CBP, που είναι συνενεργοποιητής της p53 (Demarest et al., 2002).

Μια ακόμα λειτουργία της E6 σημαντική για την δημιουργία αθάνατων κυττάρων είναι η ενεργοποίηση της καταλυτικής υπομονάδας της τελομεράσης hTERT. Η E6 φαίνεται να ενεργοποιεί την έκφραση της hTERT μέσω συνδυαστικής δράσης των c-myc και Sp1. Αυτή η εκτεταμένη ενεργοποίηση της hTERT έχει σαν

αποτέλεσμα την διατήρηση μεγάλων τελομερών κατά την κυτταρική διαίρεση αναστέλλοντας την κυτταρική γήρανση με άμεση συνέπεια τα κύτταρα να γίνονται αθάνατα (Artandi et al., 2005).



Εικόνα 1.11. Σχηματική αναπαράσταση της E6 πρωτεΐνης. Περιλαμβάνει τέσσερις περιοχές σύνδεσης του ψευδαργύρου με το μοτίβο Cys-X-X-Cys (μπλε χρώμα), οι οποίες είναι απαραίτητες για την μεταγραφική ενεργοποίηση, μετασηματισμό, αθανατοποίηση και σύνδεση με κυτταρικές πρωτεΐνες. Το C- τελικό άκρο (πράσινο χρώμα) περιέχει ένα μοτίβο πρόσδεσης κυτταρικών πρωτεϊνών που διαθέτουν PDZ περιοχές (Boulet, 2007).

1.6.6 Η ιική πρωτεΐνη E7

Η E7 πρωτεΐνη είναι η δεύτερη ογκοπρωτεΐνη που σχετίζεται με την δημιουργία αθάνατων κυττάρων καθώς επίσης και με την ιική παθογένεια. Έχει μέγεθος 100 αμινοξέων και περιέχει τρεις διατηρημένες περιοχές. Την CR1, CR2 και την CR3. α) Η CR1 περιοχή περιλαμβάνει το N-τελικό άκρο το οποίο είναι απαραίτητο για τον κυτταρικό μετασηματισμό και την αποδόμηση της pRB, (β) η CR2 περιέχει ένα συντηρημένο μοτίβο LXCXE και μία θέση φωσφορλίωσης της κινάσης καζεΐνης II (CKII) (γ) τέλος η CR3 αποτελείται από δύο μοτίβα δακτύλων ψευδαργύρου και εμπλέκεται στην σύνδεση της E7 με την pRb και άλλες κυτταρικές πρωτεΐνες, ενώ συμβάλει στον διμερισμό της πρωτεΐνης (Εικόνα 1.12) (Hebner, 2006, Munger et al., 2004)

Βασική λειτουργία της E7 πρωτεΐνης είναι η αλληλεπίδραση της με την οικογένεια των πρωτεϊνών του ρετινοβλαστώματος (pRb, p107 και p103) (Roman, 2006, Barrow et al., 2010). Σε επίπεδο ρύθμισης του κυτταρικού κύκλου η αποφωσφορυλιωμένη μορφή της pRb δημιουργεί σύμπλοκα με τον μεταγραφικό παράγοντα E2F/DP1, ο οποίος αναστέλλει την έκφραση γονιδίων που συμβάλουν στην προώθηση του κυτταρικού κύκλου στη φάση S καθώς επίσης και στην προώθηση του κυττάρου στην απόπτωση. Κατά την μετάβαση από την φάση G1 στην φάση S τα σύμπλοκα κυκλίνης-κινάσης φωσφορυλιώνουν την pRb με αποτέλεσμα να

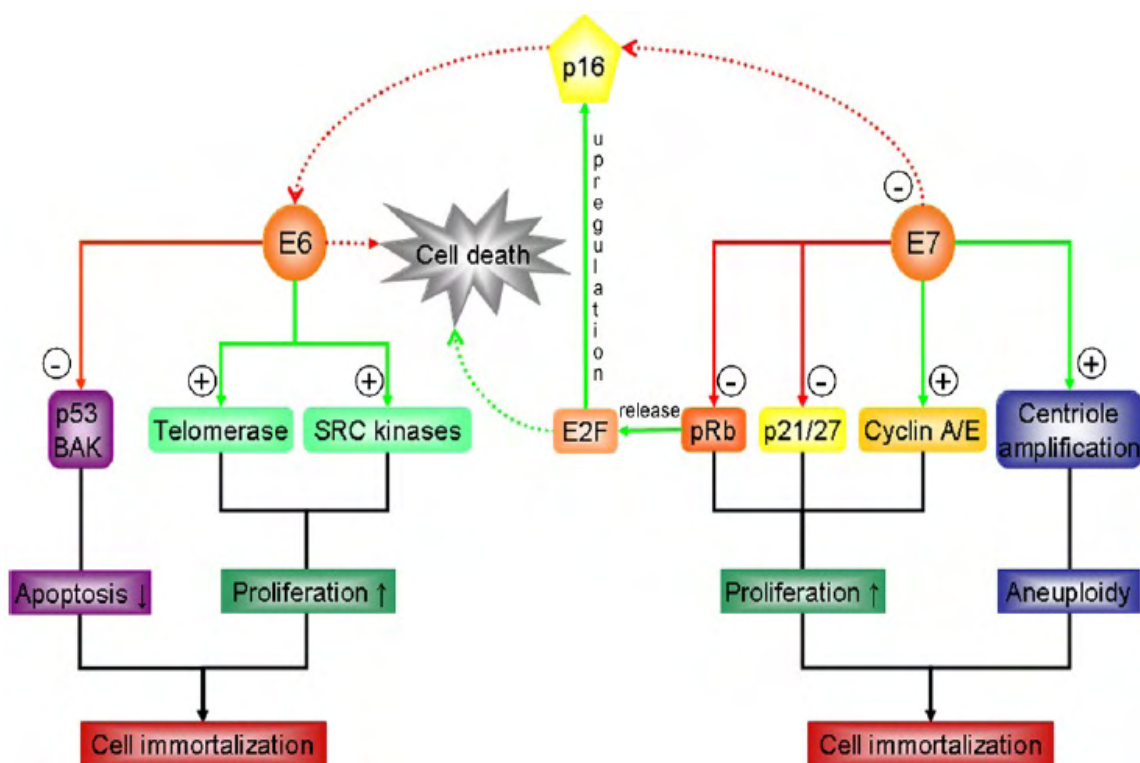
απελευθερώνεται η pRb από το σύμπλοκο που σχηματίζει με τον μεταγραφικό παράγοντα E2F με συνέπεια να επάγεται η μεταγραφή των γονιδίων που εμπλέκονται στην αντιγραφή του DNA. Όσον αφορά την δράση του ιού, η E7 συνδέεται με την pRb και επάγει την αποικοδόμηση της pRb μέσω του μονοπατιού του ουβικουτινικού πρωτεοσώματος. (Berezutskaya et al., 1997, Wang et al., 2001). Η πρόσδεση της E7 στην pRb έχει ως αποτέλεσμα την απορρύθμιση του κυτταρικού κύκλου με συνέπεια τον ανεξέλεγκτο κυτταρικό πολλαπλασιασμό.

Επιπλέον, η πρωτεΐνη E7 συνδέεται με τις κυκλίνες A και E καθώς επίσης με τους αναστολείς των κύκλινο-εξαρτώμενων κινασών (cdk) p21 και p27. Πιο συγκεκριμένα, η E7 αυξάνει τα επίπεδα των A και E κυκλινών και μπλοκάρει την δράση των p21, p27 με αποτέλεσμα να επάγεται η φωσφορυλίωση της pRb και η πρόοδος του κυτταρικού κύκλου (Funk et al., 1997, Nguyen, 2008). Επιπρόσθετη ομάδα πρωτεϊνών, που αλληλεπιδρά με την E7 είναι η απακετυλάση των ιστονών (HDACs). Σε φυσιολογικές συνθήκες η πρωτεΐνη pRb συνδέεται με τις HDACs και τις στρατολογεί στους E2F υποκινητές. Οι HDACs εκφράζονται σε όλους τους ιστούς και η βασική τους λειτουργία είναι η μετακίνηση των ακέτυλο ομάδων από τις ιστόνες. Οι HDACs μπορούν επιπρόσθετα να απακετυλιώσουν τον παράγοντα E2F αναστέλλοντας την λειτουργία του. Έχει παρατηρηθεί ότι η σύνδεση της E7 στην HDAC αναστέλλει την σύνδεση της pRb-HDAC με αποτέλεσμα να καταστέλλεται η λειτουργία της. Επίσης έχει αναφερθεί ότι η πρόσδεση της E7 στην HDAC έχει ως αποτέλεσμα την αναστολή της απακετυλίωσης του E2F. Αυτό έχει ως αποτέλεσμα την απορρύθμιση του ελέγχου σύνθεσης των HDAC με συνέπεια την εκτεταμένη αντιγραφή (Longworth, 2004).

Η E7 πρωτεΐνη των HPV τύπων υψηλού κινδύνου συμβάλει επιπλέον στην πρόκληση γενωμικής αστάθειας με το να προάγει την ενίσχυση των κεντρομερών στα κύτταρα της βασικής στοιβάδας του επιθηλίου. Ο ακριβής μηχανισμός που οδηγεί στο φαινόμενο αυτό δεν έχει πλήρως κατανοηθεί (Duensing et al., 2009, Korzeniewski et al., 2011). Βασική λειτουργία των E6 και E7 πρωτεϊνών στους περισσότερους τύπους HPV δεν είναι να προάγουν τον πολλαπλασιασμό των επιθηλιακών κυττάρων αλλά να διεγείρουν διαρκώς τον κυτταρικό κύκλο στα κύτταρα της μεσαίας επιθηλιακής στοιβάδας με στόχο να ενισχυθεί το γονιδίωμα του ιού (Doorbar, 2012).



Εικόνα 1.12. Σχηματική αναπαράσταση E7 πρωτεΐνης. Χωρίζεται σε τρεις περιοχές. α) Η CR1 περιοχή περιλαμβάνει το αμινοτελικό άκρο το οποίο είναι απαραίτητο για τον κυτταρικό μετασηματισμό και την αποδόμηση της pRb, (β) η CR2 περιέχει ένα συντηρημένο μοτίβο LXCXE και μία θέση φωσφορυλίωσης κινάσης καζεΐνης II (CKII) (γ) η CR3 περιέχει δύο μοτίβα δακτύλων ψευδαργύρου και εμπλέκεται στην σύνδεση της pRb και άλλων κυτταρικών πρωτεϊνών, ενώ συμβάλει στον διμερισμό της πρωτεΐνης (Boulet, 2007).



Εικόνα 1.13. Σχηματική αναπαράσταση των στόχων των E6 και E7 γονιδίων που οδηγούν στην δημιουργία αθάνατων κυττάρων.

1.7 Μορφή του γονιδιώματος

Το γονιδίωμα του ιού μπορεί να βρίσκεται σε μορφή επισώματος, ή ενσωματωμένο στο DNA του κυττάρου ξενιστή ή ταυτόχρονα και στις δύο μορφές. Μελέτες έχουν δείξει ότι "επίμονη" μόλυνση από HPV τύπους υψηλού κινδύνου μπορεί να οδηγήσει στην ενσωμάτωση του ιού εντός του χρωμοσώματος του κυττάρου ξενιστή (Tonon et al., 2001, Andersson et al., 2005, Pett et al., 2006, Arias Pulido et al., 2006, Huang et al., 2006, Cheung et al., 2006, De Marco et al., 2007, Briolat et al., 2007, Li et al., 2008, Cricca et al., 2009, Matovina et al., 2009). Ενσωμάτωση του ιικού DNA έχει παρατηρηθεί επιπρόσθετα σε HPV τύπους χαμηλού κινδύνου αλλά με μικρότερη συχνότητα (Coutlée et al., 2009, Wang et al., 2013). Σε μολυσμένα κύτταρα στα οποία ο ιός έχει "εγκατασταθεί", διατηρεί το γονιδίωμα του υπό την μορφή επισώματος. Το γονιδίωμα του ιού αντιγράφεται με συχνότητα μία φορά σε κάθε κυτταρικό κύκλο στην φάση S. Με αυτό τον τρόπο ο ιός εξασφαλίζει την διατήρηση της μόλυνσης στα κύτταρα της βασικής στοιβάδας (Doorbar, 2012). Η ενσωμάτωση του ιού πραγματοποιείται κατά την διαδικασία επιδιόρθωσης των ασταθών χρωμοσωμικά κυττάρων και θεωρητικά γίνεται παράλληλα σε πολλούς κυτταρικούς κλώνους σε ασταθείς περιοχές του γονιδιώματος. Στην συνέχεια πραγματοποιούνται διαδικασίες επιλογής των κυτταρικών κλώνων με αποτέλεσμα την επιλεκτική ανάπτυξη ενός ή μερικών κλώνων με την υψηλότερη έκφραση των ογκογονιδίων E6 και E7, οι οποίοι στην συνέχεια μπορεί να εξελιχθούν σε όγκο (Wentzensen., 2004, Ziegert et al., 2013).

Η ρήξη του γονιδιώματος του ιού κατά την διάρκεια της ενσωμάτωσης συμβαίνει συνήθως στο 3' άκρο του ιικού DNA, με αποτέλεσμα να διατηρείται ακέραιη η ρυθμιστική περιοχή (LCR) και τα ογκογονίδια E6 και E7. Μελέτες έχουν δείξει ότι οι συνηθέστερες θέσεις ρήξης του γονιδιώματος του ιού είναι τα ρυθμιστικά γονίδια E1 και E2 (Chen et al., 1994, Kalantari et al., 1998, Lukaszuk et al., 2003, Arias Pulido et al., 2006, Cricca et al., 2009), ενώ επιπλέον έχει παρατηρηθεί ρήξη στα γονίδια L1 και L2 (Li et al., 2013, Xu et al., 2013, Akagi et al., 2014). Πιο συγκεκριμένα, έχουν παρατηρηθεί θέσεις ρήξης σε όλο το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου, ενώ στο E2 γονίδιο έχει εντοπιστεί ρήξη με μεγαλύτερη συχνότητα μεταξύ των νουκλεοτιδίων 3243 και 3539. Η θέση αυτή συμπεριλαμβάνει την περιοχή "αρμός" (hinge region) της E2 πρωτεΐνης. (Arias-Pulido et al., 2006, Cricca et al., 2009).

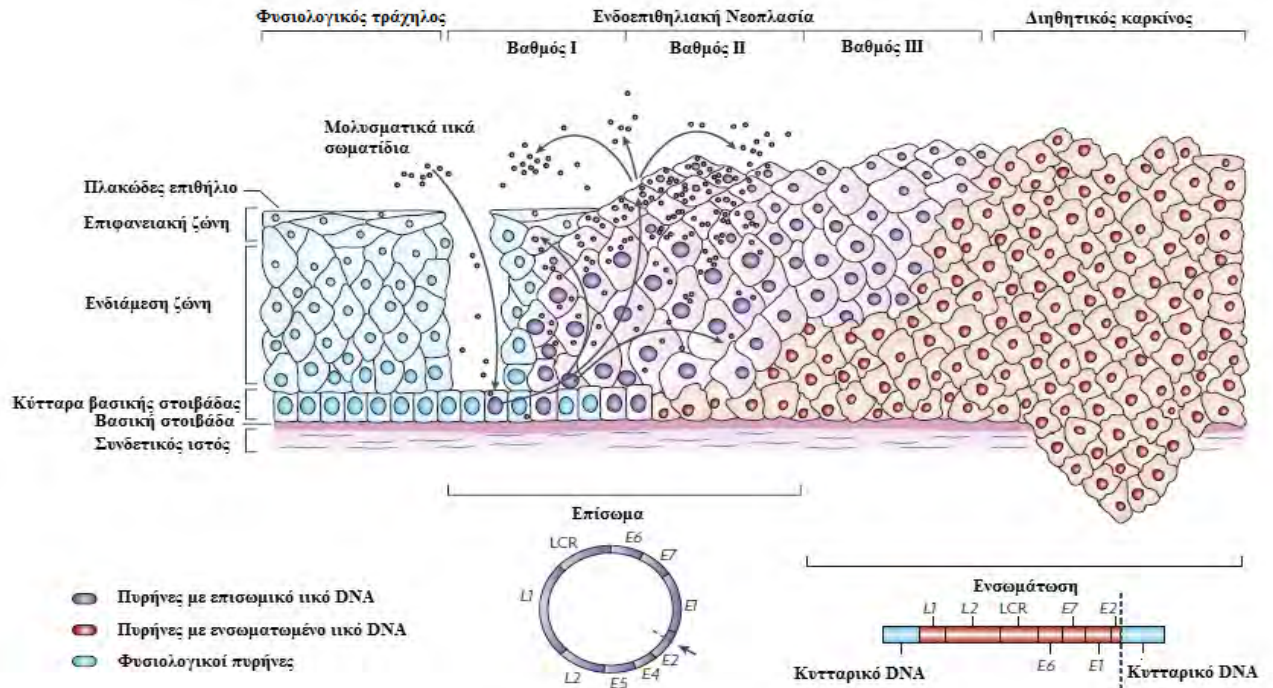
Η ρήξη των E1 και E2 γονιδίων οδηγεί πρώτον, στην απορύθμιση του ελέγχου αντιγραφής των επισωμικών αντιγράφων του ιού, καθώς δεν εκφράζεται η E1 ελικάση και δεύτερον σταθεροποιεί την έκφραση των ογκογονιδίων E6 και E7 σε υψηλά επίπεδα, καθώς δεν εκφράζεται ο καταστολέας του πρώιμου υποκινητή (E2 πρωτεΐνη) (Arias Pulido et al., 2006, Cricca et al., 2009, Hamid et al., 2009, Kadaja et al., 2009, Thierry, 2009). Η απορύθμιση αυτή οδηγεί σταδιακά στον ανεξέλεγκτο πολλαπλασιασμό των επιθηλιακών κυττάρων και στην προώθηση γενετικών λαθών στο γονιδίωμα του ξενιστή, τα οποία με την σειρά τους οδηγούν στην εμφάνιση ενδοεπιθηλιακών αλλοιώσεων και καρκίνο (Arias Pulido et al., 2006, Huang et al., 2006, Kulmala et al., 2006, Cricca et al., 2009).

Σε γενικές γραμμές έχει παρατηρηθεί ότι η ενσωμάτωση του γονιδιώματος του ιού σχετίζεται σημαντικά με υψηλότερου βαθμού δυσπλασίες (Εικόνα 1.14) (Hudelist et al., 2004, Arias Pulido et al., 2006, Li et al., 2008, Matovina et al., 2009). Τα τελευταία χρόνια έχει προταθεί ότι η ενσωμάτωση είναι ένα γεγονός το οποίο συμβαίνει νωρίς στον κύκλο ζωής του ιού γεγονός που μπορεί να οδηγήσει στην εξέλιξη της τραχηλικής ενδοεπιθηλιακής αλλοίωσης από βαθμό I (CINI) σε βαθμό II, III και κατ' επέκταση σε καρκίνο (De Marco et al., 2007, Briolat et al., 2007, Collins et al., 2009, Boulet et al., 2009, Manawapat et al., 2012). Επιπρόσθετες μελέτες έχουν δείξει ότι ο καρκίνος του τραχήλου της μήτρας μπορεί να προκληθεί από κύτταρα τα οποία περιέχουν τον ιό μόνο υπό την επισωμική του μορφή. Όσον αφορά τον HPV16 περίπου το 30% των περιπτώσεων καρκίνου του τραχήλου της μήτρας φέρουν τον ιό υπό την επισωμική του μορφή (Pett et al., 2007, Vinokurova et al., 2008). Αντιθέτως περίπου το 70% των περιπτώσεων καρκίνου του τραχήλου της μήτρας που προκαλείται από τον HPV16 φέρουν τον ιό σε ενσωματωμένη μορφή, ενώ όσον αφορά τον HPV18 το ιικό γονιδίωμα βρίσκεται μόνο ενσωματωμένο. (Fehrmann et al., 2003, Badaracco et al., 2002, Woodman et al., 2003). Επομένως, ο έλεγχος της μορφής του γονιδιώματος του ιού είναι ένας χρήσιμος βιο-δείκτης για την πρόβλεψη της εξέλιξης της τραχηλικής ενδοεπιθηλιακής αλλοίωσης και την ανάπτυξη καρκίνου (Gradissimo Oliveira et al., 2013).

Γιατί όμως ο ιός να ενσωματώνεται; Έχει παρατηρηθεί ότι τα μετάγραφα των ενσωματωμένων ογκογονιδίων περιλαμβάνουν κυτταρικές αλληλουχίες στην θέση 3'UTR. Τα χμιαρικά HPV μετάγραφα παρουσιάζουν μεγαλύτερη σταθερότητα από τα μετάγραφα που εκφράζονται από την επισωμική μορφή του ιού και έχουν μεγαλύτερο χρόνο ζωής στο κυτταρόπλασμα (Couturier et al., 1991, Jeon et al.,

1995a). Το γεγονός αυτό δίνει την δυνατότητα στον ιό να διεγείρει έντονα τον κυτταρικό κύκλο και την διαρκή είσοδο του κυττάρου στη φάση της μίτωσης, δίνοντας στα μολυσμένα κύτταρα ένα ισχυρό αναπτυξιακό πλεονέκτημα. (Jeon et al., 1995a, Doorbar, 2012). Σε δεύτερη φάση μετά τον συνεχή πολλαπλασιασμό των μολυσμένων κυττάρων προκαλείται συσσώρευση γενετικών λαθών και αποσταθεροποίηση του γονιδιώματος του ξενιστή με αποτέλεσμα την δημιουργία καρκινικών αλλοιώσεων (Duensing, 2004).

Δύο τύποι ενσωμάτωσης έχουν καταγραφεί (Jeon et al., 1995, Xu et al., 2013, Akagi et al., 2014). Ο τύπος I περιλαμβάνει ένα μοναδικό αντίγραφο του ιού ενσωματωμένο στο χρωμόσωμα. Ο τύπος II, περιλαμβάνει αλλυσομερή (concatemers) ικών γονιδιωμάτων σε μία περιοχή του κυτταρικού χρωμοσώματος. Έχει παρατηρηθεί ότι και στις δύο περιπτώσεις το γονιδίωμα του ιού αντιγράφεται παράλληλα με το κυτταρικό γονιδίωμα. Στις περιπτώσεις τύπου II συχνά παρατηρείται μόνο το ενσωματωμένο γονιδίωμα του ιού να είναι μεταγραφικά ενεργό, ενώ η επισωματική μορφή του γονιδιώματος του ιού που συσσωρεύεται στα κύτταρα δεν μεταγράφεται. Ένας πιθανός λόγος μπορεί να είναι η επιγενετική τροποποίηση των επισωματικών ικών γονιδιωμάτων (Johannsen, 2013). Παρόλα αυτά επιπρόσθετες μελέτες θα πρέπει να πραγματοποιηθούν για την εξακρίβωση του παραπάνω φαινομένου.



Εικόνα 1.14. Σχηματική αναπαράσταση της σχέσης μεταξύ ενσωμάτωσης του ιού και τραχηλικής ενδοεπιθηλιακής αλλοίωσης βαθμού I, II, III, και διηθητικού καρκίνου (Woodman, 2007).

1.8 Θέσεις ενσωμάτωσης

Εκτός από τις θέσεις που έχει υποστεί ρήξη το γονιδίωμα του ιού, σημαντικό ρόλο στην ανάπτυξη καρκινικών αλλοιώσεων παίζει και η θέση ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα (Woodman, 2007, Wentzensen., 2004). Ο μηχανισμός την ενσωμάτωσης περιλαμβάνει το σπάσιμο της διπλής έλικας του DNA του ιού και του κυτταρικού DNA και την συνένωση τους μέσω κυτταρικών πρωτεϊνών. Παρόλα αυτά, ο ακριβής μηχανισμός δεν έχει εξακριβωθεί πλήρως (Winder et al., 2007). Η ενσωμάτωση του γονιδιώματος του HPV προκαλεί δομικές αλλαγές στο χρωμόσωμα όπου ο ιός έχει ενσωματωθεί, όπως διπλασιασμοί, ελλείψεις, μετατοπίσεις (Peter et. al, 2010, Xu et al., 2013, Akagi et al., 2014). Μέχρι σήμερα περισσότερες από 250 θέσεις ενσωμάτωσης του ιού έχουν καταγραφεί σε όλα τα χρωμοσώματα (Wentzensen, 2004, Yu et al., 2005, Matovina et al., 2009). Μελέτες υποστηρίζουν ότι η ενσωμάτωση συμβαίνει στις κοινές εύθραυστες θέσεις του κυτταρικού γονιδιώματος (Common fragile sites- CFS), και συχνά στην περιοχή διαφόρων γονιδίων που εμπλέκονται στην ανάπτυξη καρκίνου, όπως το ογκογονίδιο Myc (Ferber et al., 2003, Thorland et al., 2000, 2003).

Οι κοινές εύθραυστες θέσεις (Common Fragile Sites - CFS) είναι περιοχές οι οποίες παρουσιάζουν ελλείψεις, ρήξεις, αναδιατάξεις, διπλασιασμούς και άλλα χαρακτηριστικά ασταθούς DNA στα μεταφασικά χρωμοσώματα μετά από μερική αναστολή της αντιγραφής του DNA. Οι θέσεις αυτές είναι σταθερές στα σωματικά κύτταρα, ενώ σε περιπτώσεις καρκίνου συμβάλουν στην γενωμική αστάθεια του κυττάρου. Οι περιοχές αυτές εκτείνονται από μερικές εκατοντάδες χιλιάδες βάσεις έως εκατομμύρια βάσεις και περιέχουν περιοχές DNA με μεγάλη ελαστικότητα (Glover, 2006). Η πιο κοινή εύθραυστη θέση είναι η FRA3B (3p14.2) η οποία υφίσταται τροποποιήσεις σε διάφορους τύπους καρκίνου μεταξύ των οποίων και ο καρκίνος του τραχήλου της μήτρας. Η περιοχή αυτή έχει μέγεθος 4 Mbp και στο κέντρο της τοποθετείται το ογκοκατασταλτικό γονίδιο FHIT. Ενσωμάτωση του HPV16 εντός του ιντρονίου 3 του συγκεκριμένου γονιδίου έχει εντοπιστεί σε περιπτώσεις καρκίνου του τραχήλου της μήτρας (Ferber et al., 2003).

Αρκετά γονίδια γνωστά για την επίδραση που ασκούν στην ανάπτυξη του καρκίνου έχει παρατηρηθεί ότι διαρρηγνύονται λόγω ενσωμάτωσης του HPV. Μερικά από αυτά είναι τα Myc, TP63, NR4A2, APM-1, FANCC, TNFAIP2 και hTERT (Wentzensen, 2004, Schmitz et al., 2012a). Έχει παρατηρηθεί ότι το 25% των περιπτώσεων ενσωμάτωσης του γονιδιώματος του HPV18 συμβαίνει κοντά στην

περιοχή Myc στην θέση 8q24.2. το οποίο συνορεύει μεταξύ δύο CFS περιοχών FRA8C (8q24.1) και FRAD (8q24.3) (Ferber et al., 2003). Επιπλέον η ενσωμάτωση του HPV16 κοντά στο ογκογονίδιο Myc συναντάται στο 10% των περιπτώσεων καρκίνου του τραχήλου της μήτρας και σχετίζεται με την αυξημένη έκφρασή του ογκογονιδίου (Herrick et al., 2005, Wentzensen, 2004, Peter et al., 2006).

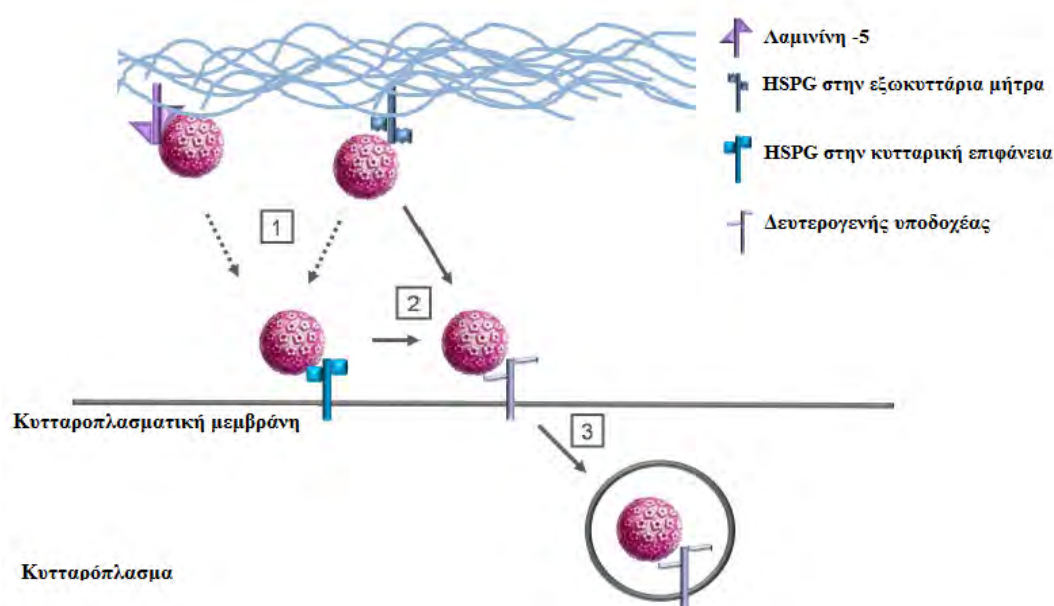
Πρόσφατες μελέτες έδειξαν ότι η ενσωμάτωση του ιού δεν γίνεται εντελώς τυχαία (Schmitz et al., 2012a, 2012b) και οι θέσεις ενσωμάτωσης τείνουν να εντοπίζονται σε επαναλαμβανόμενες περιοχές του κυτταρικού γονιδιώματος, (Li et al., 2013). Επιπλέον πρόσφατη μελέτη απέδειξε ότι πολλές περιοχές του γονιδιώματος του ιού παρουσιάζουν σημαντική ομοιότητα πάνω από 50 νουκλεοτίδια με τα γονίδια στα οποία ενσωματώνονται καθώς και με και τις θέσεις ενσωμάτωσης, γεγονός που δίνει σημαντικές πληροφορίες για τον μηχανισμό ενσωμάτωσης του ιού και για τις θέσεις του κυτταρικού χρωμοσώματος που εμπλέκονται στην ενσωμάτωση (Schmitz et al., 2012b). Παρόλα αυτά επιπρόσθετες μελέτες απαιτούνται έτσι ώστε να οδηγηθούμε σε ασφαλέστερα συμπεράσματα για τον τρόπο με τον οποίο ο ιός επιλέγει τις θέσεις ενσωμάτωσης και τον μοριακό μηχανισμό σύμφωνα με τον οποίο ο ιός ενσωματώνεται εντός του κυττάρου ξενιστή.

Συμπερασματικά, εκτός από την εκτεταμένη ενεργοποίηση των ογκογονιδίων E6 και E7, η θέση στο κυτταρικό χρωμόσωμα όπου ο ιός έχει ενσωματωθεί παίζει πολύ σημαντικό ρόλο στον ρυθμό ανάπτυξης και στην φύση του όγκου (Matovina et al., 2009), ενώ η ανίχνευση θέσεων ενσωμάτωσης του ιού σε κλινικά δείγματα μπορεί να δώσει καθοριστικές πληροφορίες για την θεραπεία του ασθενή.

1.9 Ο κύκλος ζωής των Papilloma ιών του ανθρώπου

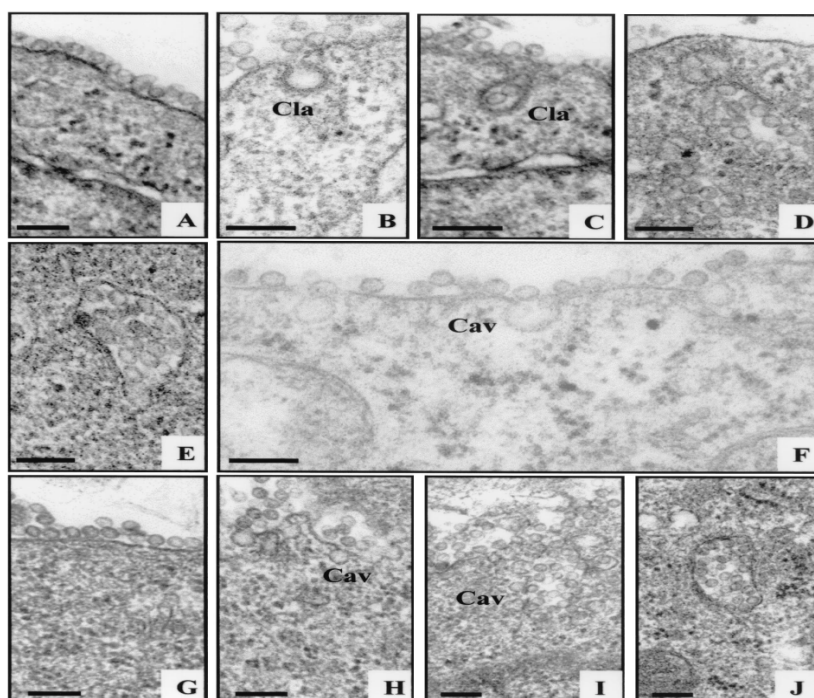
1.9.1 Είσοδος του ιού στο κύτταρο

Ο κύκλος ζωής των HPV ακολουθεί το πρόγραμμα διαφοροποίησης των κερατινοκυττάρων του ξενιστή. Τα ιικά σωματίδια προσβάλλουν τους επιθηλιακούς ιστούς μέσω μικροτραυματισμών, με αποτέλεσμα ο ιός να εισβάλει στα κύτταρα της βασικής στοιβάδας. Ο ιός εισέρχεται εντός των βασικών επιθηλιακών κυττάρων, μέσω της σύνδεσης του με την θεϊκή ηπαρίνη (HSPG) η οποία αποτελεί συστατικό των πρωτεογλυκανών, που αλληλεπιδρούν με την κυτταρική μεμβράνη (Shafti-Keramat et al., 2003, Johnson et al., 2009). Η θεϊκή ηπαρίνη βρίσκεται στην εξωκυττάρια μήτρα (ECF) και στην επιφάνεια πολλών κυττάρων (Εικόνα 1.15). Όπως συμβαίνει με πολλούς άλλους ιούς, οι HPV απαιτούν την παρουσία και δεύτερου υποδοχέα για την αποτελεσματική είσοδό του στο κύτταρο ξενιστή, όπως είναι οι Α6 – ιντεγκρίνες (Shafti-Keramat et al., 2003, Surviladze et al., 2012). Επιπλέον έχει αποδειχθεί, ότι η λαμινίνη-5 είναι ένας ακόμα παράγοντας ικανός να λειτουργεί ως υποδοχέας των HPV στην επιφάνεια των κερατινοκυττάρων (Εικόνα 1.15) (Culp et al., 2006).



Εικόνα 1.15. Πιθανό μοντέλο αλληλεπίδρασης του ιικού καψιδίου με την εξωκυττάρια μήτρα και την επιφάνεια του κυττάρου. 1) Η θεϊκή ηπαρίνη λειτουργεί ως υποδοχέας του ιού βρίσκεται στην εξωκυττάρια μήτρα και στην επιφάνεια των κυττάρων. Τα καψίδια επίσης αλληλεπιδρούν με την λαμινίνη-5 2) Σύνδεση του ιού με δευτερογενή υποδοχέα. 3) Είσοδος του ιού στο κυτταρόπλασμα με ενδοκυττάρωση (Horvath, 2010).

Στην συνέχεια ο ιός μετά την πρόσδεση στον υποδοχέα του, εισέρχεται εντός του κυττάρου με ενδοκυττάρωση σε κυστίδια κλαθρίνης ή καβεολίνης, ανάλογα με τον τύπο του ιού (Εικόνα 1.16). Έχει προταθεί ότι οι HPV16 και HPV58 εισέρχονται σε κυστίδια κλαθρίνης ενώ ο HPV-31 συνδέεται με κυστίδια καβεολίνης (Bousarghin et al., 2003). Έπειτα τα ιικά σωματίδια αποσυναρμολογούνται στα ενδοσώματα ή/και στα λυσοσώματα και το ιικό DNA μεταφέρεται στον πυρήνα με την βοήθεια της καψιδιακής πρωτεΐνης L2 ενώ η L1 πρωτεΐνη αποδομείται εντός των ενδοσωμάτων . (Bergant Marusic et al., 2012, Schelhaas et al., 2012).

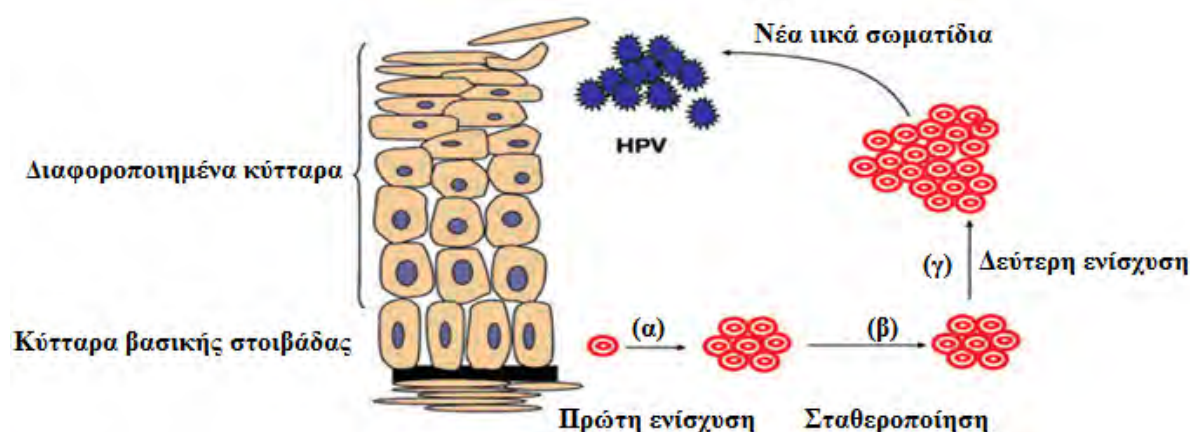


Εικόνα 1.16. Ηλεκτρονική μικρογραφία της πρόσδεσης και της ενδοκυττάρωσης του HPV16 και HPV31 σε κυστίδια κλαθρίνης και καβεολίνης αντίστοιχα. (A-E) HPV16 ιικά σωματίδια, (F-J) HPV31 ιικά σωματίδια (Bousarghin et al., 2003).

1.9.2 Αντιγραφή του ιικού γονιδιώματος

Οι HPV αντιγράφουν το γονιδίωμα τους σε διαφορετικές περιόδους κατά την διάρκεια του κύκλου ζωής τους και με διαφορετική συχνότητα. Σε πρώτο επίπεδο έχει προταθεί, ότι αμέσως μετά την μόλυνση των κυττάρων από τον ιό παρατηρείται ενίσχυση του ιικού γονιδιώματος περίπου διακοσίων αντιγράφων σε κάθε κύτταρο (Maglennon, 2011). Σε αυτό το σημείο το ιικό γονιδίωμα παραμένει εντός του πυρήνα ως επίσωμα. Η επόμενη φάση του πολλαπλασιασμού του ιικού DNA συμβαίνει κατά την διαίρεση των κυττάρων της βασικής στοιβάδας στην φάση S. Στην φάση αυτή το

ικό DNA αντιγράφεται μαζί με το DNA του κυττάρου ξενιστή, δημιουργώντας δύο αντίγραφα ένα σε κάθε θυγατρικό κύτταρο με σκοπό την σταθεροποίηση του ικού γονιδιώματος (McBride, 2008, Pyeon et al., 2009, Parish et al., 2006). Η τελευταία φάση αντιγραφής του ικού DNA συμβαίνει στα κύτταρα των ανώτερων στοιβάδων, όπου παρατηρείται έξαρση της αντιγραφής του ιού με σκοπό την σύνθεση του ικού DNA που θα εισέλθει εντός των νεοσχηματιζόμενων ικών σωματιδίων (Hoffmann et al., 2006).



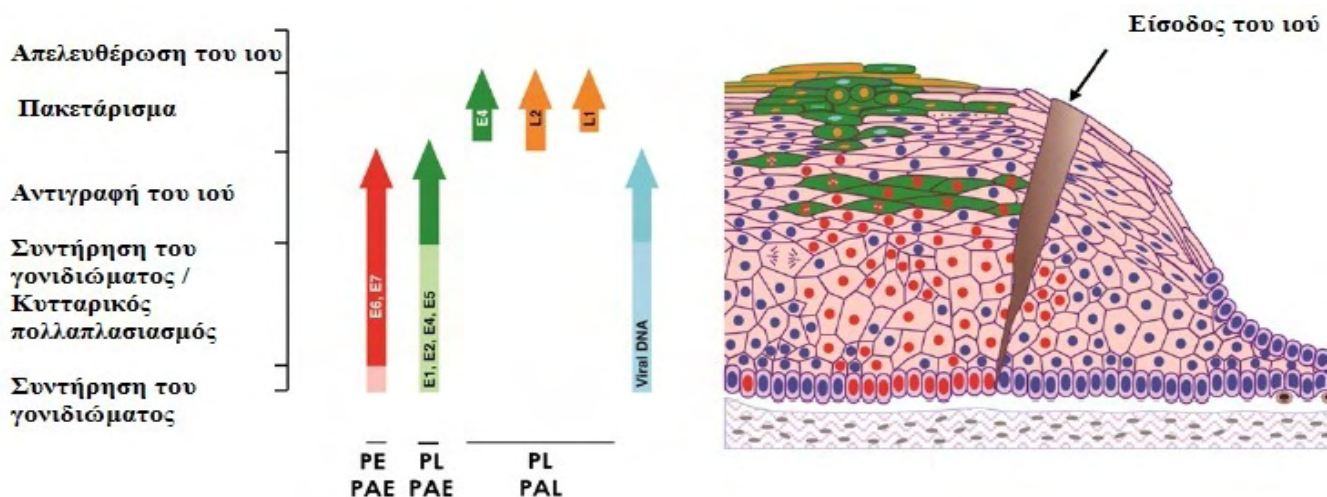
Εικόνα 1.17. α) Η πρώτη ενίσχυση του ικού γονιδιώματος γίνεται αμέσως μετά την μόλυνση των κυττάρων της βασικής στοιβάδας, με την παραγωγή περίπου 200 ικών αντιγράφων σε κάθε κύτταρο. β) Στην συνέχεια σταθεροποιείται του ικό γονιδίωμα και η αντιγραφή του γίνεται παράλληλα με την αντιγραφή του μολυσμένου κυττάρου. γ) Τέλος στα μολυσμένα κύτταρα των ανώτερων στοιβάδων παρατηρείται μια παρατεταμένη αντιγραφική δραστηριότητα του ιού με στόχο την παραγωγή πολυάριθμων μορίων ικού DNA για την σύνθεση νέων μολυσματικών ικών σωματιδίων.

1.9.3 Μεταγραφή του ικού γονιδιώματος

Η έκφραση των γονιδίων του ιού εξαρτάται από το πρόγραμμα διαφοροποίησης των κερατινοκυττάρων. Ο μηχανισμός με τον οποίο η διαφοροποίηση των κερατινοκυττάρων ρυθμίζει την έκφραση των γονιδίων των HPV δεν έχει πλήρως εξακριβωθεί. Έχει παρατηρηθεί ότι στα αδιαφοροποίητα κερατινοκύτταρα ή τα ενδιάμεσα διαφοροποιημένα κερατινοκύτταρα εκφράζονται έξι μη δομικές ρυθμιστικές πρωτεΐνες (E1, E2, E4, E5, E6, και E7) από την πρόιμη περιοχή του ικού γονιδιώματος. Τα έξι πρώιμα ανοιχτά πλαίσια ανάγνωσης των γονιδίων E1, E2, E4, E5, E6, E7 βρίσκονται υπό τον άμεσο έλεγχο του πρώιμου μεταγραφικού υποκινητή

(p97 για τον HPV16, HPV31 και p105 για τον HPV18) για την εγκαθίδρυση της μόλυνσης και την αντιγραφή του DNA του ιού στα πρώτα στάδια της εισόδου του ιού στο κύτταρο (Εικόνα 1.18) (Zheng, 2006).

Από την άλλη πλευρά, ο όψιμος υποκινητής ενεργοποιείται ειδικά σε διαφοροποιημένες στοιβάδες του επιθηλίου. Ο όψιμος υποκινητής τοποθετείται ανοδικά του E1 γονιδίου (θέση 670 για τον HPV16) και ενεργοποιεί την έκφραση των γονιδίων E1 και E2 κατά την διάρκεια της διαφοροποίησης των κυττάρων με κύριο σκοπό την ενίσχυση του DNA του ιού (Εικόνα 3.3) (Grassmann et al., 1996). Στο επιθηλιακό κύτταρο που πραγματοποιείται ενίσχυση του ιικού γονιδιώματος, έχει παρατηρηθεί ότι εκφράζονται παράγοντες διαφοροποίησης όπως κερατίνες 1 και 10 (δερματικό επιθήλιο) ή 4 και 13 (επιθήλιο του βλεννογόνου). Επιπλέον, ο όψιμος υποκινητής συμβάλει στην έκφραση των καψιδιακών πρωτεϊνών L1 και L2, οι οποίες είναι απαραίτητες για τον σχηματισμό του ιικού καψιδίου και το πακετάρισμα του ιικού γονιδιώματος στα νεοσχηματιζόμενα ιικά σωματίδια (Εικόνα 1.18) (Doorbar, 2012). Ακόμα παραμένει απαραίτητο να ξεκαθαριστεί ο μηχανισμός με τον οποίο ρυθμίζεται ο όψιμος υποκινητής κατά την διαδικασία διαφοροποίησης των επιθηλιακών κυττάρων. Πέρα από την έλεγχο του όψιμου υποκινητή από κυτταρικούς ρυθμιστικούς παράγοντες, έχει παρατηρηθεί για τον HPV16, ότι τα διαφορετικά επίπεδα μεθυλίωσης σε θέσεις της ρυθμιστικής περιοχής LCR σε διαφοροποιημένα επιθηλιακά κύτταρα συμβάλουν στον έλεγχο της μεταγραφής του ιού. Πιο συγκεκριμένα, σε πλήρη διαφοροποιημένα επιθηλιακά κύτταρα υπάρχει μειωμένη μεθυλίωση στην περιοχή ενίσχυσης της μεταγραφής και αύξηση της μεθυλίωσης στην περιοχή του πρώιμου υποκινητή και πιο συγκεκριμένα στην θέση πρόσδεσης της E2 πρωτεΐνης (Vinokurova et al., 2011).



Εικόνα 1.18. Σχηματική αναπαράσταση της έκφρασης των γονιδίων του HPV16 σε ενδοεπιθηλιακές αλλοιώσεις. Τα πρώιμα ικά γονίδια E1, E2, E4, E5, E6, E7 εκφράζονται αμέσως μετά την μόλυνση των κυττάρων της βασικής στοιβάδας και βρίσκονται κάτω από τον αυστηρό έλεγχο του πρώιμου υποκινητή PE (κόκκινο χρώμα). Τα προϊόντα των γονιδίων αυτών εμπλέκονται στην σταθεροποίηση και ενίσχυση του ιικού γονιδιώματος (E1, E2) και στον ανεξέλεγκτο κυτταρικό πολλαπλασιασμό (E5, E6, E7). Στην όψιμη φάση της ζωής του ιού εκφράζονται τα όψιμα γονίδια L1, L2 τα οποία είναι υπεύθυνα για την σύσταση του ιικού καψιδίου και το πακετάρισμα του γονιδιώματος. Η έκφραση των γονιδίων αυτών βρίσκεται κάτω από την επίδραση του όψιμου υποκινητή PL (πράσινο χρώμα). Επίσης τα γονίδια E1, E2, E4, E5 εκφράζονται σε μεγάλες ποσότητες κατά την όψιμη φάση της ζωής του ιού ενώ συνεχίζουν να εκφράζονται και στα πλήρως διαφοροποιημένα κερατινοκύτταρα με την βοήθεια του όψιμου υποκινητή PL (πράσινο χρώμα) (Doorbar, 2012).

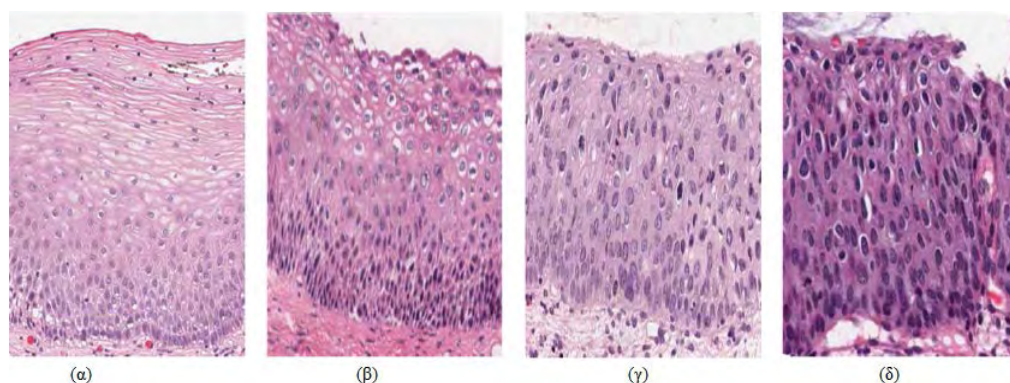
1.9.4 Απελευθέρωση του ιού

Το τελευταίο στάδιο του κύκλου ζωής του ιού περιλαμβάνει την τοποθέτηση των αντιγράφων του ιικού DNA εντός των νεοσχηματιζόμενων καψιδίων και την σύνθεση των νέων ιικών σωματιδίων. Η συναρμολόγηση των νέων μολυσματικών ιικών σωματιδίων πραγματοποιείται στην ανώτερη επιθηλιακή στοιβάδα (Εικόνα 1.18). Για την συναρμολόγηση τους απαιτούνται οι καψιδιακές πρωτεΐνες L1, L2, οι οποίες θα κατασκευάσουν το καψίδιο καθώς επίσης και η πρωτεΐνη E2, η οποία συμβάλει στην τοποθέτηση του ιικού DNA εντός των καψιδίων (Buck et al., 2004). Η ωρίμανση των ιικών σωματιδίων συμβαίνει όταν τα μολυσμένα κύτταρα ανέρθουν στην ανώτερη επιθηλιακή στοιβάδα. Τελικά ο ιός απελευθερώνεται από το κερατινοποιημένο κάλυμμα που έχει σχηματιστεί στην κυτταρική επιφάνεια με την βοήθεια της πρωτεΐνης E1^{E4} χωρίς να προκαλεί λύση των κερατινοκυττάρων (Lehr E et al 2004).

1.10 Ιστολογικές αλλοιώσεις

Οι HPV ανάλογα με τον τροπισμό τους προσβάλλουν το επιθήλιο του δέρματος και το επιθήλιο του βλεννογόνου. Οι δερματικοί HPV τύποι προκαλούν καλοήγη δερματικά θηλώματα, όπως το κοινό θήλωμα, ακροχορδώδες θήλωμα, επίπεδο θήλωμα και σπάνια εμπλέκονται στην πρόκληση καρκίνου (Pffister, 2003). Οι HPV που προσβάλλουν το επιθήλιο του βλεννογόνου προκαλούν αλλοιώσεις στο ουρογεννητικό σύστημα, στον ανώτερο αναπνευστικό σωλήνα, τον οισοφάγο και αποτελούν τον κύριο αιτιολογικό παράγοντα για την ανάπτυξη τραχηλικής ενδοεπιθηλιακής νεοπλασίας και καρκίνου του τραχήλου της μήτρας (Zur Hausen, 1996). Ο καρκίνος του τραχήλου της μήτρας αποτελεί τον τρίτο πιο συχνά εμφανιζόμενο τύπο καρκίνου που προσβάλλει τον γυναικείο πληθυσμό μετά τον καρκίνο του μαστού και του εντέρου και αποτελεί βασικό αντικείμενο έρευνας τις τελευταίες δεκαετίες (Jemal, 2011).

Οι τραχηλικές ενδοεπιθηλιακές αλλοιώσεις και δείγματα καρκίνου του τραχήλου της μήτρας, αποτελούν κύριο υλικό μελέτης της παρούσας διδακτορικής διατριβής. Οι τραχηλικές ενδοεπιθηλιακές αλλοιώσεις αρχικά πήραν τον ορισμό δυσπλασία και διαχωρίστηκαν περαιτέρω σε ήπιο, μέτριο, σοβαρό βαθμό και *in situ* καρκίνο (Reagan, 1953). Για να γίνει ο διαχωρισμός πιο ακριβής το 1968 οι επιστήμονες εισήγαγαν τον όρο τραχηλική ενδοεπιθηλιακή νεοπλασία (Cervical Intraepithelial Neoplasia - CIN) (Richart, 1967). Η τραχηλική ενδοεπιθηλιακή νεοπλασία χωρίζεται σε τρεις βαθμούς I, II και III, οι οποίοι αναφέρονται στην ήπια, μεσαία και σοβαρή μορφή δυσπλασίας, αντίστοιχα (He, 2011, De et al., 2013). Η διάγνωση ανωμαλιών του τραχήλου γίνεται με πολλές διαδικασίες, όπως το τεστ Pap, κολποσκόπηση για την οπτική παρατήρηση του τραχήλου, ενώ όταν υπάρχει διαθέσιμος ιστός τραχήλου γίνεται ιστοπαθολογική ανάλυση (De et al., 2013). Πρακτικά η τελική διάγνωση της αλλοίωσης πραγματοποιείται με την ιστοπαθολογική ανάλυση του τραχηλικού ιστού. Η σοβαρότητα της νεοπλασίας βασίζεται στο εύρος του επιθηλίου με παθολογικά κύτταρα. Για παράδειγμα, στην ενδοεπιθηλιακή νεοπλασία τύπου I ανώριμα παθολογικά κύτταρα βρίσκονται στο κάτω 1/3 άκρο του επιθηλίου, στην ενδοεπιθηλιακή νεοπλασία τύπου II τα ανώριμα παθολογικά κύτταρα βρίσκονται στο μισό επιθήλιο και στην περίπτωση της νεοπλασίας τύπου III τα ανώριμα παθολογικά κύτταρα καλύπτουν όλη την έκταση του επιθηλίου (Εικόνα 1.19) (He, 2011, De et al., 2013).



Εικόνα 1.19. Παραδείγματα τραχηλικής ενδοεπιθηλιακής νεοπλασίας (CIN). α) Φυσιολογικός ιστός, (β) τραχηλική ενδοεπιθηλιακή νεοπλασία τύπου I (CINI), τραχηλική ενδοεπιθηλιακή νεοπλασία τύπου II (CINII), τραχηλική ενδοεπιθηλιακή νεοπλασία τύπου III (CINIII) (De et al., 2013).

Ένα επιπλέον σύστημα ταξινόμησης δύο κατηγοριών έχει δημιουργηθεί για να περιγράψει όλες τις μη διηθητικές τραχηλικές αλλοιώσεις. Οι αλλοιώσεις που προκαλούνται από ένα ολοκληρωμένο κύκλο ζωής του ιού αναφέρονται ως χαμηλού βαθμού δυσπλασίες (*low-grade*) και περιλαμβάνουν αλλοιώσεις της τάξης CINI, ενώ οι πιο σοβαρές αλλοιώσεις που αποτελούν τον πρόδρομο εμφάνισης καρκίνου αναφέρονται ως υψηλού βαθμού αλλοιώσεις (*high-grade*) και περιλαμβάνουν αλλοιώσεις τάξης CINII και CINIII. Το σύστημα ορολογίας Bethesda το οποίο δημιουργήθηκε από το Εθνικό Ινστιτούτο Καρκίνου των Ηνωμένων Πολιτειών Αμερικής (NCI) για την αναφορά τραχηλικών αλλοιώσεων μετά από την εξέταση Pap, έχει υιοθετήσει τους όρους χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (*low-grade squamous intraepithelial lesion, LSIL*) και υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (*high-grade squamous intraepithelial lesion, HSIL*) (Solomon et al., 2002).

Σε μοριακό επίπεδο η μόλυνση από HPV αποτελεί το εναρκτήριο γεγονός για την πρόκληση τραχηλικής ενδοεπιθηλιακής αλλοίωσης. Παρόλα αυτά η εξέλιξη της τραχηλικής ενδοεπιθηλιακής νεοπλασίας σε καρκίνο είναι μια μακρά διαδικασία (Heise, 2003). Μετά την είσοδο του ιού στο κύτταρο η έκφραση του ιικού γονιδιώματος μπορεί να κατασταλεί (π.χ λόγω μεθυλίωσης), με αποτέλεσμα να προκαλείται ασυμπτωματική μόλυνση και το ιικό γονιδίωμα να παραμένει στην βασική στοιβάδα του επιθηλίου. Εναλλακτικά, η μόλυνση μπορεί να οδηγήσει στην έκφραση του γονιδιώματος του ιού οδηγώντας στην δημιουργία νέων ιικών σωματιδίων και την έξοδο του ιού από την ανώτερη επιθηλιακή στοιβάδα (CINI), ή στην απορύθμιση της έκφρασης των γονιδίων του ιού και την δημιουργία υψηλού

βαθμού ενδοεπιθηλιακής αλλοίωσης (CINI/CINIII). Η "επίμονη" υψηλού βαθμού αλλοίωση συνδέεται με αυξανόμενο ρίσκο ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα και την πρόοδο της αλλοίωσης σε καρκίνο (Doorbar, 2012).

1.11 Επιδημιολογία

Ο καρκίνος του τραχήλου της μήτρας είναι ο τρίτος πιο συχνά εμφανιζόμενος τύπος καρκίνου που εμφανίζεται στον γυναικείο πληθυσμό με περισσότερα από 530.000 νέα περιστατικά κάθε χρόνο και πάνω από 270.000 θανάτους ετησίως σε όλο τον κόσμο. Περισσότερο από το 85% των περιπτώσεων θανάτου συμβαίνει σε αναπτυσσόμενες χώρες (Forman et al., 2012). Η Ινδία, η δεύτερη μεγαλύτερη χώρα σε πληθυσμό στον κόσμο καταλαμβάνει το 27% των συνολικών θανάτων από καρκίνο του τραχήλου της μήτρας. Επιπλέον από τους υψηλότερους αριθμούς περιστατικών καταλαμβάνουν η Δυτική και Νότια Αφρική, η Νότιο-Κεντρική Ασία και η Νότια Αμερική. Το μεγάλο ποσοστό καρκίνου του τραχήλου της μήτρας που καταλαμβάνουν οι αναπτυσσόμενες χώρες και οι πληθυσμοί με χαμηλή ιατρική περίθαλψη οφείλεται στην έλλειψη εξετάσεων που συμβάλουν στην ανίχνευση προκαρκινικών αλλοιώσεων (Jemal, 2011).

Όπως και πολλά από τα υπόλοιπα σεξουαλικά μεταδιδόμενα νοσήματα, οι HPV συνήθως προκαλούν σε αρχικό στάδιο ασυμπτωτική μόλυνση του ασθενή. Έτσι η λοίμωξη δεν διαπιστώνεται, δεν αντιμετωπίζεται, με αποτέλεσμα ο ιός να μπορεί να μεταδίδεται στον πληθυσμό. Στις περισσότερες μελέτες η συνηθέστερη ηλικία HPV μόλυνσης είναι μικρότερη των 25 ετών, ενώ η συνηθέστερη ηλικία εμφάνισης καρκίνου κυμαίνεται στην ηλικία των 50 ετών. Παρόλα αυτά ενδοεπιθηλιακές αλλοιώσεις μπορούν να εμφανιστούν σε οποιαδήποτε χρονική στιγμή μεταξύ τριών εβδομάδων και οχτώ μηνών μετά την μόλυνση, με την μεγαλύτερη συχνότητα εμφάνισης των συμπτωμάτων στους τρεις μήνες (Munoz et al., 2004). Οι περισσότερες μολύνσεις από HPV εξαλείφονται μόνες τους, ενώ οι περισσότερες αλλοιώσεις μπορούν να απομακρυνθούν αυτόματα χωρίς ιατρική παρέμβαση. Παρόλα αυτά υπάρχει κίνδυνος για όλες τις γυναίκες, η μόλυνση από HPV να γίνει χρόνια και οι προ-καρκινικές αλλοιώσεις να εξελιχθούν σε καρκίνο του τραχήλου της μήτρας. Χρειάζεται ένα χρονικό περιθώριο 15 - 20 ετών σε ανοσολογικά υγιείς γυναίκες να αναπτύξουν καρκίνο του τραχήλου της μήτρας ενώ σε γυναίκες με αδύναμο

ανοσολογικό σύστημα μπορεί να αναπτυχθεί καρκίνος μέσα σε ένα χρονικό διάστημα 5 έως 10 χρόνων (WHO).

Επιδημιολογικές μελέτες έχουν δείξει, ότι οι πιο επικρατείς τύποι HPV που προκαλούν καρκίνο του τραχήλου της μήτρας σε παγκόσμια κλίμακα είναι οι HPV16 και HPV18 σε ποσοστό 56,6% και 16%, αντίστοιχα. Ακολουθούν οι HPV58, 33, 45, 31, 52, 59, 39, 51, 56 (Εικόνα 1.20) (Li et al., 2011, de Sanjose et al., 2010). Παρόλα αυτά παρατηρείται γεωγραφική και ειδική για κάθε χώρα μεταβλητότητα. Όπως για παράδειγμα στην υπό-Σαχάρια Αφρική και την Λατινική Αμερική οι HPV16 και HPV18 είναι υπεύθυνοι για το 65% των περιπτώσεων διηθητικού καρκίνου. Επιπλέον, στην Δυτική-Κεντρική Ασία το 82% των περιπτώσεων καρκίνου του τραχήλου οφείλονται στους τύπους HPV16 και HPV18 ενώ στην Ανατολική Ασία μόνο το 68% των περιπτώσεων σχετίζεται με αυτούς τους δύο τύπους (Li et al., 2010). Ακόμη έχει αναφερθεί ότι HPV τύποι κατανέμονται σε διαφορετικές περιοχές με διαφορετική συχνότητα. Ένα χαρακτηριστικό παράδειγμα είναι ο HPV58, ο οποίος παρατηρείται περίπου στο 10% των περιπτώσεων καρκίνου του τραχήλου της μήτρας στην Ανατολική Ασία, καταλαμβάνοντας το υψηλότερο ποσοστό εμφάνισης του στη συγκεκριμένη περιοχή από οπουδήποτε αλλού στον κόσμο (Li et al., 2011).



Εικόνα 1.20. Παγκόσμια κατανομή HPV τύπων σε περιπτώσεις καρκίνου του τραχήλου της μήτρας, χρησιμοποιώντας δημοσιευμένα δεδομένα από το 1990 έως το 2010 (Li et al., 2011).

Κεφάλαιο 2

2. Υλικά και Μέθοδοι

2.1 Κλινικά δείγματα

Στην παρούσα διδακτορική διατριβή χρησιμοποιήθηκαν κλινικά δείγματα τραχήλου της μήτρας, που προήρθαν από τον γυναικείο ελληνικό πληθυσμό. Τα δείγματα συλλέχθηκαν από το εργαστήριο Ιολογίας του Αντικαρκινικού Νοσοκομείου Αθηνών "Άγιος Σάββας", το εργαστήριο Κυτταροπαθολογίας του Πανεπιστημιακού Γενικού Νοσοκομείου Λάρισας και από ιδιωτικό κυτταροπαθολογικό εργαστήριο. Όλα τα τραχηλικά δείγματα χαρακτηρίστηκαν από τους Κυτταροπαθολόγους ως χαμηλού βαθμού (low-grade), ή υψηλού βαθμού (high-grade) ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και καρκίνος του τραχήλου της μήτρας (Πίνακας 2.1, 2.2, 2.3). Η ανίχνευση HPV DNA και η ταυτοποίηση του HPV16 έγινε στο Εργαστήριο Μοριακής Ιολογίας - Μικροβιολογίας του τμήματος Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας.

Συνολικά μελετήθηκαν ογδόντα τέσσερα κλινικά δείγματα θετικά για μόλυνση με HPV16. Από τα ογδόντα τέσσερα δείγματα που εξετάστηκαν, δεκαεπτά δείγματα προήρθαν από βιοψίες τραχηλικού ιστού μονιμοποιημένου σε παραφίνη, οκτώ δείγματα προήρθαν από τραχηλικά επιχρίσματα και πενήντα εννέα δείγματα προήρθαν από ThinPrep. Από την κυτταροπαθολογική ανάλυση των δειγμάτων τριάντα τρία δείγματα έχουν αναφερθεί ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (*high-grade squamous intraepithelial lesion, HSIL*) και σαράντα δείγματα ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (*low-grade squamous intraepithelial lesion, LSIL*) (Πίνακας 2.1, 2.2). Επιπλέον, 11 περιπτώσεις καρκίνου του τραχήλου της μήτρας χρησιμοποιήθηκαν στην παρούσα μελέτη, εκ των οποίων οι εννέα προέκυψαν από βιοψία τραχηλικού ιστού μονιμοποιημένου σε παραφίνη και δύο περιπτώσεις προήρθαν από ThinPrep (Πίνακας 2.3).

Υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (n=33)		
Κωδικοποίηση δειγμάτων	Είδος δειγμάτων	Προέλευση
A3	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A4	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A6	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A13	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A31	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL1	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL4	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL5	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL6	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL7	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL8	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL9	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL10	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL11	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL12	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL13	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL14	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL15	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
ATT4	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT5	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT9	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT16	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT20	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT21	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT24	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT79	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
D2	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D4	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D13	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D14	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D15	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D16	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο
D30	Ιστός σε παραφίνη	Ιδιωτικό εργαστήριο

Πίνακας 2.1. Τριάντα τρία κλινικά δείγματα υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης πλακώδους επιθηλίου που εξετάστηκαν στα πλαίσια της παρούσας εργασίας. Στον πίνακα αναγράφεται η κωδικοποίηση των δειγμάτων, το είδος των δειγμάτων και η προέλευσή τους.

Χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου (n=40)		
Κωδικοποίηση δειγμάτων	Είδος δειγμάτων	Προέλευση
A17	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A24	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
A38	Τραχηλικό επίχρισμα	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL2	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL3	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
ATT8	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT6	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT10	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT12	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT18	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT22	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT23	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT26	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT45	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT46	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT51	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT54	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT78	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT80	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT91	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
ATT98	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
747	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
793	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
796	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
858	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
868	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
1143	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
1620	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2121	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2127	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2148	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2186	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2209	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2216	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2270	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2279	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
2290	ThinPrep	Αντικαρκινικό Νοσοκομείο Αθηνών "Άγιος Σάββας"
D7	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D34	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D42	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο

Πίνακας 2.2. Σαράντα κλινικά δείγματα χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης πλακώδους επιθηλίου που εξετάστηκαν στη παρούσα εργασία. Στον πίνακα αναγράφεται η κωδικοποίηση των δειγμάτων, το είδος των δειγμάτων και προέλευσή τους.

Καρκίνος του τραχήλου		
Κωδικοποίηση δειγμάτων	Είδος δειγμάτων	Προέλευση
D58	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D59	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D60	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D61	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D62	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D63	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D64	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D65	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
D66	Ιστός σε παραφίνη	Ιδιωτικό Εργαστήριο
PNL7	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας
PNL16	ThinPrep	Πανεπιστημιακό Γενικό Νοσοκομείο Λάρισας

Πίνακας 2.3. Έντεκα κλινικά δείγματα καρκίνου που εξετάστηκαν στα πλαίσια της παρούσας εργασίας. Στο πίνακα αναγράφεται η κωδικοποίηση των δειγμάτων, το είδος των δειγμάτων και η προέλευσή τους.

2.2 Εκχύλιση DNA

Τρεις διαφορετικές μέθοδοι εκχύλισης χρησιμοποιήθηκαν για την απομόνωση DNA από τα κλινικά δείγματα, ανάλογα με το είδος του δείγματος (ThinPrep, τραχηλικά επιχρίσματα, τραχηλικός ιστός μονιμοποιημένος σε παραφίνη).

2.2.1 Εκχύλιση DNA (ThinPrep)

Για τα δείγματα από ThinPrep η απομόνωση του DNA πραγματοποιήθηκε με την μέθοδο της Θειοκυανιούχου Γουανιδίνης (GuSCN) (Casas et al., 1996). Σε πλαστικό σωλήνα των 2ml αναμιγνύονται 100μl δείγματος, 10μl γλυκογόνο (100mg/ml) και 300μl διαλύματος Θειοκυανιούχου Γουανιδίνης (GuSCN). Ακολουθεί ισχυρή ανάδευση και επώαση σε θερμοκρασία δωματίου (18-25°C) για 20 min, ώστε να επιτευχθεί η λύση των κυτταρικών μεμβρανών για την απελευθέρωση των νουκλεϊκών οξέων. Στην συνέχεια, προστίθενται 400 μl ισοπροπανόλης (διατηρημένης στους -20°C) και μετά από ισχυρή ανάδευση ακολουθεί επώαση για 20 min στους -20°C. Ακολουθεί φυγοκέντρηση για 10min στις 14.000 rcf στους +4°C και απομάκρυνση του υπερκείμενου. Το ίζημα επαναδιαλύεται με ισχυρή ανάδευση 500μl αιθανόλης 70% (διατηρημένης στους -20°C). Ακολουθεί φυγοκέντρηση για 10min στις 14.000 rcf στους +4°C. Το υπερκείμενο απομακρύνεται πλήρως και το ίζημα επαναδιαλύεται σε 100 μl διπλά απιονισμένου νερού (ddH₂O) (Sigma, USA)

απαλλαγμένου ριβονουκλεασών (RNase, Dnase free). Το απομονωμένο DNA αποθηκεύτηκε στους -20°C μέχρι την περαιτέρω χρήση του.

Διαλύματα:

Διάλυμα Θειοκυανιούχου Γουανιδίνης: 4M GuSCN, 0.5% N-lauroyl sacrosine, 1mM dithiothreitol, 25 mM sodium citrate.

2.2.2 Εκχύλιση DNA (Τραχηλικά επιχρίσματα)

Η απομόνωση DNA από το πλακάκι πάνω στο οποίο είχαν μονιμοποιηθεί κύτταρα έγινε με την χρήση πρωτεΐνάσης K (Puramen et al., 1996). Απομακρύνονται τα κύτταρα από το πλακάκι με νυστέρι και τοποθετούνται σε πλαστικό σωλήνα των 2ml. Πραγματοποιείται ισχυρή ανάδευση με 0.5 έως 1.5 ml διαλύματος I, ανάλογα με την ποσότητα των κυττάρων που έχουν απομονωθεί. Ακολουθεί φυγοκέντρηση στις 13.000 rpm για 1 min στους $+4^{\circ}\text{C}$ και απομάκρυνση του υπερκείμενου. Το ίζημα αναδιαλύεται με ισχυρή ανάδευση 1 ml διαλύματος I και ακολουθεί επώαση στους 4°C για 1 ώρα. Στην συνέχεια, πραγματοποιείται φυγοκέντρηση στις 13.000 rpm για 10 min στους $+4^{\circ}\text{C}$ και το ίζημα επαναδιαλύεται με 50 μl διαλύματος II. Ακολουθεί προσθήκη 1 μl πρωτεΐνάσης K (20 $\mu\text{g}/\mu\text{l}$) και επώαση στους 55°C για 60 min. Για την απενεργοποίηση της πρωτεΐνάσης K πραγματοποιείται επώαση του μείγματος στους 95°C για 10 min. Τέλος ακολουθεί φυγοκέντρηση των σωληναρίων στις 13.000 rpm για 5 min στους $+4^{\circ}\text{C}$ και συλλογή του υπερκείμενου. Το DNA των δειγμάτων αποθηκεύτηκε στους -20°C για μακροπρόθεσμη χρήση.

Διαλύματα:

Διάλυμα I: 10 mM Tris Base, 1mM EDTA

Διάλυμα II: 50 mM KCl, 10 mM Tris Base, 2.5 mM MgCl_2 , 0.5% Tween 20, pH 8.3.

2.2.3 Εκχύλιση DNA (Ιστός μονιμοποιημένος σε παραφίνη)

Για τον τραχηλικό ιστό μονιμοποιημένο σε παραφίνη η απομόνωση του DNA έγινε με την χρήση της πρωτεΐνάσης K. Τομές τραχηλικού ιστού μονιμοποιημένου σε παραφίνη, διαμέτρου 5-10 μm απομακρύνονται με την χρήση νυστεριού και τοποθετούνται σε πλαστικό σωλήνα των 2ml. Ακολουθεί ισχυρή ανάδευση με 400 μl ξυλόλης και φυγοκέντρηση στις 13.000 rpm για 5min. Το υπερκείμενο απομακρύνεται και ακολουθεί ισχυρή ανάδευση του ιζήματος με 400 μl αιθανόλης (96%). Στη συνέχεια, πραγματοποιείται φυγοκέντρηση στις 13.000 rpm για 5min

στους +4°C και απομάκρυνση του υπερκειμένου. Ακολουθεί η προσθήκη 50-200μl διαλύματος εκχύλισης 1X και ολονύχτια επώαση στους 37°C. Για την απενεργοποίηση της πρωτεΐνης K πραγματοποιείται επώαση στους 95°C για 15 min. Ακολουθεί σύντομη φυγοκέντρηση του διαλύματος στους +4°C, συλλογή του υπερκειμένου και αποθήκευση του DNA στους -20°C.

Διαλύματα

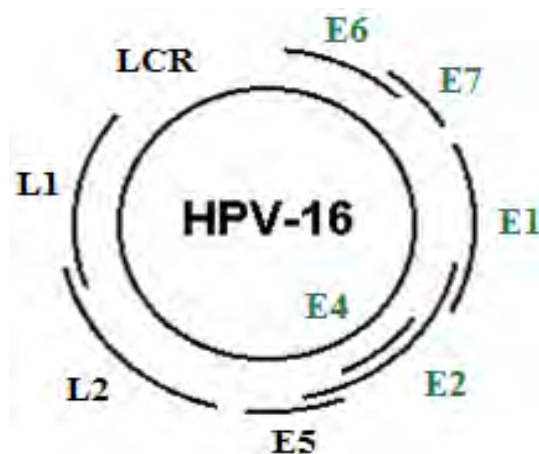
Διάλυμα εκχύλισης: 10μl πρωτεΐνη K (20μg/μl), 100μl PCR ρυθμιστικού διαλύματος 10 X (Stratagene, La Jolla USA), 890μl διπλά απιονισμένο νερό ελεύθερο νουκλεασών (Sigma, USA).

2.3 Έλεγχος της εκχύλισης

Για να επιβεβαιωθεί η παρουσία κυτταρικού DNA στα εκχυλισμένα δείγματα και για να διαπιστωθεί ότι δεν υπάρχουν αναστολείς της PCR, πραγματοποιήθηκε PCR για την ενίσχυση τμήματος του γονιδίου της β-ακτίνης μεγέθους 498 bp (Li et al., 2008). Επιπλέον, τα δείγματα εξετάστηκαν για την παρουσία HPV16 DNA χρησιμοποιώντας την μέθοδο Nested Multiplex PCR (Sotlar et al., 2004).

2.4 Μοριακή ανάλυση του γονιδιώματος του HPV16

Στη παρούσα διδακτορική διατριβή πραγματοποιήθηκε νουκλεοτιδική, αμινοξική, φυλογενετική και εξελικτική μελέτη των γονιδίων E1, E2, E4, E6 και E7 της πρώιμης περιοχής του γονιδιώματος του HPV16. Η ανάλυση πραγματοποιήθηκε, με την βοήθεια της μεθόδου PCR (Polymerase Chain Reaction).



Εικόνα 2.1. Απεικόνιση του γονιδιώματος του HPV16. Με πράσινο χρώμα έχουν τονιστεί τα γονίδια που χρησιμοποιήθηκαν για την μοριακή ανάλυση του γονιδιώματος του ιού.

2.4.1 Σχεδιασμός εκκινητικών μορίων

Τα εκκινητικά μόρια που χρησιμοποιήθηκαν για την ενίσχυση των γονιδίων E1, E2, E4, E6 και E7 προήρθαν από την διεθνή βιβλιογραφία, ή σχεδιάστηκαν στα πλαίσια της παρούσας διδακτορικής διατριβής με τη βοήθεια του προγράμματος Primer3, Whitehead Institute (<http://bioinfo.ut.ee/primer3-0.4.0>). Οι εκκινητές που σχεδιάστηκαν στα πλαίσια της παρούσας εργασίας βασίστηκαν στην πρότυπη αλληλουχία του HPV16 γονιδιώματος, η οποία είναι διαθέσιμη στην τράπεζα δεδομένων NCBI GenBank (www.ncbi.nlm.nih.gov) με αριθμό καταχώρισης NC_001526.

2.4.2 PCR

Όλες οι αντιδράσεις της PCR για την ενίσχυση των E1, E2, E4, E6 και E7 γονιδίων πραγματοποιήθηκαν σε τελικό όγκο 50μl. Σε όλες τις αντιδράσεις χρησιμοποιήθηκαν 3μl DNA από κάθε δείγμα, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 50pmol/μl), 10μl 5X ρυθμιστικού διαλύματος (Green GoTaq Flexi buffer, Promega), 4μl MgCl₂ (25mM), 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 1.25U Taq DNA πολυμεράσης (GoTaq® DNA Polymerase, Promega) και ddH₂O μέχρι τελικό όγκο 50μl.

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολούθησε η εφαρμογή διαφορετικών συνθηκών αποδιάταξης, υβριδοποίησης και επιμήκυνσης, για κάθε ζεύγος ξεχωριστά για 40 κύκλους. Όλες οι αντιδράσεις περιλαμβάνουν ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων (Πίνακας 2.4).

2.4.3 Auto Nested PCR

Για να αυξηθεί το σήμα ενίσχυσης του γονιδίου στόχου λόγω χαμηλού ικού φορτίου, ή για να επιβεβαιωθεί η ύπαρξη ρήξης του γονιδιώματος του HPV16 στις περιπτώσεις των γονιδίων E1 και E2 πραγματοποιήθηκε Auto Nested PCR με την χρήση των ίδιων εκκινητικών μορίων και με DNA μήτρα το προϊόν της πρώτης PCR. Το μείγμα της αντίδρασης περιλαμβάνει 2μl DNA από κάθε ενισχυμένο προϊόν, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 50pmol/μl), 10μl 5X ρυθμιστικού διαλύματος (Green GoTaq Flexi buffer, Promega), 4μl MgCl₂ (25mM),

5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 1.25U Taq DNA πολυμεράσης (GoTaq® DNA Polymerase, Promega) και ddH₂O μέχρι τελικό όγκο 50μl.

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολούθησε η εφαρμογή διαφορετικών συνθηκών αποδιάταξης, υβριδοποίησης και επιμήκυνσης, για κάθε ζεύγος ξεχωριστά για 25 κύκλους. Όλες οι αντιδράσεις περιλαμβάνουν ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων.

Εκκινητές	Αλληλουχία (5'-3')	Θέση	Μέγεθος Προϊόντων	Συνθήκες	Βιβλιογραφία		
E1 γονίδιο							
E1F1	CAGGTACCAATGGGGAAGA	878	645bp	95°C 2 min 95°C 20 sec 50°C 30 sec 72°C 20 sec 72°C 5 min	Arias Pulido et al., 2006 Arias Pulido et al., 2006 Arias Pulido et al., 2006 Arias Pulido et al., 2006		
E1R3	GAAAAACTCACCCCGATAAC	1523	488bp				
E1F4	GTTATACGGGGTGAATTTTC	1503					
E1R5	GCAATTTCACTATCGICTACTATG	1991	533bp				
E1F6	GGGCCTACGATAATGACATAGTA	1952					
E1R7	CCAATGCATTTCTTAAATTGTC	2485					
E1F8	CCCTGTTGGAACATACATAGATGAC	2443	405bp				
E2R9	GTCTATATGGTCACGTAGGTCTGTA	2848					
E2 γονίδιο							
HPV16 2734	GACGAGGACAAGGAAAACGA	2734	1176 bp	95°C 2 min 95°C 40 sec 56°C 30 sec 72°C 40 sec 72°C 5 min	Παρούσα διατριβή Παρούσα διατριβή		
HPV16 3912	GCACACAAAGCAAAGCAAAA	3912	632bp				
5' άκρο E2 γονιδίου							
HPV16 2732	GAGGACGAGGACAAGGAAAA	2732	632bp				
HPV16 3364	CGTTGCTGCTAAACACAGATG	3364					
3' άκρο του E2 γονιδίου							
HPV16 3312	ATGCGGGTGGTCAGGTAAT	3312	598bp				
HPV16 3912	GCACACAAAGCAAAGCAAAA	3912					
E4 γονίδιο							
HPV16 3312	ATGCGGGTGGTCAGGTAAT	3312	598bp	95°C 2 min 95°C 40 sec 50°C 30 sec 72°C 40 sec 72°C 5 min	Παρούσα διατριβή		
HPV16 3912	GCACACAAAGCAAAGCAAAA	3912					
E6-E7 γονίδια							
HPV16 41	AGCGACCCAGAAAGTTACCA	123	716bp	95°C 2 min 95°C 40 sec 55°C 50 sec 72°C 1 min 72°C 5 min	Παρούσα διατριβή		
HPV16 757	ATGGGGCACACAATTCCTAG	839					

Πίνακας 2.4. Παρουσίαση των εκκινητών που χρησιμοποιήθηκαν από τη βιβλιογραφία και των εκκινητών που σχεδιάστηκαν στη παρούσα εργασία με τη βοήθεια του προγράμματος Primer 3, Whitehead Institute (<http://bioinfo.ut.ee/primer3-0.4.0>). Στον πίνακα αναγράφεται η αλληλουχία των εκκινητών και οι διαφορετικές συνθήκες ενίσχυσης των E1, E2, E4, E6 και E7 γονιδίων.

2.4.4 Ηλεκτροφόρηση προϊόντων της PCR

Η επιβεβαίωση των αποτελεσμάτων της PCR και της Auto Nested PCR έγινε με ηλεκτροφόρηση των προϊόντων σε πήκτωμα αγαρόζης (Invitrogen, UK), συγκέντρωσης 2% σε ρυθμιστικό διάλυμα TBE 1X (Tris, Boric acid, EDTA). Στο διάλυμα προστίθεται βρωμιούχου αιθιδίου (EtBr₂), ώστε η τελική του συγκέντρωση να είναι 1μg/ml. Από το προϊόν κάθε αντίδρασης χρησιμοποιήθηκαν 10 μl, τα οποία αναμίχθηκαν με 2μl χρωστικής (κυανό της βρωμοφαινόλης σε 40% w/v σουκρόζης σε TBE) και μεταφέρθηκαν στο πήκτωμα αγαρόζης. Για τον προσδιορισμό του μήκους των επιθυμητών προϊόντων χρησιμοποιήθηκε μάρτυρας μοριακού βάρους (100 bp DNA Ladder, Invitrogen UK). Η ηλεκτροφόρηση πραγματοποιήθηκε σε 120Volts για περίπου 1h. Στη συνέχεια το πήκτωμα αγαρόζης τοποθετείται σε τράπεζα εκπομπής υπεριώδους ακτινοβολίας (Foto UV15, Fotodyne, Hartland WI) και φωτογραφίζεται με ψηφιακή μηχανή (Olympus digital camera).

Διαλύματα:

5X TBE: 5.4% Tris Base, 2.75% H₃BO₃ (Merck, Germany) και 10μM EDTA διαλύονται σε 1 λίτρο απιονισμένο νερό.

Κυανό της βρωμοφαινόλης: 1mM EDTA, 0.25% κυανό της βρωμοφαινόλης (SIGMA, USA), 40% σουκρόζη.

2.4.5 Απομόνωση των προϊόντων από το πήκτωμα αγαρόζης

Ολόκληρη η ποσότητα της αντίδρασης της PCR ή της Auto Nested PCR ηλεκτροφορήθηκε σε πήκτωμα αγαρόζης 2% που περιείχε βρωμιούχο εθίδιο σε συγκέντρωση 1μg/ml. Οι επιθυμητές ζώνες απομακρύνθηκαν από το πήκτωμα και μεταφέρθηκαν σε πλαστικούς σωλήνες των 2ml. Ακολούθησε καθαρισμός των προϊόντων από το πήκτωμα αγαρόζης με τη χρήση του Gel extraction kit (Macherey–Nagel, Germany), σύμφωνα με τις οδηγίες του κατασκευαστή. Όλα τα προϊόντα της αντίδρασης υποβλήθηκαν σε μοριακή κλωνοποίηση.

2.4.6 Μοριακή κλωνοποίηση των ενισχυμένων γονιδίων και αλληλούχιση

Για να μελετηθούν οι αλληλουχίες των γονιδίων E1, E2, E4 E6 και E7, δύο ανεξάρτητες PCR πραγματοποιήθηκαν για κάθε περιοχή, για κάθε κλινικό τραχηλικό

δείγμα ξεχωριστά. Στη συνέχεια, τα προϊόντα από τις δύο ανεξάρτητες αντιδράσεις υποβλήθηκαν σε μοριακή κλωνοποίηση και εξετάστηκαν τρεις κλώνοι από κάθε διαφορετικό ενισχυμένο προϊόν. Η διαδικασία αυτή πραγματοποιήθηκε για κάθε γονίδιο ξεχωριστά. Η μοριακή κλωνοποίηση έγινε για την αποφυγή ύπαρξης μειγμάτων διαφορετικών στελεχών του HPV16 καθώς επίσης και τον εντοπισμό ταυτόχρονης μόλυνσης από περισσότερο από μία ενδοτυπική παραλλαγή του ιού.

Η διαδικασία της κλωνοποίησης πραγματοποιήθηκε με το πλασμιδιακό φορέα pGEM T-easy vector System (Promega, Madison, USA), ο οποίος εισήχθη σε έτοιμα δεκτικά κύτταρα *Escherichia coli* DH5a. Η απομόνωση του ανασυνδυασμένου πλασμιδιακού φορέα από μετασχηματισμένες βακτηριακές αποικίες πραγματοποιήθηκε με το Nucleospin plasmid kit (Macherey-Nagel GmbH, Duren, Germany), σύμφωνα με τις οδηγίες του κατασκευαστή. Στη συνέχεια πραγματοποιήθηκε αλληλούχιση και των δύο κλώνων του ενθέματος από την εταιρία Macrogen (Seoul, Korea). Για κάθε αντίδραση αλληλούχισης των κλωνοποιημένων δειγμάτων χρησιμοποιήθηκαν ως εκκινητές οι επαγωγείς T7 και SP6 του πλασμιδιακού φορέα που βρίσκονται εκατέρωθεν της θέσης ένθεσης του τμήματος DNA.

2.4.7 Έλεγχος των αλληλουχιών

Οι αλληλουχίες παραλήφθησαν υπό την μορφή χρωματογραφήματος. Το χρωματογράφημα κάθε αλληλουχίας ελέγχθηκε για την ποιότητα του και για την ύπαρξη λαθών που ίσως πραγματοποιήθηκαν κατά την διάρκεια της αλληλούχισης. Οι κλωνοποιημένες αλληλουχίες κάθε δείγματος εξετάστηκαν ως προς την ειδικότητα τους μέσω ομοπαράθεσης με την πρότυπη αλληλουχία του γονιδιώματος του HPV16 (HPV16R), η οποία είναι διαθέσιμη στη νουκλεοτιδική βάση δεδομένων (Los Alamos National Laboratory). Η ομοπαράθεση πραγματοποιήθηκε με την χρήση του αλγόριθμου Muscle (Edgar et al., 2004) και την βοήθεια του προγράμματος MEGA v.5 (Tamura et al., 2007). Στη συνέχεια, οι αλληλουχίες κάθε δείγματος μελετήθηκαν ως προς την μεταξύ τους ομοιότητα για κάθε γονίδιο του HPV16 ξεχωριστά (E1, E2, E4, E6 και E7). Μια αντιπροσωπευτική αλληλουχία χρησιμοποιήθηκε στις περιπτώσεις που οι κλωνοποιημένες αλληλουχίες του δείγματος παρουσίαζαν μεταξύ τους ομοιότητα 100% για μια συγκεκριμένη περιοχή. Στην συνέχεια,

πραγματοποιήθηκε φυλογενετική και εξελικτική ανάλυση όλων των περιοχών που εξετάστηκαν στην παρούσα διατριβή, ενώ παράλληλα εντοπίστηκαν νουκλεοτιδικές και αμινοξικές αλλαγές, οι οποίες συσχετίστηκαν με τον βαθμό της ενδοεπιθηλιακής αλλοίωσης και την φυλογενετική προέλευση του ιού.

2.4.8 Ομοπαράθεση νουκλεοτιδικών και αμινοξικών αλληλουχιών

Για την ανίχνευση νουκλεοτιδικών και αμινοξικών μεταλλάξεων πραγματοποιήθηκε για κάθε γονίδιο ξεχωριστά (E1, E2, E4, E6 και E7) πολλαπλή ομοπαράθεση μεταξύ των κλωνοποιημένων αλληλουχιών των δειγμάτων με την πρότυπη αλληλουχία του HPV16 γονιδιώματος (HPV16R). Η πολλαπλή ομοπαράθεση έγινε με την χρήση του αλγόριθμου Muscle (Edgar et al., 2004) και την βοήθεια του προγράμματος MEGA v.5 (Tamura et al., 2007).

2.4.9 Κατασκευή φυλογενετικών δέντρων

Η κατασκευή φυλογενετικών δέντρων πραγματοποιήθηκε με τη μέθοδο Μέγιστης Πιθανοφάνειας (Maximum Likelihood, ML) και σε ορισμένες περιπτώσεις με τη μέθοδο της Μπεϋζιανής Συμπερασματολογίας (Bayesian Interference, BI) και της Μέγιστης Φειδωλότητας (Maximum Parsimony, MP). Οι συγκεκριμένες μέθοδοι υπολογίζουν δεδομένα με διαφορετικές καταστάσεις πληροφοριακών χαρακτήρων. Με τις συγκεκριμένες μεθόδους επιλέγονται κάθε φορά οι μικρότερες εξελικτικές αποστάσεις (νουκλεοτιδικές ή αμινοξικές υποκαταστάσεις) για την κατασκευή φυλογενετικού δέντρου, σε σχέση με όλα τα πιθανά δέντρα και θεωρείται το πιο πιθανό σωστό δέντρο. Τα φυλογενετικά δέντρα που προέκυψαν με τις μεθόδους Μέγιστης Πιθανοφάνειας και Μέγιστης Φειδωλότητας, αξιολογήθηκαν με την μέθοδο της επανασυμπίπτουσας ιχνηλάτησης (bootstrap ανάλυση). Η επανασυμπίπτουσα ιχνηλάτηση είναι μια στατιστική μέθοδος, κατά την οποία χρησιμοποιούνται πολλά τυχαία μέρη δειγματικών χαρακτήρων, όπως λόγου χάρη νουκλεοτιδικές θέσεις και γίνονται πολλές επαναλαμβανόμενες αναλύσεις, με νέα τεχνητά δεδομένα, για τον υπολογισμό της πιστότητας ενός δέντρου. Η μέθοδος αυτή συναντά την πιθανότερη πραγματική κατάσταση στα δέντρα που δεν αλλάζουν ή αλλάζουν λίγο και για αυτό αντανακλούν την αναζητούμενη πιστότητα. Τα φυλογενετικά δέντρα Μέγιστης

Πιθανοφάνειας και Μέγιστης Φειδωλότητας κατασκευάστηκαν και αξιολογήθηκαν με το πρόγραμμα MEGA v.5 (1000 bootstrap values).

Αντίθετα από τις μεθόδους Μέγιστης Φειδωλότητας και Μέγιστης Πιθανοφάνειας, η Μπεϋζιανή Συμπερασματολογία δεν παράγει ένα μόνο δέντρο ή ένα σύνολο δέντρων. Η Μπεϋζιανή Συμπερασματολογία χρησιμοποιεί την πιθανότητα των δέντρων με τη χρήση του αλγόριθμου MCMC (Markov Chain Monte Carlo) για ένα δείγμα δέντρων σε αναλογία με την πιθανότητά τους, έτσι ώστε να παραχθεί ένα αξιόπιστο δείγμα δέντρων. Τα φυλογενετικά δέντρα που προέκυψαν με την Μπεϋζιανή Συμπερασματολογία αξιολογήθηκαν με την εκτίμηση των εκ των υστέρων πιθανοτήτων (Posterior Probabilities). Η εκ των υστέρων πιθανότητα εκφράζει την πιθανότητα κάθε δέντρου δεδομένου α) της εκ των προτέρων γνώσης, β) ενός μοντέλου νουκλεοτιδικής υποκατάστασης, γ) των παρατηρούμενων δεδομένων.

Για την κατασκευή φυλογενετικών δέντρων με την μέθοδο της Μπεϋζιανής Συμπερασματολογίας πραγματοποιήθηκε Μπεϋζιανή ανάλυση χρησιμοποιώντας τον αλγόριθμο MCMC (Markov Chain Monte Carlo) με την βοήθεια του προγράμματος BEAST v.1.6.2 (Drummond, 2007) για την εκτίμηση των εκ των υστέρων πιθανοτήτων. Ο αλγόριθμος MCMC έχει την ιδιότητα να συγκλίνει προς μία κατάσταση ισορροπίας ανεξάρτητα από το σημείο εκκίνησης. Τα αποτελέσματα της ανάλυσης αξιολογήθηκαν με την χρήση του προγράμματος Tracer v1.5, εντός του προγράμματος BEAST v.1.6.2. Η απόκλιση των αποτελεσμάτων πραγματοποιήθηκε με τις τιμές ESS (Effective Sample Size). Στη συνέχεια, το φυλογενετικό δέντρο κατασκευάστηκε με το πρόγραμμα TreeAnnotator V.1.6.2 και FigTree v.1.3.1, το οποίο διατίθεται στο πρόγραμμα BEAST v.1.6.2. Η Μπεϋζιανή ανάλυση και η μέθοδος Μέγιστης Πιθανοφάνειας πραγματοποιήθηκαν με την χρήση του κατάλληλου μοντέλου νουκλεοτιδικής υποκατάστασης, όπως αυτό υπολογίστηκε από το πρόγραμμα FindModel, το οποίο διατίθεται στο διαδίκτυο (<http://.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>). Τα μοντέλα νουκλεοτιδικής υποκατάστασης μετατρέπουν την παρατηρούμενη απόσταση σε πραγματική εξελικτική απόσταση.

Η κατασκευή φυλογενετικών δέντρων πραγματοποιήθηκε για κάθε περιοχή του γονιδιώματος του HPV16 ξεχωριστά (E1, E2, E4, E6, E7). Η φυλογενετική ανάλυση έγινε μετά από πολλαπλή ομοπαράθεση, μεταξύ των αλληλουχιών κάθε γονιδίου, της πρότυπης αλληλουχίας του γονιδιώματος του HPV16 και των πρότυπων αλληλουχιών

των ενδοτυπικών παραλλαγών του HPV16 γονιδιώματος. Η διαδικασία αυτή πραγματοποιήθηκε με στόχο να ομαδοποιηθούν οι αλληλουχίες στις πέντε διαφορετικές ενδοτυπικές γενιές παραλλαγών του γονιδιώματος του HPV16 και να ενσωματωθούν στα τέσσερα διαφορετικά φυλογενετικά κλαδιά. Επιπλέον εξετάστηκε εάν όλες οι περιοχές του ιού μπορούν να συμβάλουν στον ξεκάθαρο διαχωρισμό των τεσσάρων διαφορετικών φυλογενετικών κλαδιών που σχηματίζονται από τις διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του ιού.

2.4.10 Ανάλυση επιλεκτικής πίεσης

Η αξιολόγηση της επιλεκτικής πίεσης που επιδρά στα γονίδια E1, E2, E4, E6 και E7, πραγματοποιήθηκε μετά από πολλαπλή ομοπαράθεση, μεταξύ των αλληλουχιών κάθε γονιδίου ξεχωριστά και της πρότυπης αλληλουχίας του γονιδιώματος του HPV16. Ο προσδιορισμός της επιλεκτικής πίεσης πραγματοποιήθηκε με τον υπολογισμό του ρυθμού dN/dS. Ο ρυθμός dN/dS είναι μία διαδεδομένη στατιστική μέθοδος που χρησιμοποιείται ευρέως, λόγω της απλότητας και της αξιοπιστίας της. Ο λόγος dN/dS είναι ο ρυθμός που προσδιορίζει τον αριθμό των μη συνώνυμων υποκαταστάσεων ανά μη-συνώνυμη θέση (dN) προς τον αριθμό των συνώνυμων υποκαταστάσεων ανά συνώνυμη θέση (dS) (Kryazhimskiy, 2008). Οι υποκαταστάσεις σε συνώνυμες (σιωπηλές) και μη συνώνυμες (μη σιωπηλές) θέσεις διαφοροποιούνται μεταξύ ουδέτερων και ενεργών εξελικτικών δυνάμεων που ασκούνται σε γονίδια. Ο ρυθμός dN/dS δίνει ένδειξη για τον τρόπο και την δύναμη της εξέλιξης. Πιο συγκεκριμένα, περίσσεια μη συνώνυμων μεταλλάξεων (dN/dS > 1) αποτελεί ένδειξη θετικής επιλεκτικής πίεσης (δρά η φυσική επιλογή), ενώ περίσσεια συνώνυμων μεταλλάξεων (dN/dS < 1) δείχνει αρνητική επιλεκτική πίεση (αναστολή της φυσικής επιλογής). Επιπλέον, καμία διαφορά μεταξύ συνώνυμων και μη συνώνυμων μεταλλάξεων (dN/dS = 1) αποτελεί ένδειξη ουδετερότητας (τυχαία σταθεροποίηση ουδέτερων μεταλλάξεων) (Yang et al., 2002, Nielsen, 2005, Hurst, 2009, Akashi et al., 2012).

Ο ρυθμός dN/dS υπολογίστηκε με την μέθοδο της Μέγιστης Πιθανοφάνειας μέσω του προγράμματος SLAC (Single Likelihood Ancestor Counting) που είναι διαθέσιμο στον διακομιστή Datamonkey (Suzuki et al., 1999, Pond et al., 2005). Επιπλέον, για να προσδιοριστεί η ύπαρξη θετικής επιλεκτικής πίεσης (dN/dS > 1) σε

συγκεκριμένα κωδικόνια εντός των γονιδίων, χρησιμοποιήθηκε η μέθοδος Μέγιστης Πιθανοφάνειας και Μπεϋζιανής Συμπερασματολογίας. Πιο συγκεκριμένα, ο προσδιορισμός θετικής επιλεκτικής πίεσης με την μέθοδο Μέγιστης Πιθανοφάνειας πραγματοποιήθηκε με το πρόγραμμα FEL (Fixed Effects Likelihood model) στον διακομιστή Datamonkey, ενώ ο προσδιορισμός θετικής επιλεκτικής πίεσης με την Μπεϋζιανή Συμπερασματολογία πραγματοποιήθηκε με την βοήθεια του διακομιστή Selecton 2.4 (Stern et al., 2007).

Για να εντοπιστούν σήματα ανασυνδυασμού χρησιμοποιήθηκε το πρόγραμμα SBP (Single Recombination Breakpoint), το οποίο είναι διαθέσιμο στο διακομιστή Datamonkey. Για την οπτική αναπαράσταση του ανασυνδυασμού χρησιμοποιήθηκε το πρόγραμμα SimPlot v.3.5.1 (Ray, 1998). Η ποσοστιαία ομοιότητα των ανασυνδυασμένων αλληλουχιών με τις πρότυπες αλληλουχίες των διαφορετικών ενδοτυπικών παραλλαγών του HPV16 που συμμετέχουν στον ανασυνδυασμό πραγματοποιήθηκε με τη μέθοδο ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2/analysis).

2.4.11 Προσδιορισμός του μέσου ρυθμού υποκατάστασης των γονιδίων

Επιπλέον, προσδιορίστηκε ο μέσος ρυθμός υποκατάστασης για το πρώτο δεύτερο και τρίτο νουκλεοτίδιο όλων των κωδικονίων ενός γονιδίου. Για την διαδικασία αυτή πραγματοποιήθηκε Μπεϋζιανή ανάλυση, χρησιμοποιώντας τον αλγόριθμο MCMC (Markov Chain Monte Carlo) με την βοήθεια του προγράμματος BEAST v.1.6.2. Τα αποτελέσματα της ανάλυσης αξιολογήθηκαν με την χρήση του προγράμματος Tracer v1.5 εντός του προγράμματος BEAST v.1.6.2. Η απόκλιση των αποτελεσμάτων πραγματοποιήθηκε με τις τιμές ESS (Effective Sample Size). Η Μπεϋζιανή ανάλυση πραγματοποιήθηκε με την χρήση του κατάλληλου μοντέλου υποκατάστασης όπως αυτό υπολογίστηκε από το πρόγραμμα FindModel (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html) για κάθε γονίδιο ή περιοχή γονιδίου ξεχωριστά.

2.5 Προσδιορισμός της μορφής του γονιδιώματος του HPV16

Ο προσδιορισμός της μορφής του γονιδιώματος του HPV16 πραγματοποιήθηκε σε δύο βήματα. Πρώτον, προσδιορίστηκαν οι πιο συχνές θέσεις ρήξης του γονιδιώματος του ιού, μέσω χαρτογράφησης των E1 και E2 γονιδίων. Δεύτερον, οι πιο κοινές θέσεις ρήξης του HPV16 DNA χρησιμοποιήθηκαν για να προσδιοριστεί ο λόγος των τιμών E1/E6 και E2/E6. Με βάση τον λόγο των αντιγράφων των γονιδίων, προσδιορίστηκε το εύρος των τιμών που αντιπροσωπεύουν την μορφή του γονιδιώματος του ιού (επισωματική, ενσωματωμένη, ή μεικτή μορφή).

2.5.1 Χαρτογράφηση των E1 και E2 γονιδίων

Ο προσδιορισμός των θέσεων ρήξης εντός των E1 και E2 γονιδίων πραγματοποιήθηκε με την μέθοδο της PCR, χρησιμοποιώντας αλληλεπικαλυπτόμενα εκκινητικά μόρια από την βιβλιογραφία. Τα εκκινητικά μόρια που χρησιμοποιήθηκαν ενισχύουν σε επιμέρους τμήματα ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης των E1 και E2 γονιδίων, με σκοπό να προσδιοριστούν τα σημεία στα οποία το γονιδίωμα του ιού έχει υποστεί ρήξη, λόγω ενσωμάτωσης του HPV16 DNA στο κυτταρικό γονιδίωμα (Πίνακας 2.5).

Όλες οι αντιδράσεις της PCR για την ενίσχυση των επιμέρους τμημάτων των E1 και E2 γονιδίων, πραγματοποιήθηκαν σε τελικό όγκο 50μl. Σε όλες τις αντιδράσεις χρησιμοποιήθηκαν 3μl DNA από κάθε δείγμα, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 50pmol/μl), 5μl 10X ρυθμιστικού διαλύματος (Stratagene, La Jolla, USA) περιεκτικότητας 2 mM MgCl₂, 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 2.5U Paq DNA πολυμεράσης (Paq5000™ DNA Polymerase, Stratagene, La Jolla, USA) και ddH₂O μέχρι τελικό όγκο 50μl.

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολούθησε η εφαρμογή διαφορετικών συνθηκών αποδιάταξης, υβριδοποίησης και επιμήκυνσης, για κάθε ζεύγος ξεχωριστά για 40 κύκλους. Όλες οι αντιδράσεις περιλαμβάνουν ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων (Πίνακας 2.5). Η επιβεβαίωση των αποτελεσμάτων έγινε σε πήκτωμα αγαρόζης συγκέντρωσης 2%.

Εκκινητές	Αλληλουχία (5'-3')	Θέση	Μέγεθος Προϊόντων	Συνθήκες	Βιβλιογραφία	
E2 γονίδιο						
a1	AGGACGAGGACAAGGAAAA	2735	475 bp	95°C 2 min 95°C 20 sec 50°C 30 sec 72°C 20 sec 72°C 5 min	Li et al., 2008	
a2	ACTTGACCCTCTACCACAGTACT	3210				
b1	TTGTGAAGAAGCATCAGTAACT	3172				
b2	TAAAGTATTAGCATCACCTT	3649	477 bp		40 cycles	Li et al., 2008
c1	GTAATAGTAACTACACCCATA	3597	276 bp			
c2	GGATGCAGTATCAAGATTTGTT	3873				Li et al., 2008
E1 γονίδιο						
E1F1	CAGGTACCAATGGGGAAGA	878	241bp	95°C 2 min 95°C 20 sec 55°C 30 sec 72°C 10 sec 72°C 5 min	Arias Pulido et al., 2006	
E1R1	ACCTGTA CTGCATCTCTATGTTGT	1118				
E1F2	GACAGCACATGCGTGTGTTAC	1059	265bp			Arias Pulido et al., 2006
E1R2	AGTCTCATGGCGCCCTTC	1323				
E1F3	CGCCATGAGACTGAAACAC	1312	212bp			Arias Pulido et al., 2006
E1R3	GAAAACTCACCCCGTATAAC	1523				
E1F4	GTTATACGGGGTGAGTTTTTC	1503	282bp			Arias Pulido et al., 2006
E1R4	ATCATCATACACATTGGAGACA	1784				
E1F5	GTGTGTCTCCAATGTGTATGATG	1760	232bp			Arias Pulido et al., 2006
E1R5	GCAATTTCACTATCGTCTACTATG	1991				
E1F6	GGCCTACGATAATGACATAGTA	1952	235bp			Arias Pulido et al., 2006
E1R6	TGCTTCCAATCACCTCCAT	2186				
E1F7	GGAGGTGATTGGAAGCAA	2170	316bp			Arias Pulido et al., 2006
E1R7	CCAATGCATTTCTTAAATTGTC	2485				
E1F8	CCCTGTTGGA ACTACATAGATGAC	2443	301bp			Arias Pulido et al., 2006
E1R8	CCTCGTCCTCGTGCAA ACT	2743				
E1F9	CCAGTGTATGAGCTTAAATGATAAGA	2662	187bp			Arias Pulido et al., 2006
E2R9	GTCTATATGGTACCGTAGGTCTGTA	2848				

Πίνακας 2.5. Παρουσίαση των αλληλοεπικαλυπτόμενων εκκινητών που χρησιμοποιήθηκαν από τη βιβλιογραφία για την ενίσχυση ολόκληρων των γονιδίων E1 και E2. Στον πίνακα αναγράφεται η αντίστοιχη αλληλουχία των εκκινητών και οι διαφορετικές συνθήκες ενίσχυσης των επιμέρους τμημάτων των E1 και E2 γονιδίων.

2.5.2 Έλεγχος ακεραιότητας του HPV16 γονιδιώματος

Για να ελέγχθη εάν το DNA του HPV16 έχει κατακερματιστεί, λόγω επεξεργασίας των δειγμάτων (κυρίως για τις περιπτώσεις τραχηλικού ιστού μονιμοποιημένου σε παραφίνη) και να εξακριβωθεί ότι η ρήξη προέρχεται από την ενσωμάτωση του ιού στο κυτταρικό DNA, πραγματοποιήθηκε ενίσχυση τμήματος του E6 γονιδίου, μεγέθους 354 ζευγών βάσεων σε όλα τα δείγματα με την μέθοδο της PCR. Το τμήμα του E6 γονιδίου ενισχύθηκε με την χρήση των εκκινητικών μορίων HPV-16 1 5'-GCACAGAGCTGCAAACAACACT-3' / HPV-16 353 5'-GGTCCACCGACCCCTTATATTT-3', οι οποίοι σχεδιάστηκαν στη παρούσα διατριβή.

Η αντίδραση της PCR για την ενίσχυση του τμήματος του E6 γονιδίου, πραγματοποιήθηκε σε τελικό όγκο 50μl. Στην αντίδραση χρησιμοποιήθηκαν 3μl DNA από κάθε δείγμα, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 50pmol/μl), 5μl 10X ρυθμιστικού διαλύματος (Stratagene, La Jolla, USA) περιεκτικότητας 2 mM MgCl₂, 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 2.5U Paq DNA πολυμεράσης (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA) και ddH₂O μέχρι τελικό όγκο 50μl.

Η αντίδραση πραγματοποιήθηκε με ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολούθησαν 40 κύκλοι αποδιάταξης στους 95°C για 40sec, υβριδοποίησης στους 58°C για 20sec και επιμήκυνσης στους 72 °C για 20sec. Τέλος η αντίδραση περιλαμβάνει ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων. Η επιβεβαίωση των αποτελεσμάτων έγινε σε πήκτωμα αγαρόζης συγκέντρωσης 2%.

2.5.3 Προσδιορισμός ευαισθησίας της μεθόδου.

Ο προσδιορισμός της ευαισθησίας των αντιδράσεων της PCR για την ενίσχυση των E1, E2 και E6 γονιδίων, πραγματοποιήθηκε με την χρήση του ανασυνδυασμένου πλασμιδίου pE6-E2, που περιέχει την γενωμική περιοχή μεταξύ των γονιδίων E6 έως και E2 (θέση 123 έως 3912).

Για τη κατασκευή του πλασμιδίου pE6-E7 ενισχύθηκε η αλληλουχία μεταξύ των νουκλεοτιδίων 123 έως 3912 με την βοήθεια των εκκινητικών μορίων HPV-16 41 5'-AGCGACCCAGAAAGTTACCA-3' / HPV-16 3912

GCACACAAAGCAAAGCAAAA. Τα εκκινητικά μόρια σχεδιάστηκαν στη παρούσα διατριβή και η αντίδραση πραγματοποιήθηκε σε εκχυλισμένο δείγμα που χαρακτηρίστηκε ως πρότυπο στέλεχος (796)

Η αντίδραση της PCR για την ενίσχυση της περιοχής μεταξύ των γονιδίων E6 και E2 πραγματοποιήθηκε σε τελικό όγκο 50μl. Η αντίδραση περιείχε 3μl DNA, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 50pmol/μl), 5μl 10X ρυθμιστικού διαλύματος (GeneON) περιεκτικότητας 1.5 mM MgCl₂, 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 2.5U Taq DNA πολυμεράσης (Maximo Taq DNA Polymerase, GeneON) και ddH₂O μέχρι τελικό όγκο 50μl.

Η αντίδραση πραγματοποιείται με ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 5min και ακολουθούν 40 κύκλοι αποδιάταξης στους 95°C για 40sec, υβριδοποίησης στους 55°C για 40sec και επιμήκυνσης στους 72 °C για 4min. Τέλος η αντίδραση περιλαμβάνει ένα τελικό στάδιο επιμήκυνσης στους 72°C για 10min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων.

Στη συνέχεια, το ενισχυμένο προϊόν μεγέθους 3791 ζευγών βάσεων υποβλήθηκε σε μοριακή κλωνοποίηση με την χρήση του StataClone PCR cloning kit (Stratagene, Agilent technologies), σύμφωνα με τις οδηγίες του κατασκευαστή. Η απομόνωση του ανασυνδυασμένου πλασμιδιακού φορέα από μετασχηματισμένες βακτηριακές αποικίες πραγματοποιήθηκε με το Nucleospin plasmid kit (Macherey-Nagel GmbH, Duren, Germany), σύμφωνα με τις οδηγίες του κατασκευαστή. Στη συνέχεια, πραγματοποιήθηκε αλληλούχιση και των δύο κλώνων του ενθέματος από την εταιρία Macrogen (Seoul, Korea) για να διαπιστωθεί η ειδικότητα του προϊόντος. Για την αντίδραση της αλληλούχισης χρησιμοποιήθηκαν ως εκκινητές οι επαγωγείς T3 και T7 του πλασμιδιακού φορέα που βρίσκονται εκατέρωθεν της θέσης ένθεσης του τμήματος DNA. Ο αριθμός των αντιγράφων του pE6-E2 προσδιορίστηκε μέσω Ποσοτικής PCR Πραγματικού Χρόνου, ενισχύοντας τμήμα του γονιδίου E6 (ενότητα 2.5.2).

Σειριακές υποδεκαλάσιες αραιώσεις από 10⁶ έως 1 αντίγραφο του ανασυνδυασμένου πλασμιδίου (pE6-E2) υποβλήθηκαν σε PCR με όλα τα ζεύγη εκκινητικών μορίων που χρησιμοποιήθηκαν στη μέθοδο χαρτογράφησης των E1 και E2 γονιδίων, ξεχωριστά. Οι συνθήκες που εφαρμόστηκαν για κάθε ζεύγος εκκινητών περιγράφηκαν στην ενότητα 2.5.1 (Πίνακας 2.5). Η διαδικασία πραγματοποιήθηκε

για να προσδιοριστεί και να βελτιωθεί η ευαισθησία της μεθόδου και εφαρμόστηκε τρεις φορές για κάθε ζεύγος εκκινητικών μορίων.

2.6 Προσδιορισμός της μορφής του γονιδιώματος του HPV16 με Ποσοτική PCR Πραγματικού Χρόνου (quantitative - Real Time PCR)

Ο προσδιορισμός της μορφής του γονιδιώματος του HPV16 (επισωμακή μορφή, ενσωματωμένη μορφή, μεικτή μορφή) πραγματοποιήθηκε με την μέθοδο της Ποσοτικής PCR Πραγματικού Χρόνου και τη χρήση φθορίζουσας χρωστικής SYBR Green I. Για να προσδιοριστεί η μορφή του γονιδιώματος του ιού πραγματοποιήθηκε ποσοτικοποίηση των γονιδίων E1, E2, E6 και υπολογίστηκε ο λόγος των τιμών E1/E6 και E2/E6 για κάθε δείγμα ξεχωριστά. Η βασική αρχή της μεθόδου βασίζεται στην υπόθεση ότι τα γονίδια E1, E2 και E6 παρουσιάζουν ίση ποσότητα αντιγράφων όταν ο ιός βρίσκεται στην επισωμακή του μορφή, ενώ το γεγονός αυτό αναιρείται όταν το γονιδίωμα του ιού υφίσταται ρήξη στη περιοχή των γονιδίων E1 ή/και E2 κατά την διαδικασία ενσωμάτωσης του ιού στο γονιδίωμα του κυττάρου ξενιστή.

Η επιλογή εκκινητικών μορίων για την ποσοτικοποίηση των γονιδίων E1 και E2 βασίστηκε στον εντοπισμό θέσεων που παρουσιάζουν με μεγαλύτερη συχνότητα ρήξη, όπως προέκυψε από την χαρτογράφηση των θέσεων ρήξης των γονιδίων E1 και E2. Πιο συγκεκριμένα, η πιο συχνή θέση ρήξης του E2 γονιδίου εντοπίζεται στην περιοχή μεταξύ των νουκλεοτιδίων 3172 και 3649, ενώ για το E1 γονίδιο η πιο συχνή θέση εντοπίζεται στο 5' άκρο του γονιδίου μεταξύ των νουκλεοτιδίων 1059 και 1323. Επομένως, για τον ποσοτικό προσδιορισμό του E2 γονιδίου χρησιμοποιήθηκε ζεύγος εκκινητών, που ενισχύει τη περιοχή μεταξύ των νουκλεοτιδίων 3243 και 3539 (E2F12-E2R12), ενώ η ποσοτικοποίηση του E1 γονιδίου πραγματοποιήθηκε με ζεύγος εκκινητών που ενισχύουν την περιοχή μεταξύ των νουκλεοτιδίων 1059 και 1323 (E1F2-E1R2) (Πίνακας 2.6). Οι αντιδράσεις σε κάθε δείγμα πραγματοποιήθηκαν σε συνολικό αριθμό 500 αντιγράφων κυτταρικού DNA, όπως αυτό υπολογίστηκε μετά την ποσοτικοποίηση τμήματος του γονιδίου GAPDH.

2.6.1 Κατασκευή πλασμιδίων

Για τον προσδιορισμό των αντιγράφων των E1, E2, E6 και GAPDH γονιδίων με την μέθοδο της Ποσοτικής PCR Πραγματικού Χρόνου δημιουργήθηκαν πρότυπες

καμπύλες με την χρήση ανασυνδυασμένων πλασμιδίων. Δύο διαφορετικά ανασυνδυασμένα πλασμίδια κατασκευάστηκαν. Ένα πλασμίδιο που περιέχει την περιοχή μεταξύ των νουκλεοτιδίων E6 και E2 (pE6-E2) (ποσοτικοποίηση E1, E2 και E6) και ένα πλασμίδιο που φέρει τμήμα του GAPDH γονιδίου (pGAPDH) (ποσοτικοποίηση του κυτταρικού DNA).

Για τη κατασκευή του πλασμιδίου pE6-E2 πραγματοποιήθηκε η διαδικασία που αναφέρεται στην ενότητα 2.5.3. Για τη κατασκευή του πλασμιδίου pGAPDH, πραγματοποιήθηκε PCR για την ενίσχυση τμήματος του GAPDH γονιδίου σε εκχυλισμένο δείγμα κυτταρικής σειράς CaSki (Li et al., 2008). Στη συνέχεια το προϊόν της αντίδρασης υποβλήθηκε σε μοριακή κλωνοποίηση με το StataClone PCR cloning kit (Stratagene, Agilent technologies), σύμφωνα με τις οδηγίες του κατασκευαστή.

Επιπρόσθετα, για την πραγματοποίηση πειραμάτων προσομοίωσης (ενότητα 2.6.3), κατασκευάστηκε ένα ακόμα πλασμίδιο μετά από ενίσχυση τμήματος του E6 γονιδίου προερχόμενο από το κλινικό δείγμα 796, το οποίο περιέχει το πρότυπο στέλεχος του ιού. Το πλασμίδιο αυτό κατασκευάστηκε μετά από ενίσχυση τμήματος του E6 γονιδίου με την χρήση των εκκινητικών μορίων HPV16 1 5'-GCACAGAGCTGCAAACAACACT-3' / HPV16 353 5'-GGTCCACCGACCCCTTATATT-3', όπως περιγράφηκε στην ενότητα 2.5.2. Στη συνέχεια το προϊόν της αντίδρασης υποβλήθηκε σε μοριακή κλωνοποίηση με το StataClone PCR cloning kit (Stratagene, Agilent technologies), σύμφωνα με τις οδηγίες του κατασκευαστή.

2.6.2 Ποσοτική PCR Πραγματικού Χρόνου

Η κατασκευή πρότυπης καμπύλης για τα γονίδια E1, E2, E6 και GAPDH πραγματοποιήθηκε ξεχωριστά για κάθε γονίδιο στόχο. Όλες οι αντιδράσεις πραγματοποιήθηκαν με το SYBR® FAST qPCR Kit (KAPA, Boston, Massachusetts, United States) και στη συσκευή Mx3000P® (Stratagene). Κάθε αντίδραση περιελάμβανε 3μl πλασμιδιακού DNA, 2μl εκκινητικών μορίων (1μl απο τον καθένα με συγκέντρωση 10pmol/μl), 10μl 2X μείγματος αντίδρασης (KAPA SYBR® FAST qPCR Master Mix), 0.4μl χρωστικής αναφοράς ROX και ddH₂O μέχρι τελικό όγκο 25μl.

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 1min και ακολουθούν 40 κύκλοι αποδιάταξης στους 95°C για 3sec, υβριδοποίησης στους 60°C για 20sec και επιμήκυνσης στους 60°C για 10sec. Τέλος η αντίδραση περιλαμβάνει ένα τελικό στάδιο κατασκευής καμπύλης τήξης των ενισχυμένων προϊόντων, που κυμαίνεται από 55°C έως 95°C. Η συλλογή των αποτελεσμάτων έγινε στους 72°C στα 510 nm (Πίνακας 2.6).

Η πρότυπη καμπύλη που κατασκευάστηκε για το γονίδιο GAPDH πραγματοποιήθηκε μέσω ενίσχυσης υποδεκαπλάσιων αραιώσεων του πλασμιδίου pGAPDH, που κυμαίνονταν από 1 έως 10⁸ αντίγραφα του γονιδίου. Επιπλέον τρεις διαφορετικές πρότυπες καμπύλες κατασκευάστηκαν για τους στόχους των E1, E2 και E6 γονιδίων χρησιμοποιώντας το πλασμίδιο pE6-E2, το οποίο περιλαμβάνει την περιοχή μεταξύ των γονιδίων E6 και E2. Οι πρότυπες καμπύλες κατασκευάστηκαν με την χρήση υποδεκαπλάσιων αραιώσεων του πλασμιδίου pE6-E2, που κυμαίνονταν από 1 έως 10⁸ αντίγραφα του πλασμιδίου. Όλες οι αντιδράσεις για την κατασκευή των πρότυπων καμπυλών πραγματοποιήθηκαν τρεις φορές, ενώ όλα τα κλινικά δείγματα εξετάστηκαν εις διπλούν. Ο αριθμός των αντιγράφων των E1, E2 και E6 γονιδίων καθορίστηκε για κάθε δείγμα, με βάση τον αριθμό των κυττάρων (500 αντίγραφα του κυτταρικού γονιδιώματος), όπως προέκυψε από τη ποσοτικοποίηση του GAPDH γονιδίου και τα αποτελέσματα εκφράστηκαν ως E1, E2 και E6 αντίγραφα/ 500 κύτταρα.

Εκκινητές	Αλληλουχία (5'-3')	Θέση	Προϊόν	Βιβλιογραφία
E1 στόχος				
E1F2	GACAGCACATGCGTTGTTTAC	1059	265bp	Arias Pulido et al., 2006
E1R2	AGTCTCATGGCGCCCTTC	1323		
E2 στόχος				
E2F12	GAATACGAACATATTTTGTGCAG	3243	297 bp	Arias Pulido et al., 2006
E2R12	CTGAGTCTCTGTGCAACAACCTTA	3539		
E6 στόχος				
HPV16 1	GCACAGAGCTGCAAACAACCT	150	353 bp	Παρούσα διατριβή
HPV16 353	GGTCCACCGACCCCTTATATT	502		
GAPDH στόχος				
GAPDH F	ACGGATTTGGTCGATTGGG		250bp	Li et al., 2008
GAPDH R	TGATTTTGGAGGGATCTCGC			

95°C 1 min
 95°C 3 sec
 60°C 20 sec
 60°C 10 sec
 55°C-95°C Αποδιάταξη

} 40 cycles

Πίνακας 2.6. Παρουσίαση των εκκινητών για τη ποσοτικοποίηση των E1, E2, E6 και GAPDH γονιδίων. Στον πίνακα αναγράφεται η αντίστοιχη αλληλουχία των εκκινητών και οι συνθήκες ενίσχυσης της Ποσοτικής PCR Πραγματικού Χρόνου.

2.6.3 Καθορισμός του εύρους των τιμών E1/E6 και E2/E6

Για να προσδιοριστεί η μορφή του γονιδιώματος του ιού για κάθε κλινικό τραχηλικό δείγμα, υπολογίστηκε το εύρος των τιμών E1/E6 και E2/E6 που διαχωρίζουν την επισωματική, ενσωματωμένη και μεικτή μορφή του γονιδιώματος του ιού. Η πειραματική προσέγγιση των ορίων πραγματοποιήθηκε με την δημιουργία διαλυμάτων πλασμιδίων, που περιείχαν μια σταθερή ποσότητα (100 αντίγραφα) του πλασμιδίου pE6-E2 και διαφορετικές ποσότητες του πλασμιδίου pE6, που κυμαίνονταν από 0 έως 10^5 αντίγραφα (Πίνακας 2.7). Το διάλυμα πλασμιδίου που περιείχε μόνο το πλασμίδιο pE6-E2, αξιολογήθηκε ως η επισωματική μορφή του γονιδιώματος του ιού. Αντιθέτως το διάλυμα πλασμιδίου που περιείχε μόνο το πλασμίδιο pE6 αξιολογήθηκε ως η ενσωματωμένη μορφή του πλασμιδίου. Ο αριθμός των αντιγράφων των E1, E2 και E6 γονιδίων υπολογίστηκε σε όλα τα διαλύματα πλασμιδίων και επαναλήφθηκε τρεις φορές για κάθε διάλυμα πλασμιδίου. Ο λόγος των τιμών E1/E6 και E2/E6 υπολογίστηκε για όλα τα διαλύματα πλασμιδίων ξεχωριστά.

Διαλύματα πλασμιδίων		
	pE6-E2	pE6
Διάλυμα 1	100 αντίγραφα	0 αντίγραφα
Διάλυμα 2	100 αντίγραφα	10 αντίγραφα
Διάλυμα 3	100 αντίγραφα	100 αντίγραφα
Διάλυμα 4	100 αντίγραφα	1000 αντίγραφα
Διάλυμα 5	100 αντίγραφα	10000 αντίγραφα
Διάλυμα 6	100 αντίγραφα	100000 αντίγραφα

Πίνακας 2.7. Περιγραφή των διαλυμάτων των πλασμιδίων που χρησιμοποιήθηκαν για τον καθορισμό του εύρους των τιμών που διαχωρίζουν την επισωματική, ενσωματωμένη και μεικτή μορφή του γονιδιώματος του HPV16. Ο λόγος των τιμών E1/E6 και E2/E6 υπολογίστηκε για όλα τα διαλύματα πλασμιδίων.

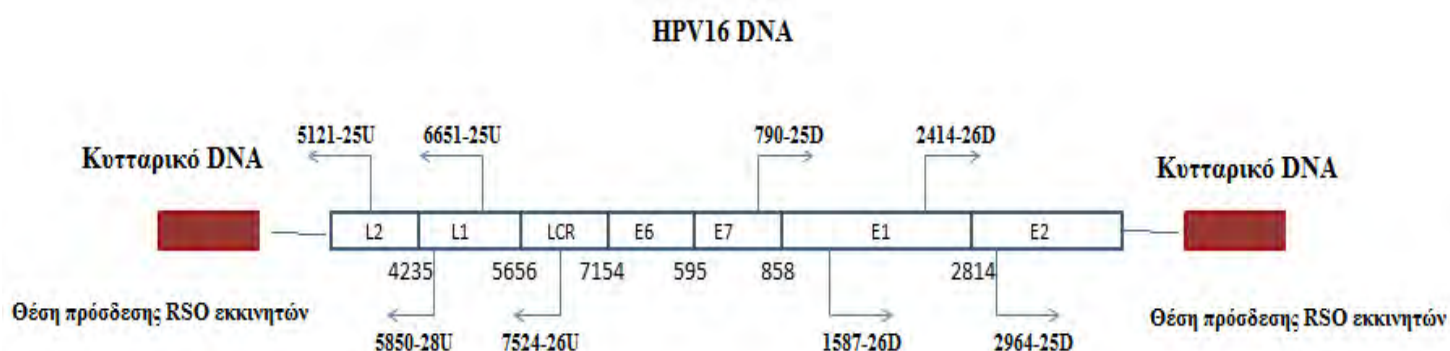
2.6.4 Στατιστική μελέτη των αποτελεσμάτων

Η στατιστική επεξεργασία των δεδομένων πραγματοποιήθηκε με την μέθοδο χ^2 (Chi-square test) με διόρθωση του Yates (Yates' correction) ή με τον ακριβή έλεγχο Fisher (Fisher's exact test) σε 2 x 2 πίνακα συνάφειας (Contingency table) που δημιουργείται με βάση τις παρατηρούμενες τιμές και τη βοήθεια του προγράμματος

GraphPad Prism v. 4.00 (GraphPad Software, San Diego, CA), το οποίο διατίθεται στο διαδίκτυο. Η μέθοδος του χ^2 ανήκει στην κατηγορία των ονομαστικών μεταβλητών και χρησιμοποιείται στη περίπτωση που τα δείγματα είναι ανεξάρτητα μεταξύ τους, ενώ ο ακριβής έλεγχος Fisher χρησιμοποιείται όταν ο αριθμός των δειγμάτων είναι μικρός. Η τιμή P θεωρήθηκε στατιστικά σημαντική σε τιμές μικρότερες του 0.05. Η τιμή P καθορίζει το παρατηρούμενο επίπεδο της στατιστικής σημαντικότητας που υπολογίζεται με τη χρήση του κατάλληλου στατιστικού μοντέλου.

2.7 Προσδιορισμός της θέσης ενσωμάτωσης του HPV16 στο κυτταρικό χρωμόσωμα

Ο προσδιορισμός της θέσης ενσωμάτωσης του ιού στο κυτταρικό DNA πραγματοποιήθηκε με την τεχνική RS-PCR (Restriction Site PCR) (Yu et al., 2005). Η μέθοδος RS-PCR χρησιμοποιείται για την ενίσχυση τμημάτων DNA, όταν μόνο η μία πλευρά της αλληλουχίας-στόχου είναι γνωστή (στη συγκεκριμένη περίπτωση το γονιδίωμα του ιού). Η μέθοδος περιλαμβάνει τη χρήση ειδικών εκκινητών, που αναγνωρίζουν συγκεκριμένες θέσεις στα γονίδια E7, E1, E2, L1 και L2 του HPV16 και συνδυάζονται στην αντίδραση με εκκινητές που αποκαλούνται εκκινητές αναγνώρισης θέσεων περιορισμού (Restriction site oligonucleotide primers, RSO) (Εικόνα 2.2). Οι εκκινητές RSO είναι σχεδιασμένοι να ενισχύουν άγνωστες αλληλουχίες που συνορεύουν στη θέση ενσωμάτωσης του HPV16 γονιδιώματος. Η αλληλουχία των RSO εκκινητών περιλαμβάνει, μία θέση αναγνώρισης του υποκινητή T7, δέκα τυχαία νουκλεοτίδια και έξι νουκλεοτίδια συμπληρωματικά της αντίστοιχης θέσης αναγνώρισης από περιοριστικά ένζυμα (Πίνακας 2.8). Η μέθοδος πραγματοποιείται με μια πρώτη ενίσχυση του ενσωματωμένου ιικού γονιδιώματος με RS-PCR και ακολουθεί Nested RS-PCR. Για τον προσδιορισμό της θέσης ενσωμάτωσης του ιού χρησιμοποιούνται τέσσερις διαφορετικοί RSO εκκινητές, ειδικοί για να προσδεθούν σε θέσεις αναγνώρισης από τα περιοριστικά ένζυμα *Bam*HI, *Eco*RI, *Sau*3AI και *Xba*I. Στη παρούσα εργασία, όλοι οι ειδικοί εκκινητές για τα γονίδια του HPV16 συνδυάστηκαν με τους τέσσερις RSO εκκινητές (Πίνακας 2.8).



Εικόνα 2.2. Σχηματική αναπαράσταση της θέσης πρόσδεσης των εκκινητών της μεθόδου RS-PCR στο γονιδίωμα του HPV16

2.7.1 RS-PCR

Κάθε αντίδραση RS-PCR πραγματοποιήθηκε σε τελικό όγκο 25μl. Σε όλες τις αντιδράσεις χρησιμοποιήθηκαν 3μl DNA από κάθε δείγμα, 1μl από τον ειδικό HPV16 εκκινιτή (συγκέντρωσης 2pmol/μl) 1μl από τον RSO εκκινιτή (συγκέντρωσης 20pmol/μl), 5μl 5X ρυθμιστικού διαλύματος A περιεκτικότητας 1.5mM MgCl₂ (KAPA2G Buffer A, KAPA), 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 0.5U Taq DNA πολυμεράσης (KAPA2G Robust DNA Polymerase, KAPA) και ddH₂O μέχρι τελικό όγκο 25μl

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολουθούν 30 κύκλοι αποδιάταξης στους 95°C για 30sec, υβριδοποίησης στους 45°C για 1min και επιμήκυνσης στους 72°C για 2min. Τέλος η αντίδραση περιλαμβάνει ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων. Στη συνέχεια τα ενισχυμένα προϊόντα υποβλήθηκαν σε Nested RS-PCR. (Πίνακας 2.8)

2.7.2 Nested RS-PCR

Η αντίδραση Nested RS-PCR πραγματοποιήθηκε σε τελικό όγκο 25μl. Σε όλες τις αντιδράσεις χρησιμοποιήθηκαν 2μl από το προϊόν της πρώτης RS-PCR, 1μl από τον ειδικό HPV16 εκκινιτή (συγκέντρωσης 2pmol/μl), ο οποίος υβριδίζεται σε θέση εσωτερικά του πρώτου ειδικού εκκινιτή (RS-PCR), 1μl από τον RSO εκκινιτή (συγκέντρωσης 20pmol/μl), 5μl 5X ρυθμιστικού διαλύματος A περιεκτικότητας 1.5mM MgCl₂ (KAPA2G Buffer A, KAPA), 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 0.5U Taq DNA πολυμεράσης (KAPA2G Robust DNA Polymerase, KAPA) και ddH₂O μέχρι τελικό όγκο 25μl

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολουθούν 35 κύκλοι αποδιάταξης στους 95°C για 30sec, υβριδοποίησης στους 55°C για 1min και επιμήκυνσης στους 72°C για 2min. Τέλος η αντίδραση περιλαμβάνει ένα τελικό στάδιο επιμήκυνσης στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων.

Η επιβεβαίωση των αποτελεσμάτων έγινε σε πήκτωμα αγαρόζης συγκέντρωσης 2%. Στη συνέχεια τα ενισχυμένα προϊόντα υποβλήθηκαν σε μοριακή κλωνοποίηση με το με το πλασμιδιακό φορέα pGEM T-easy vector System (Promega,

Madison, USA). Η απομόνωση του ανασυνδυασμένου πλασμιδιακού φορέα από μετασχηματισμένες βακτηριακές αποικίες πραγματοποιήθηκε με το Nucleospin plasmid kit (Macherey-Nagel GmbH, Duren, Germany), σύμφωνα με τις οδηγίες του κατασκευαστή. Στη συνέχεια πραγματοποιήθηκε αλληλούχιση και των δύο κλώνων του ενθέματος από την εταιρία MacroGen (Amsterdam, Netherlands). Για κάθε αντίδραση αλληλούχισης των κλωνοποιημένων δειγμάτων χρησιμοποιήθηκαν ως εκκινητές οι επαγωγείς T7 και SP6 του πλασμιδιακού φορέα που βρίσκονται εκατέρωθεν της θέσης ένθεσης του τμήματος DNA.

Εκκινητές	Αλληλουχία (5'-3')
RS-PCR	
HPV16-768-24D	ACAAAGCACACACGTAGACATTTCG
HPV16-1545-26D	AGTAATAAATCAACGTGTTGCGATTG
HPV16-2386-25D	TTTGTTTACAACCATTAGCAGATGC
HPV16-2929-24D	GTGCCAACACTGGCTGTATCAAAG
HPV16-5156-27U	TACCAATTCTACTGTACCTAATGCCAG
HPV16-5883-27U	ACTTATTGGGGTCAAGGTAATGTATTC
HPV16-6686-25U	AGTAGATATGGCAGCACATAATGAC
HPV16-7591-22U	GTTGGCAAGCAGTGCAGGTCAG
Nested RS_PCR	
HPV16-790-25D	CGTACTTTGGAAGACCTGTAAATGG
HPV16-1587-26D	GGACTTACACCCAGTATAGCTGCAG
HPV16-2414-26D	AATAGGTATGTTAGATGATGCTACAG
HPV16-2964-25D	ACAAGCAATTGAACTGCAACTAACG
HPV16-5121-25U	GAGGTAAATGCTGGCCTATGTAAAG
HPV16-5850-28U	CCCTGTATTGTAATCCTGATACTTTAGG
HPV16-6651-25U	TGCGTGTAGTATCAACAACAGTAAAC
HPV16-7524-26U	TTAAACCATAGTTGCTGACATAGAAC
RSO εκκινητές	
RSO-Bam	TAATACGACTCACTATAGGGAGANNNNNNNNNNNGGATCC
RSO-Eco	TAATACGACTCACTATAGGGAGANNNNNNNNNNNGAATTC
RSO-Xba	TAATACGACTCACTATAGGGAGANNNNNNNNNNTCTAGA
RSO-Sau	TAATACGACTCACTATAGGGAGANNNNNNNNNNGATC

Πίνακας 2.8. Η κωδικοποίηση των εκκινητών περιλαμβάνει την θέση πρόσδεση στο ιικό γονιδίωμα, ενώ η ένδειξη U και D δηλώνει την κατεύθυνση των εκκινητών. Πιο συγκεκριμένα με U σημειώνονται οι εκκινητές που έχουν προσανατολισμό 5'-3', ενώ με D σημειώνονται οι εκκινητές που έχουν προσανατολισμό 3'-5'. Επιλέον στον πίνακα αναγράφονται οι RSO εκκινητές, οι οποίοι αποτελούνται από μία θέση αναγνώρισης του υποκινητή T7, δέκα τυχαία νουκλεοτίδια (N) και έξι νουκλεοτίδια συμπληρωματικά της αντίστοιχης θέσης αναγνώρισης από τα περιοριστικά ένζυμα *Bam*HI, *Eco*RI, και *Xba*I *Sau*3AI, αντίστοιχα.

2.7.3 Ανάλυση των αλληλουχιών

Η ταυτοποίηση των αλληλουχιών που προέκυψαν με τη μέθοδο RS-PCR για τον προσδιορισμό της θέσης ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα, πραγματοποιήθηκε με την βοήθεια του προγράμματος NCBI BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990) και τον αλγόριθμο Mega-BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

2.7.4 Επιβεβαίωση αναδιατάξεων του HPV16 DNA

Η ανάλυση των αποτελεσμάτων της RS-PCR έδειξε φαινόμενα αναδιάταξης του HPV16 DNA. Οι αναδιατάξεις περιελάμβαναν ρήξη των γονιδίων E1, E2 και συνένωση με το γονίδιο L1 με αντίθετο προσανατολισμό. Για να εξακριβωθεί ότι η αναδιάταξη αυτή δεν έχει προέρθει από λάθος της RS-PCR, πραγματοποιήθηκε PCR με τη χρήση εκκινητών οι οποίοι σχεδιάστηκαν ειδικά για αυτό το σκοπό και στοχεύουν εκατέρωθεν της θέσης αναδιάταξης. Οι εκκινητές σχεδιάστηκαν στα πλαίσια της παρούσας διατριβής με τη βοήθεια του προγράμματος Primer3, Whitehead Institute (<http://bioinfo.ut.ee/primer3-0.4.0>) και με βάση τις αλληλουχίες αναδιάταξης για κάθε περίπτωση ξεχωριστά. Επομένως για την επιβεβαίωση της αναδιάταξης E7,E1-L1 σχεδιάστηκαν δύο ζεύγη εκκινητών για την ενίσχυση διαφορετικών δειγμάτων, που παρουσιάζουν μεταλλάξεις στη νουκλεοτιδική αλληλουχία του L1 γονιδίου: RSE1F1-RSE1R1 και RSE1F1-RSE1R2. Επίσης, για την αναδιάταξη E2-L2 σχεδιάστηκε το ζεύγος των εκκινητών RSE2F1-RSE2R2 (Πίνακας 2.9).

Όλες οι αντιδράσεις της PCR για την ενίσχυση των τμημάτων αναδιάταξης του ιικού DNA (E7, E1-L1 και E2-L1) πραγματοποιήθηκαν σε τελικό όγκο 50μl. Σε όλες τις αντιδράσεις χρησιμοποιήθηκαν 3μl DNA, 2μl εκκινητικών μορίων (1μl από τον καθένα με συγκέντρωση 25pmol/μl), 5μl 10X ρυθμιστικού διαλύματος (GeneON) περιεκτικότητας 1.5 mM MgCl₂, 5μl μείγματος νουκλεοτιδίων (dNTPs 10mM, Invitrogen, USA), 2.5U Taq DNA πολυμεράσης (Maximo Taq DNA Polymerase, GeneON) και ddH₂O μέχρι τελικό όγκο 50μl.

Σε κάθε αντίδραση πραγματοποιήθηκε ένα αρχικό στάδιο αποδιάταξης του DNA στόχου στους 95°C για 2min και ακολούθησε η εφαρμογή διαφορετικών συνθηκών αποδιάταξης, υβριδοποίησης και επιμήκυνσης, για κάθε ζεύγος ξεχωριστά για 40 κύκλους. Όλες οι αντιδράσεις περιλαμβάνουν ένα τελικό στάδιο επιμήκυνσης

στους 72°C για 5min για πλήρη σύνθεση των μη ολοκληρωμένων νεοσυντιθέμενων προϊόντων (Πίνακας 2.9). Η επιβεβαίωση των αποτελεσμάτων έγινε σε πήκτωμα αγαρόζης συγκέντρωσης 2%, ενώ η εξακρίβωση της ειδικότητας των αλληλουχιών πραγματοποιήθηκε μετά από μοριακή κλωνοποίηση και τη χρήση του πλασμιδιακού φορέα pGEM T-easy vector System (Promega, Madison, USA), όπως αναφέρθηκε στην ενότητα 2.4.6.

Εκκινητές	Αλληλουχία (5'-3')	Θέση	Προϊόν	Βιβλιογραφία	
E2-L1					
RSE2F1	ATGCGGGTGGTCAGGTAA	3315	193bp	95°C 2 min 95°C 40 sec 55°C 40 sec 72°C 20 sec 72°C 5 min } 40 cycles Παρούσα διατριβή	
RSE2R1	CGGCTTTGGTGCTATGGACT	6266			
E7, E1-L1					
RSE1F1	TGTGTGCCCCATCTGTCT	828	250bp		95°C 2 min 95°C 40 sec 50°C 40 sec 72°C 20 sec 72°C 5 min } 40 cycles Παρούσα διατριβή
RSE1R1	TTGTCCACCATTAGAGTTACCC	6211			
RSE1F1	TGTGTGCCCCATCTGTCT	828	248bp		
RSE1R2	TTGTCCACCATTAGAGTTAC	6209			

Πίνακας 2.9. Παρουσίαση των εκκινητών για τη επιβεβαίωση των γεγονότων αναδιάταξης του HPV16 DNA. Στον πίνακα αναγράφεται η αντίστοιχη αλληλουχία των εκκινητών που στοχεύουν σε περιοχές εκατέρωθεν της θέσης αναδιάταξης και οι συνθήκες της PCR

Κεφάλαιο 3

3. Αποτελέσματα

Τα αποτελέσματα της παρούσας διατριβής έχουν χωριστεί σε τέσσερις ενότητες. Οι ενότητες περιλαμβάνουν α) την ανάλυση των αποτελεσμάτων από την μοριακή μελέτη των E1, E2, E4, E6 και E7 γονιδίων, β) τη χαρτογράφηση των E1 και E2 γονιδίων για τον προσδιορισμό των θέσεων ρήξης του HPV16 γονιδιώματος, γ) το προσδιορισμό της μορφής του γονιδιώματος του ιού και δ) τον προσδιορισμό θέσεων ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα.

3.1 Μοριακή ανάλυση του γονιδιώματος του HPV16

Στη παρούσα διδακτορική διατριβή πραγματοποιήθηκε μοριακή ανάλυση της πρώιμης περιοχής του γονιδιώματος του HPV16, με στόχο να προσδιοριστούν γονιδιακές τροποποιήσεις, όπως μεταλλάξεις, διπλασιασμοί, ελλείψεις, ενθέσεις και να συσχετιστούν με το βαθμό της ενδοεπιθηλιακής αλλοίωσης. Για το σκοπό αυτό πραγματοποιήθηκε μελέτη των πρώιμων γονιδίων E1, E2, E4, E6 και E7 με τη μέθοδο της PCR, χρησιμοποιώντας εκκινητικά μόρια τα οποία σχεδιάστηκαν στα πλαίσια της παρούσας εργασίας ή επιλέχθηκαν από την υπάρχουσα βιβλιογραφία (Πίνακας 2.4).

3.1.1 Νουκλεοτιδική ανάλυση των E6 και E7 γονιδίων

Για την ανάλυση των E6 και E7 γονιδίων χρησιμοποιήθηκαν συνολικά πενήντα τρία κλινικά δείγματα (σαράντα έξι δείγματα από ThinPrep και επτά τραχηλικά επιχρίσματα). Από τα πενήντα τρία δείγματα, είκοσι τρία δείγματα χαρακτηρίστηκαν από τους κυτταροπαθολόγους ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και είκοσι οκτώ δείγματα χαρακτηρίστηκαν ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Επιπρόσθετα δύο περιπτώσεις χαρακτηρίστηκαν ως καρκίνος του τραχήλου της μήτρας. Για την ανάλυση των E6 και E7 γονιδίων πραγματοποιήθηκε PCR χρησιμοποιώντας τους εκκινητές HPV16 41 - HPV16 757, οι οποίοι ενισχύουν τη περιοχή μεταξύ των γονιδίων E6 και E7 (Πίνακας 2.4).

Η νουκλεοτιδική ανάλυση του Ε6 γονιδίου έδειξε ότι το Ε6 γονίδιο είναι μια αρκετά πολυμορφική περιοχή του γονιδιώματος του ιού. Η πιο συχνά εμφανιζόμενη μετάλλαξη που παρατηρήθηκε στο Ε6 γονίδιο είναι η μετάλλαξη T350G, η οποία προκαλεί την αμινοξική αλλαγή L83V. Η μετάλλαξη T350G έχει συσχετιστεί από προηγούμενες μελέτες με την ανάπτυξη υψηλότερου βαθμού ενδοεπιθηλιακών αλλοιώσεων και καρκίνο και θεωρείται σημαντικός δείκτης πρόβλεψης της εξέλιξης της ενδοεπιθηλιακής αλλοίωσης (Yamada et al., 1997, Andersson et al., 2000, Grodzki et al., 2006). Στη παρούσα μελέτη η μετάλλαξη T350G παρουσίασε ίση κατανομή μεταξύ των δειγμάτων υψηλού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Πιο συγκεκριμένα, η μετάλλαξη T350G ανιχνεύτηκε στο 69.6% (16/23) των δειγμάτων υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, ενώ εντοπίστηκε στο 64.3% (18/28) των περιπτώσεων χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (Πίνακας 3.1, 3.2). Επιπλέον η μετάλλαξη T350G ανιχνεύτηκε σε δύο περιπτώσεις καρκίνου (PNL7 και PNL16).

Επιπρόσθετα εννέα νουκλεοτιδικές μεταλλάξεις A184G, G201A, G219A, A280T, T302A, A336G, G514A, C523T, C539T αναφέρθηκαν για πρώτη φορά στα πλαίσια της παρούσας διατριβής (Πίνακας 3.1). Πιο συγκεκριμένα οι μεταλλάξεις A184G, G201A, G219A, T302A, A336G, προκαλούν τις αμινοξικές αλλαγές I27M, C33Y, R39Q, L67I, και H78R αντίστοιχα και εντοπίζονται στο N- τελικό άκρο της Ε6 ογκοπρωτεΐνης (Ghittoni et al., 2010, Boulet et al., 2007). Επιπλέον οι μεταλλάξεις G514A, C539T προκαλούν τις αμινοξικές αλλαγές M137I και R146C αντίστοιχα. Οι μεταλλάξεις αυτές εντοπίζονται στο C-τελικό άκρο της πρωτεΐνης (Ghittoni et al., 2010, Boulet et al., 2007). Τέλος οι μεταλλάξεις A280T και C523T χαρακτηρίστηκαν ως σιωπηλές μεταλλάξεις (Πίνακας 3.1)

Όσον αφορά τη σύνδεση των νέων μεταλλάξεων στο Ε6 γονίδιο με το βαθμό της ενδοεπιθηλιακής αλλοίωσης παρατηρήθηκε ότι οι μεταλλάξεις A184G, G201A, G219A, A280T, C523T, και C539T ανιχνεύονται σε υψηλού και χαμηλού βαθμού δυσπλασίες, ενώ οι μεταλλάξεις T302A, A336G, G514A ανιχνεύονται μόνο σε χαμηλού βαθμού δυσπλασίες (Πίνακας 3.1). Παράλληλα, παρατηρήθηκε ότι οι καινούργιες μεταλλάξεις που κυκλοφορούν στον Ελληνικό πληθυσμό εντοπίζονται σε δείγματα που φέρουν ταυτόχρονα τη μετάλλαξη T350G και χαρακτηρίζονται ως HPV16 Ευρωπαϊκή ενδοτυπική παραλλαγή. Παρόλα αυτά η μετάλλαξη G219A ανιχνεύτηκε ταυτόχρονα στις ενδοτυπικές παραλλαγές HPV16 African type II και European.

Η νουκλεοτιδική ανάλυση του E7 γονιδίου έδειξε ότι το ανοιχτό πλαίσιο ανάγνωσης του E7 γονιδίου είναι πιο συντηρημένο. Παρόλα αυτά, επτά νέες μεταλλάξεις C565T, T596A, G685A, G813A, G814A, G815A και C818T εντοπίστηκαν και αναφέρθηκαν για πρώτη φορά στα πλαίσια της παρούσας διατριβής. Οι μεταλλάξεις C565T, T596A, G685A, G813A, G814A, G815A και C818T προκαλούν τις αμινοξικές αλλαγές H2Y, M12K, A42T, M84I, G85S, G85D και T86I, αντίστοιχα. Πιο συγκεκριμένα, οι αμινοξικές αλλαγές H2Y και M12K εντοπίζονται στη περιοχή CR1 της E7 ογκοπρωτεΐνης (Munger et al., 2004). Επιπλέον, οι αμινοξικές αλλαγές M84I, G85S, G85D, και T86I εντοπίστηκαν στη C-τελική περιοχή της πρωτεΐνης (Ghittoni et al., 2010) (Πίνακας 3.1).

Όσον αφορά τη σύνδεση των νέων μεταλλάξεων στο E7 γονίδιο με το βαθμό της ενδοεπιθηλιακής αλλοίωσης, παρατηρήθηκε ότι οι μεταλλάξεις T596A, G685A, G814A, G815A και C818T ανιχνεύονται σε δείγματα υψηλού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (Πίνακας 3.1). Αντίθετα, οι μεταλλάξεις C565T και G813A εντοπίζονται μόνο σε δείγματα χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (Πίνακας 3.1). Επιπλέον, οι καινούργιες μεταλλάξεις στο E7 γονίδιο που κυκλοφορούν στον Ελληνικό πληθυσμό εντοπίζονται σε δείγματα που φέρουν ταυτόχρονα τη μετάλλαξη T350G και χαρακτηρίζονται ως HPV16 Ευρωπαϊκή ενδοτυπική παραλλαγή. Η παρουσία των νέων μεταλλάξεων στα γονίδια E6 και E7 που ανιχνεύτηκαν στην Ευρωπαϊκή παραλλαγή του γονιδιώματος του HPV16, χρειάζεται να διερευνηθεί περαιτέρω με σκοπό να προσδιοριστεί ο ρόλος τους στη δομή και τη λειτουργία των E6 και E7 ογκοπρωτεϊνών.

3.1.1.1 Φυλογενετική ανάλυση του HPV16 DNA

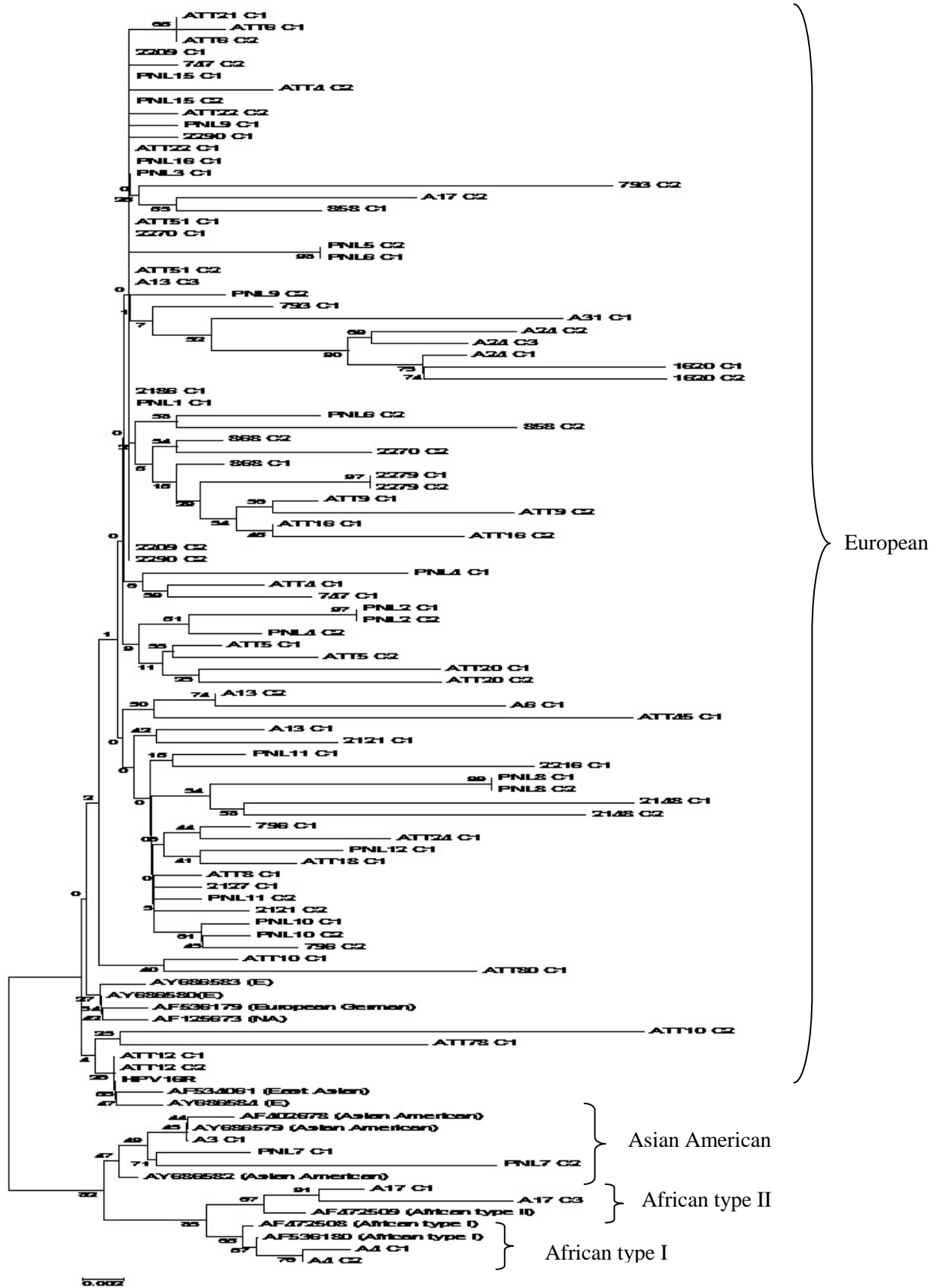
Η φυλογενετική ταξινόμηση του HPV16 γονιδιώματος στις πέντε διαφορετικές ενδοτυπικές γενιές παραλλαγών έχει βασιστεί σε συγκεκριμένους ενδοτυπικούς πολυμορφισμούς εντός των γονιδίων E6 και E7. Πιο συγκεκριμένα, πρότυπο στέλεχος (κανένας πολυμορφισμός), ενδοτυπική παραλλαγή European E-G350 (T350G), E-C109G (T109C, T350G), E-G131G (A131G, T350G), Asian (T178G, A647G), African type I (G132C, C143G, G145T, T286A, A289G, C335T, T789C, T795G), African type II (T109C, G132T, C143G, G145T, T286A, A289G, C335T, G403G, A647G, T789C, T795G) και Asian American (G145T, T286A,

A289G, C335T, T350G, A532G, T732C, T789C, T795G). (Yamada et al., 1995, 1997, Wheeler et al., 1997, Swan et al., 2005, Wu et al., 2006).

Από τη νουκλεοτιδική ανάλυση των E6 και E7 γονιδίων προέκυψε ότι η πιο συχνά εμφανιζόμενη αλληλουχία των γονιδίων E6 και E7 είναι η αλληλουχία της HPV16 Ευρωπαϊκής παραλλαγής (European, T350G) και ακολουθεί η αλληλουχία του πρότυπου στελέχους. Πιο συγκεκριμένα, η Ευρωπαϊκή παραλλαγή εντοπίστηκε στο 65.2% των δειγμάτων υψηλού βαθμού αλλοίωσης και στο 64.2% των δειγμάτων χαμηλού βαθμού αλλοίωσης, ενώ το πρότυπο στέλεχος εντοπίστηκε στο 26.1% των δειγμάτων υψηλού βαθμού αλλοίωσης και στο 35.7% των δειγμάτων χαμηλού βαθμού αλλοίωσης (Πίνακας 3.1, 3.2). Επιπλέον, η ενδοτυπική παραλλαγή HPV16 Asian American εντοπίστηκε στη περίπτωση καρκίνου PNL7 και στο δείγμα χαμηλού βαθμού αλλοίωσης A3. Μεικτή μόλυνση ανιχνεύτηκε στο δείγμα A17, καθώς παρατηρήθηκε ταυτόχρονη μόλυνση από την ενδοτυπική παραλλαγή HPV16 African type II και European. Τέλος το δείγμα υψηλού βαθμού αλλοίωσης A4 χαρακτηρίστηκε ως ενδοτυπική παραλλαγή HPV16 African type I. Με βάση την ανάλυση συγκεκριμένων ενδοτυπικών πολυμορφισμών παρατηρήθηκε ότι μη Ευρωπαϊκές παραλλαγές του γονιδιώματος του ιού κυκλοφορούν στον Ελληνικό πληθυσμό και ανιχνεύονται σε υψηλού βαθμού και χαμηλού βαθμού αλλοιώσεις (Πίνακας 3.1, 3.2).

Η ομαδοποίηση των διαφορετικών στελεχών HPV16 που απομονώθηκαν από τον Ελληνικό πληθυσμό, παρουσιάζεται στο φυλογενετικό δέντρο που κατασκευάστηκε με τη μέθοδο Μέγιστης Πιθανοφάνειας (1000 bootstrap values, Μοντέλο Υποκατάστασης: GTR + G), μετά από πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του E6 γονιδίου του HPV16, τις αλληλουχίες που απομονώθηκαν στα πλαίσια της παρούσας εργασίας και τις αλληλουχίες των ενδοτυπικών παραλλαγών, European: AF125673, AY686583, AY686580, AY686584, AF536179 (European-German type), Asian: AF534061 African type I: AF472508, AF536180, African type II: AF472509 και Asian American: AF402678, AY686579 και AY686582 (Εικόνα 3.1). Όπως φαίνεται στο φυλογενετικό δέντρο, σχηματίζονται τέσσερα διαφορετικά φυλογενετικά κλαδιά που ομαδοποιούν τις πέντε διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16. Σε ένα φυλογενετικό κλαδί ομαδοποιείται το πρότυπο στέλεχος HPV16R μαζί με τις HPV16 ενδοτυπικές παραλλαγές Asian και European, ενώ σε τρία διαφορετικά φυλογενετικά κλαδιά ομαδοποιούνται οι HPV16 ενδοτυπικές παραλλαγές African type I, African

type II και Asian American, αντίστοιχα. (Εικόνα 3.1) (Yamada et al., 1995, 1997, Wheeler et al., 1997, Swan et al., 2005, Wu et al., 2006).



Εικόνα 3.1. Φυλογενετικό δέντρο με τη μέθοδο της Μέγιστης Πιθανοφάνειας (1000 bootstrap values), με βάση την αλληλουχία του Ε6 γονιδίου.

3.1.1.2 Εξελικτική μελέτη των γονιδίων E6 και E7

Η μελέτη της επιλεκτικής πίεσης του E6 γονιδίου έδειξε ότι $dN/dS=1$ και ο μέσος ρυθμός υποκατάστασης του πρώτου δεύτερο και τρίτου νουκλεοτιδίου είναι $CP1=0.826$, $CP2=1.015$, $CP3=1.159$. Για το E7 γονίδιο ο λόγος του E7 γονιδίου $dN/dS=1.25$ και ο μέσος ρυθμός υποκατάστασης $CP1=0.728$, $CP2=1.2$ και $CP3=1.07$. Επομένως ουδέτερη πίεση ασκείται στο γονίδιο E6 και θετική επιλεκτική πίεση επιδρά στο E7 γονίδιο. Επιπλέον, θετική επιλεκτική πίεση εντοπίστηκε στο κωδικόνιο 83 (L83V) στο γονίδιο E6 μέσω Μπεϋζιανής ανάλυσης, ενώ με την μέθοδο Μέγιστης Πιθανοφάνειας δεν εντοπίστηκε θέση με θετική επιλεκτική πίεση. Επιπλέον, στο γονίδιο E7 η μέθοδος της Μέγιστης Πιθανοφάνειας και η Μπεϋζιανή Συμπερασματολογία έδειξαν ότι στο κωδικόνιο 85 (G85S, G85D) ασκείται θετική επιλεκτική πίεση, αλλά μόνο με τη Μπεϋζιανή Συμπερασματολογία η υπόθεση αυτή ήταν στατιστικά σημαντική. Ο προσδιορισμός της επιλεκτικής πίεσης πραγματοποιήθηκε μετά από πολλαπλή ομοπαράθεση των αλληλουχιών της παρούσας μελέτης με τη πρότυπη αλληλουχία του HPV16 γονιδιώματος και τις πρότυπες αλληλουχίες των HPV16 ενδοτυπικών παραλλαγών, European: AF125673, AY686583, AY686580, AY686584, AF536179 (European-German type), Asian: AF534061 African type I: AF472508, AF536180, African type II: AF472509 και Asian American: AF402678, AY686579 και AY686582, όπως περιγράφηκαν στην εργασία του Carvajal-Rodriguez (2008).

3.1.2 Ανίχνευση νέων αλληλουχιών των E6 και E7 γονιδίων

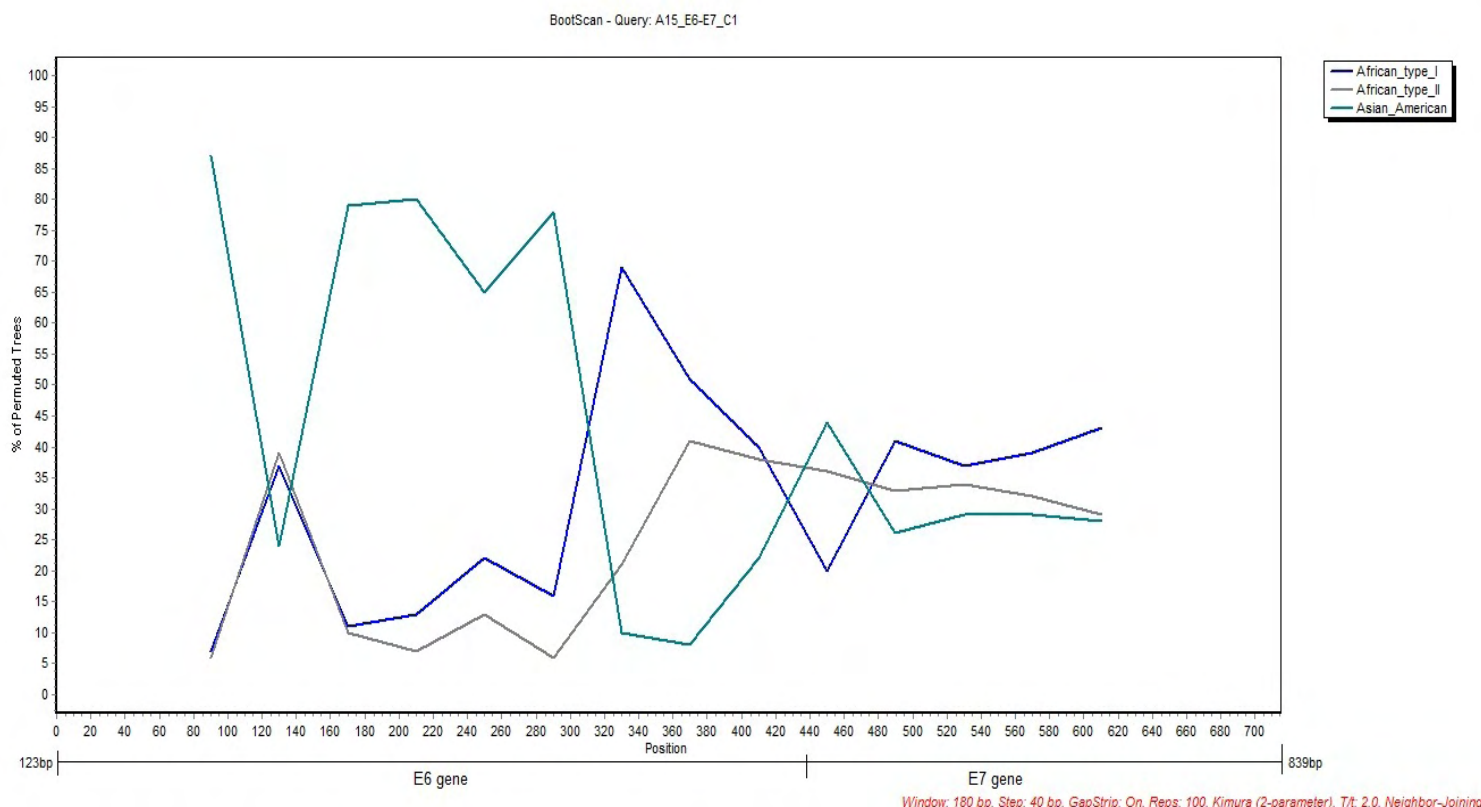
Η νουκλεοτιδική ανάλυση της περιοχής μεταξύ των γονιδίων E6 και E7 αποκάλυψε την ύπαρξη δύο νέων αλληλουχιών E6-E7, οι οποίες ανιχνεύτηκαν στα δείγματα A15 (υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης), A38 και 1143 (χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης). Πιο συγκεκριμένα, η αλληλουχία E6-E7 του δείγματος A15 έδειξε ομοιότητα με δύο διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16 σε διαφορετικές περιοχές των γονιδίων E6 και E7. Μεταξύ των νουκλεοτιδίων 123-350 το δείγμα A15, παρουσιάζει νουκλεοτιδική ομοιότητα με την ενδοτυπική παραλλαγή HPV16 Asian American (ενδοτυπικοί πολυμορφισμοί: G145T, T286A, A289G, C335T, T350G) με μέγιστη ομοιότητα 98% μέσω ClustalW2 ανάλυση. Επιπλέον, μεταξύ των νουκλεοτιδίων 350-839 παρουσιάζει νουκλεοτιδική ομοιότητα με την ενδοτυπική παραλλαγή HPV16 African type I (ενδοτυπικοί

πολυμορφισμοί: T789C, T795G) με μέγιστη νουκλεοτιδική ομοιότητα 99% (Πίνακας 3.3). Το πρόγραμμα SBP έδειξε ένδειξη ανασυνδυασμού στη θέση 405 (κωδικόνιο 101) εντός της αλληλουχίας του Ε6 γονιδίου σύμφωνα με το κριτήριο AIC (AIC=92.3% και cAIC=90%) μέσω πολλαπλής ομοπαράθεσης μεταξύ των κλωνοποιημένων αλληλουχιών του δείγματος A15, τη πρότυπη αλληλουχία του HPV16 γονιδιώματος και τις πρότυπες αλληλουχίες των HPV16 ενδοτυπικών παραλλαγών European, Asian American και African type I. Επιπρόσθετα η ανάλυση Bootscan έδειξε ότι το δείγμα A15 ομαδοποιείται με τις ενδοτυπικές παραλλαγές Asian American και African type I σε διαφορετικές περιοχές της αλληλουχίας των δύο γονιδίων (Εικόνα 3.2)

Η νουκλεοτιδική ανάλυση στα δείγματα χαμηλού βαθμού αλλοίωσης A38 και 1143 μεταξύ των νουκλεοτιδίων 123-143 έδειξε τις νουκλεοτιδικές αλλαγές G132T, C143G, οι οποίες περιέχονται στην ενδοτυπική παραλλαγή HPV16 African type II (Πίνακας 3.3). Αντίθετα μεταξύ των νουκλεοτιδίων 144-839 μόνο η μετάλλαξη T350G ανιχνεύτηκε και οι αλληλουχίες των δειγμάτων έδειξαν νουκλεοτιδική ομοιότητα με την Ευρωπαϊκή παραλλαγή του HPV16 (Πίνακας 3.3). Επίσης, σύμφωνα με το πρόγραμμα SBP παρατηρήθηκε ένδειξη ανασυνδυασμού στη θέση 143 (κωδικόνιο 14) εντός της αλληλουχίας του Ε6 γονιδίου μετά από πολλαπλή ομοπαράθεση των αλληλουχιών των δειγμάτων A38 και 1143, της πρότυπης αλληλουχίας του HPV16 γονιδιώματος και των πρότυπων αλληλουχιών των ενδοτυπικών παραλλαγών HPV16 Asian American, African type I, African type II και European (AIC=100% και cAIC=100%).

Κλώνοι	Ε6 γονίδιο										Ε7 γονίδιο			Αριθμοί καταχώρησης στη Genbank	
	1	1	1	1	2	2	3	3	4	5	6	7	7		
	3	3	4	4	8	8	3	5	0	3	4	3	8	9	
	1	2	3	5	6	9	5	0	3	2	7	2	9	5	
	A	G	C	G	T	A	C	T	A	A	A	T	T	T	
A15 E6-E7 C1	-	-	-	T	A	G	T	G	-	-	-	-	C	G	JN205116
A15 E6-E7 C2	-	-	-	T	A	G	T	G	-	-	-	-	C	G	JN205117
A38 E6-E7 C1	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205111
A38 E6-E7 C2	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205112
A38 E6-E7 C3	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205113
1143 E6-E7 C1	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205114
1143 E6-E7 C2	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205115
European	G	-	-	-	-	-	-	G	-	-	-	-	-	-	
African type I	-	C	G	T	A	G	T	-	-	-	-	-	C	G	
African type II	-	T	G	T	A	G	T	-	G	-	G	-	C	G	
Asian American	-	-	-	T	A	G	T	G	-	G	-	C	C	G	
Asian	-	-	-	-	-	-	-	-	-	-	G	-	-	-	

Πίνακας 3.3. Νουκλεοτιδική ανάλυση των δειγμάτων A15, A38 και 1143. Τα δείγματα A38 και 1143 φέρουν τους πολυμορφισμούς G132T και C143G, που ανιχνεύονται στην ενδοτυπική παραλλαγή HPV16 African type II, και τον πολυμορφισμό T350G που ανιχνεύεται στην Ευρωπαϊκή ενδοτυπική παραλλαγή. Επιπρόσθετα το δείγμα A15 μεταξύ των νουκλεοτιδίων 123 and 350 παρουσιάζει ομοιότητα με την ενδοτυπική παραλλαγή HPV16 Asian American, ενώ μεταξύ των νουκλεοτιδίων 350 και 839 παρουσιάζει ομοιότητα με την ενδοτυπική παραλλαγή HPV16 African type I

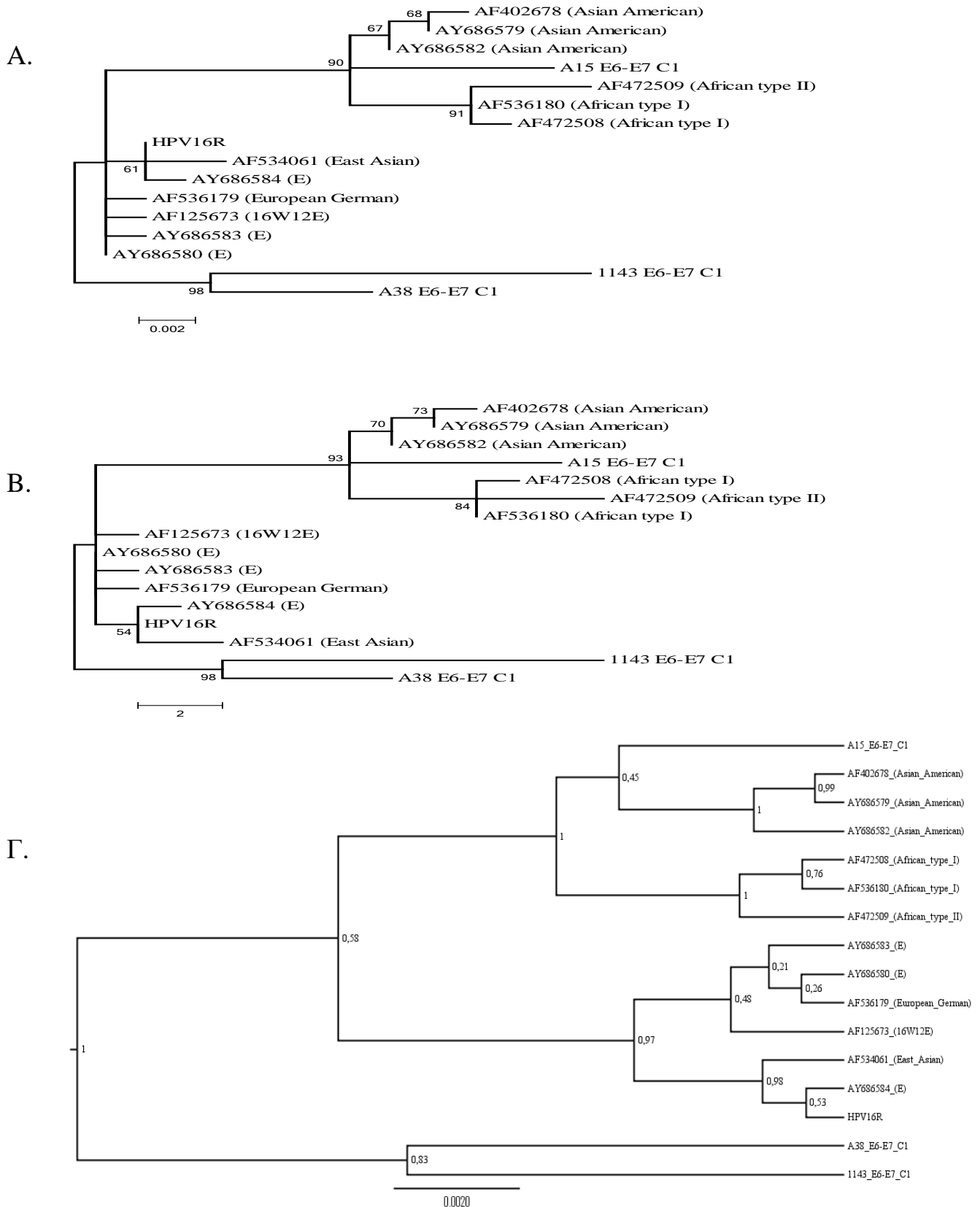


Εικόνα 3.2. Bootscan ανάλυση του δείγματος A15 με τις ενδοτυπικές παραλλαγές African type I, African type II και Asian American. Ο άξονας των Ψ δείχνει το ποσοστό των τιμών bootstrap και ο άξονας των X την θέση των γονιδίων E6 και E7. Η ανάλυση δείχνει την ομαδοποίηση του δείγματος A15 με τις ενδοτυπικές παραλλαγές Asian American και African type I σε διαφορετικές θέσεις της περιοχής E6 και E7.

3.1.2.1 Μελέτη της επιλεκτικής πίεσης των ενδοτυπικών πολυμορφισμών των E6 και E7 γονιδίων

Στο γονίδιο E6 θετική επιλεκτική πίεση ($dN/dS > 1$) ανιχνεύτηκε στα κωδινόνια 10 (G132T), 14 (C143G) και 83 (T350G), μόνο μέσω Μπεϋζιανής ανάλυσης μετά από πολλαπλή ομοπαράθεση των νέων αλληλουχιών με τη πρότυπη αλληλουχία του HPV16 γονιδιώματος και τις αλληλουχίες των HPV16 ενδοτυπικών παραλλαγών (DeFilippis et al., 2002, Chen et al., 2005, Lee et al., 2008, Carvajal-Rodriguez, 2008). Επιπλέον δεν εντοπίστηκαν θέσεις που ασκείται θετική επιλεκτική πίεση εντός του γονιδίου E7. Με βάση την αμινοξική μελέτη των ογκογονιδίων E6 και E7 που πραγματοποιήθηκε στη παρούσα διατριβή και από τη βιβλιογραφία έχει προταθεί ότι οι θέσεις 10 και 14 (θετική επιλεκτική πίεση) συμβάλουν στην ταξινόμηση των διαφορετικών HPV16 παραλλαγών. Η θέση 10I διαχωρίζει τις ενδοτυπικές παραλλαγές HPV16 African type I και African type II, ενώ η θέση 14E διαχωρίζει τις Αφρικανικές παραλλαγές (African type I, African type II) από τις μη Αφρικανικές παραλλαγές του γονιδιώματος του ιού (Carvajal-Rodriguez, 2008). Συμπερασματικά, η νέα αλληλουχία που ανιχνεύτηκε στα δείγματα A38 και 1143 έχει νουκλεοτιδική ομοιότητα με την Ευρωπαϊκή ενδοτυπική παραλλαγή του ιού και εντός της αλληλουχίας της Ευρωπαϊκής παραλλαγής εντοπίζονται οι πολυμορφισμοί στα κωδικόνια 10 και 14, που ανιχνεύονται μόνο στις Αφρικανικές ενδοτυπικές παραλλαγές του HPV16 γονιδιώματος.

Για να μελετηθεί η ταξινόμηση και η σχέση των νέων αλληλουχιών με τα διαφορετικά φυλογενετικά κλαδιά του HPV16 γονιδιώματος, κατασκευάστηκαν τρία διαφορετικά φυλογενετικά δέντρα με τις μεθόδους Μέγιστης Πιθανοφάνειας, Μέγιστης Φειδωλότητας και Μπεϋζιανής Συμπερασματολογίας. Από τα φυλογενετικά δέντρα παρατηρήθηκε ότι οι αλληλουχίες E6-E7 των δειγμάτων A38 και 1143 (10I, 14E, 83V) παρουσιάζουν διαφορετική φυλογενετική ταξινόμηση και αποκλίνουν από τα τέσσερα διαφορετικά φυλογενετικά κλαδιά που οργανώνουν το γονιδίωμα του ιού στις πέντε διαφορετικές γενιές ενδοτυπικών παραλλαγών (Εικόνα 3.3). Επίσης το δείγμα A15 παρουσιάζει φυλογενετική απόκλιση από το κλαδί στο οποίο οργανώνεται η HPV16 ενδοτυπική παραλλαγή Asian American (Εικόνα 3.3).



Εικόνα 3.3. Α) Φυλογενετικό δέντρο με τη μέθοδο Μέγιστης Πιθανοφάνειας (1000 bootstrap values, Μοντέλο Υποκατάστασης: GTR + G). Β) Φυλογενετικό δέντρο με τη μέθοδο Μέγιστης Φειδωλότητας (1000 bootstrap values). Γ) Φυλογενετικό δέντρο με τη Μπεϋζιανή Συμπερασματολογία (Μοντέλο Υποκατάστασης: GTR + G).

3.1.3 Νουκλεοτιδική ανάλυση του E2 γονιδίου

Για τη νουκλεοτιδική ανάλυση του E2 γονιδίου χρησιμοποιήθηκαν συνολικά είκοσι δύο δείγματα (οχτώ τραχηλικοί ιστοί μονιμοποιημένοι σε παραφίνη, εννέα δείγματα από τραχηλικά επιχρίσματα και πέντε δείγματα ThinPrep). Από τα είκοσι δύο δείγματα έντεκα δείγματα χαρακτηρίστηκαν ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και έντεκα δείγματα χαρακτηρίστηκαν ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Το E2 γονίδιο έχει μέγεθος 1098bp και η ενίσχυση του πραγματοποιήθηκε με ένα ζευγάρι εκκινητών που ενισχύει ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του γονιδίου και επιπλέον δύο εσωτερικούς εκκινητές που ενισχύουν σε επιμέρους τμήματα ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του γονιδίου (Πίνακας 2.4).

Όλα τα δείγματα χαμηλού βαθμού αλλοίωσης έφεραν ακέραιο το E2 γονίδιο. Αντιθέτως η περιοχή μεταξύ των νουκλεοτιδίων 3312 έως 3912 δεν ενισχύθηκε στο κλινικό δείγμα D30, ενώ η περιοχή μεταξύ των νουκλεοτιδίων 2732 έως 3364 δεν ενισχύθηκε στα δείγματα D2, D4 και D14. Παράλληλα, η ενίσχυση τμήματος του γονιδίου της β-ακτίνης μεγέθους 498 bp (Li et al., 2008) σε όλα τα δείγματα ιστού μονιμοποιημένου σε παραφίνη επιβεβαίωσε, ότι η ρήξη του E2 γονιδίου δεν προέρχεται από κατακερματισμό του γονιδιώματος του ιού λόγω επεξεργασίας του ιστού, αλλά από την ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα.

Από τη νουκλεοτιδική ανάλυση του E2 γονιδίου εντοπίστηκαν μεταλλάξεις, οι οποίες έχουν αναφερθεί και έχουν συνδεθεί με τις διαφορετικές παραλλαγές του γονιδιώματος του HPV16 από προηγούμενες μελέτες (Azizi et al., 2008, Giannoudis et al., 2001, Jiang et al., 2009, Graham et al., 2000, Eriksson et al., 1999). Συνολικά εντοπίστηκαν εννέα μεταλλάξεις στην περιοχή που κωδικοποιεί για τη περιοχή διενεργοποίησης (transactivation domain), εκ των οποίων δύο μεταλλάξεις στις θέσεις 2791 και 2823 αναφέρθηκαν για πρώτη φορά στη παρούσα διατριβή. Η μετάλλαξη 2791 ανιχνεύτηκε στα δείγματα χαμηλού βαθμού αλλοίωσης D7, D34 και χαρακτηρίστηκε ως σιωπηλή μετάλλαξη ενώ η μετάλλαξη 2823 προκαλεί την αμινοξική αλλαγή S23I και ανιχνεύτηκε στα δείγματα υψηλού βαθμού αλλοίωσης A13 και D30 (Πίνακας 3.5). Η αμινοξική αλλαγή S23I τοποθετείται στη περιοχή αλληλεπίδρασης της E2 πρωτεΐνης με την E1 πρωτεΐνη (θέση 18-41). Η περιοχή αυτή εμπλέκεται στη δημιουργία E1-E2 συμπλόκου το οποίο είναι υπεύθυνο για την έναρξη της αντιγραφής του ιικού DNA (Piccini et al., 1995).

Επιπλέον δέκα μεταλλάξεις εντοπίστηκαν στη περιοχή του E2 γονιδίου που κωδικοποιεί για τη περιοχή "αρμός" (hinge region). Μεταξύ των μεταλλάξεων αυτών αναφέρθηκε μια καινούργια μετάλλαξη στη θέση 3361 (C3361T). Η μετάλλαξη C3361T δεν προκαλεί αμινοξική αλλαγή στην E2 πρωτεΐνη. Η συγκεκριμένη μετάλλαξη ανιχνεύτηκε στα δείγματα A17, A3 και A4, τα οποία έχουν ομαδοποιηθεί με τις ενδοτυπικές παραλλαγές African type II, Asian American και African type I αντίστοιχα (Πίνακας 3.4, 3.5), μέσω ανάλυσης ενδοτυπικών νουκλεοτιδικών πολυμορφισμών εντός των ογκογονιδίων E6 και E7 (ενότητα 3.1.1.1). Η μετάλλαξη αυτή δεν παρατηρήθηκε στις πρότυπες αλληλουχίες του γονιδίου E2 και των ενδοτυπικών παραλλαγών του γονιδιώματος του HPV16. Επιπρόσθετα στη περιοχή "αρμός" ανιχνεύτηκε η αμινοξική μετάλλαξη P219S (C3410T) (Πίνακας 3.4, 3.5). Η μετάλλαξη C3410T έχει συνδεθεί με υψηλότερου βαθμού δυσπλασίες και καρκίνο του τραχήλου της μήτρας (Graham et al., 2000). Στη παρούσα εργασία η αμινοξική μετάλλαξη P219S παρουσίασε ίση κατανομή μεταξύ των υψηλού και χαμηλού βαθμού ενδοεπιθηλιακών αλλοιώσεων. Τέλος επτά μεταλλάξεις (A3622T, C3684A, T3694A, T3706G, G3778T, C3787A, T3805G) ανιχνεύτηκαν στη περιοχή του E2 γονιδίου που κωδικοποιεί για την περιοχή πρόσδεσης της E2 πρωτεΐνης στο ικό DNA (DNA binding domain) (Πίνακας 3.4, 3.5).

Βαθμός αλλοίωσης	Κλώνοι	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	Ενδοτυπικές παραλλαγές	Αριθμοί καταχώρησης στη GenBank		
		7	8	8	9	1	1	2	2	3	3	3	4	4	4	5	5	5	6	6	6	7	7	7	8	8			8	0
		1	3	0	7	9	1	2	4	9	1	2	7	0	3	9	6	7	8	6	2	4	4	6	8	7			5	
		G	G	C	A	C	C	G	T	G	C	A	C	C	G	G	C	T	A	T	A	C	T	T	G	C	T			
CIN III	A3 632 C1	-	-	A	T	A	T	A	A	A	Hinge domain													Asian American	HM 162479					
CIN III	A3 632 C2	-	-	A	T	T	T	A	A	A	DNA binding domain													Asian American	HM 162480					
CIN III	A4 632 C1	-	-	A	T	A	T	A	-	A														African type I	HM 162481					
CIN III	A6 632 C1	-	-	-	-	-	-	-	-	-														European	HM 162482					
CIN III	A6 632 C2	-	-	-	-	-	-	-	-	-														European	HM 162483					
CIN III	A13 632 C1	-	T	-	-	-	-	-	-	-														European	HM 162484					
CIN III	A13 632 C2	-	T	-	-	-	-	-	-	-														European	HM 162485					
CIN III	A15 632 C1	-	-	A	-	G	T	A	A	A														Asian American	HM 162486					
CIN II	A31 632 C1	-	-	-	-	-	-	-	-	-														European	HM 162487					
CIN III	D16 632 C1	-	-	-	-	-	-	-	-	-														European	HQ 398218					
CIN I-II	D30 632 C1	-	A	-	-	-	-	-	-	-														European	HQ 398216					
CIN I-II	D30 632 C2	-	-	-	-	-	-	-	-	-														European	HQ 398217					
CIN III	A3 598 C1										-	G	G	T	A	A	A	-	C	-	-	-	-	-	-	-	C	Asian American	HM 162489	
CIN III	A3 598 C2										T	G	G	T	A	A	A	C	C	G	-	A	A	G	T	A	G	Asian American	HM 162490	
CIN III	A4 598 C1										-	G	G	T	-	A	A	-	-	G	-	A	A	G	T	A	G	African type I	HM 162491	
CIN III	A4 598 C2										T	G	G	T	-	A	A	-	-	G	-	A	A	-	T	A	-	African type I	HM 162492	
CIN III	A6 598 C1										-	-	-	T	-	-	A	C	-	G	-	A	A	-	T	A	-	European	HM 162493	
CIN III	A13 598 C1										-	-	-	T	-	A	-	-	-	-	-	-	-	-	-	-	-	European	HM 162494	
CIN III	A13 598 C2										-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM 162495	
CIN III	A15 598 C1										-	G	G	-	A	A	A	C	C	-	-	-	-	-	-	-	-	Asian American	HM 162496	
CIN III	A31 598 C1										-	-	-	-	-	A	-	-	-	G	-	A	A	G	T	A	G	European	HM 162497	
CIN II	D2 598 C1										-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ 398221	
CIN II-III	D4 598 C1										-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ 398219	
CIN III	D14 598 C1										-	-	-	T	-	-	-	-	C	-	-	-	-	G	-	-	-	European	HQ 398220	
CIN III	D16 598 C1										-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ 398222	
	Amino acid	I	S	H	E	T	H	A	G	R	S	N	P	P	A	E	T	T	S	F	V	T	T	S	W	D	V			
		0	0	0	0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3		
		1	2	3	7	3	3	4	5	6	0	0	0	1	2	3	5	5	6	7	8	1	1	1	4	4	5			
		5	3	5	4	5	6	3	6	5	2	3	8	9	0	2	4	4	1	1	9	0	3	7	1	4	0			
		-	I	Q	D	K	Y	T	-	Q	-	D	A	S	T	K	N	N	-	V	-	K	-	-	C	E	-			

Πίνακας 3.4. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων του E2 γονιδίου. Οι αριθμοί 1176, 632 και 598 αντιστοιχούν στο μέγεθος των προϊόντων και το τελευταίο γράμμα (C1 ή C2) αντιστοιχεί στον αριθμό του κλώνου. Όλα τα δείγματα χαμηλού βαθμού αλλοίωσης περιέχουν ακέραιο το γονίδιο E2.

Βαθμός αλλοίωσης	Κλώνοι	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	Ενδοτυπικές Παραλλαγές	Αριθμοί καταχώρησης στη GenBank			
		7	8	8	9	1	1	1	2	2	3	3	3	4	4	4	5	5	5	5	6	6	6	7	7			7	8	
		9	2	6	7	5	6	8	2	4	6	6	7	1	1	4	1	1	3	6	2	8	9	0	7	8	0			
		1	3	0	7	9	1	2	4	9	1	2	7	0	3	9	6	7	8	6	2	4	4	6	8	7	5			
		G	G	C	A	C	C	G	T	G	C	A	C	C	G	G	C	T	A	T	A	C	T	T	G	C	T			
CIN I	D7 632 C1	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422083		
CIN I	D7 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422085		
CIN I	D7 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422084		
CIN I	D7 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422086		
CIN I	D34 632 C1	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422087		
CIN I	D34 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422089		
CIN I	D34 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422088		
CIN I	D34 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422090		
CIN I	D42 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422091		
CIN I	D42 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422093		
CIN I	D42 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422094		
CIN I	747 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	A	-	-	-	-	T	-	-	-	-	-	European	HM 162464		
CIN I	747 1176 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-	-	-	-	European	HM 162465		
CIN I	747 1176 C3	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-	-	-	-	European	HM 162466		
CIN I	747 1176 C4	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-	-	-	-	European	HM 162467		
CIN I	747 1176 C5	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-	-	-	-	European	HM 162468		
CIN I	796 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	A	-	-	-	-	T	-	-	-	-	-	European	HM 162469		
CIN I	858 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	HM 162475		
CIN I	868 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	HM 162477		
CIN I	793 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	European	HM 162478		
CIN I	A17 1176 C1	-	-	A	-	A	T	A	-	A	-	G	G	T	-	A	A	C	-	G	-	A	A	C	T	A	G	African type II	HQ 398228	
CIN I	A17 1176 C2	-	-	A	-	A	T	A	-	A	T	G	G	T	-	A	A	C	C	G	-	A	A	C	T	A	G	African type II	HQ 398229	
CIN I	A24 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	T	-	European	HM 162463	
CIN I	A38 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM 162462	
	Amino acid	I	S	H	E	T	H	A	G	R	S	N	P	P	A	E	T	T	S	F	V	T	T	S	W	D	V			
		0	0	0	0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3		
		1	2	3	7	3	3	4	5	6	0	0	0	1	2	3	5	5	6	7	8	1	1	1	1	4	4	5		
		5	3	5	4	5	6	3	6	5	2	3	8	9	0	2	4	4	1	1	9	0	3	7	1	4	0			
		L	I	O	D	K	Y	T	-	Q	-	D	A	S	T	K	N	N	-	V	-	K	-	-	C	E	-			
		Transactivation domain										Hinge domain										DNA binding domain								

Πίνακας 3.5. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων του E2 γονιδίου. Οι αριθμοί 632 και 598 αντιστοιχούν στο μέγεθος των προϊόντων και το τελευταίο γράμμα (C1 or C2) αντιστοιχεί στον αριθμό του κλώνου. Η περιοχή μεταξύ των νουκλεοτιδίων 3312 έως 3912 δεν ενισχύθηκε στο κλινικό δείγμα D30, ενώ η περιοχή μεταξύ των νουκλεοτιδίων 2732 έως 3364 δεν ενισχύθηκε στα δείγματα D2, D4 και D14.

3.1.3.1 Εξελικτική μελέτη του E2 γονιδίου

Η μελέτη της επιλεκτικής πίεσης που ασκείται στο E2 γονίδιο πραγματοποιήθηκε σε δύο περιοχές του E2 γονιδίου ξεχωριστά. Πιο συγκεκριμένα, η περιοχή μεταξύ των νουκλεοτιδίων 2755 έως 3364 έδειξε ότι $dN/dS=0.9$ και ο μέσος ρυθμός υποκατάστασης είναι $CP1=1$, $CP2=0.956$ $CP3=1.035$, ενώ μεταξύ των νουκλεοτιδίων 3312 έως 3852 παρατηρήθηκε ότι ο λόγος $dN/dS= 0.88$ και ο μέσος ρυθμός υποκατάστασης είναι $CP1=0.882$, $CP2=1.086$, $CP3=1.05$ μετά από πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του E2 γονιδίου με τις αλληλουχίες της παρούσας εργασίας και τις αλληλουχίες του E2 γονιδίου που είναι διαθέσιμες στη τράπεζα δεδομένων GenBank με αριθμό καταχώρησης: AY686580 AY686583, AF536179, AF534061, AF472509, AF402678, AF472508, AF536180, FJ610150, FJ610151, FJ610147, FJ610149, FJ610146, AY686581, AF125673, FJ610152, FJ610148, FJ006723, AY686584, AF193429, DQ191463, AF067034, AF193425, AF193427, AF115832, DQ191465, AY686579, NC001526, AY686582, AF067035, AF115831, EU430675, EU430667, EU430683, EU118173, AF515620, AF067029, DQ191462, AF193430, AF193426, AF067032, AF067028, AF067030, K02718, DQ191461, AF193428, AF067031, EU918764, και HM057182. Τέλος, δεν ανιχνεύτηκε θέση με θετική επιλεκτική πίεση με τις μεθόδους Μέγιστης Πιθανοφάνειας και η Μπεϋζιανής Συμπερασματολογίας.

3.1.4 Νουκλεοτιδική ανάλυση του E4 γονιδίου

Για την ανάλυση του E4 γονιδίου χρησιμοποιήθηκαν συνολικά είκοσι επτά κλινικά δείγματα (έντεκα δείγματα ThinPrep, εννέα τραχηλικά επιχρίσματα και επτά τομές τραχηλικού ιστού). Από τα είκοσι επτά δείγματα, δεκαπέντε περιπτώσεις χαρακτηρίστηκαν ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και δώδεκα δείγματα χαρακτηρίστηκαν ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Το E4 γονίδιο έχει μέγεθος 288bp και η ενίσχυσή του πραγματοποιήθηκε με τη μέθοδο της PCR με τη χρήση του ζεύγους των εκκινητικών μορίων HPV16 3312- HPV16 3912 (Πίνακας 2.4).

Μέχρι στιγμής νουκλεοτιδική ανάλυση του E4 γονιδίου δεν έχει πραγματοποιηθεί σε προηγούμενες μελέτες. Από τη νουκλεοτιδική μελέτη του E4 γονιδίου παρατηρήθηκαν μεταλλάξεις οι οποίες δεν έχουν αναφερθεί στη βιβλιογραφία. Πιο συγκεκριμένα, στη περιοχή του E4 γονιδίου που κωδικοποιεί για το N-τελικό άκρο της E1^{E4} πρωτεΐνης εντοπίστηκαν οι μεταλλάξεις G3360C, C3361T, A3362G, C3377G, T3387C, C3410T, G3413A και G3449A (Πίνακας 3.6, Πίνακας 3.7) Οι μεταλλάξεις G3360C και C3361T εντοπίζονται σε χαμηλού και υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις και προκαλούν τις αμινοξικές αλλαγές A7P και A7V, αντίστοιχα (Πίνακας 3.6, Πίνακας 3.7). Επιπλέον οι μεταλλάξεις A3362G, C3377G, T3387C, C3410T, και G3449A χαρακτηρίστηκαν ως σιωπηλές μεταλλάξεις και εντοπίζονται σε χαμηλού και υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις, ενώ η σιωπηλή μετάλλαξη G3413A εντοπίστηκε μόνο σε δείγματα υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Επιπρόσθετα, δύο διαφορετικές μεταλλάξεις παρατηρήθηκαν στη θέση 3387. Πιο συγκεκριμένα, η μετάταξη T3387C (A3, A15, PNL6, PNL7) χαρακτηρίστηκε ως σιωπηλή μετάταξη ενώ η T3387A (747, A17) προκαλεί την αμινοξική αλλαγή L16I. Όσον αφορά τη σύνδεση των μεταλλάξεων αυτών με το βαθμό της ενδοεπιθηλιακής αλλοίωσης παρατηρήθηκε ότι η μετάλλαξη T3387C εντοπίζεται μόνο σε υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις, ενώ η μετάλλαξη T3387A εντοπίζεται σε χαμηλού βαθμού ενδοεπιθηλιακές δυσπλασίες. (Πίνακας 3.6, Πίνακας 3.7)

Στη περιοχή του E4 γονιδίου που κωδικοποιεί για τη κεντρική περιοχή της E1^{E4} πρωτεΐνης εντοπίστηκαν μεταλλάξεις στις θέσεις 3476 και 3516. Δύο μεταλλάξεις ανιχνεύτηκαν στη θέση 3476. Η μετάλλαξη C3476A προκαλεί την αμινοξική αλλαγή D45E, ενώ η μετάλλαξη C3476T χαρακτηρίστηκε ως σιωπηλή μετάλλαξη. Επιπλέον, η μετάλλαξη C3516A προκαλεί την αμινοξική αλλαγή L59I

στα δείγματα A4 και A6, A31, ενώ στα δείγματα A3, A15, A17, PNL6 και PNL7 παρατηρείται σε συνδυασμό με τη μετάλλαξη T3517C προκαλώντας την αμινοξική αλλαγή L59T (Πίνακας 3.6, Πίνακας 3.7).

Τέλος, στη περιοχή του E4 γονιδίου που κωδικοποιεί για την C-τελική περιοχή της E1^{E4} πρωτεΐνης, εντοπίστηκαν οι μεταλλάξεις A3538C, C3556T και T3566G, οι οποίες προκαλούν τις αμινοξικές αλλαγές Q66P, S72F, H75Q, αντίστοιχα. Η περιοχή αυτή της πρωτεΐνης είναι υπεύθυνη για τη καταστροφή της κεράτινης στοιβάδας και τον ολιγομερισμό της E1^{E4} πρωτεΐνης. Οι μεταλλάξεις A3538C και T3566G ανιχνεύτηκαν σε υψηλού βαθμού και χαμηλού βαθμού αλλοιώσεις, ενώ η μετάλλαξη C3556T ανιχνεύτηκε μόνο σε υψηλού βαθμού δυσπλασίες (Πίνακας 3.6, Πίνακας 3.7).

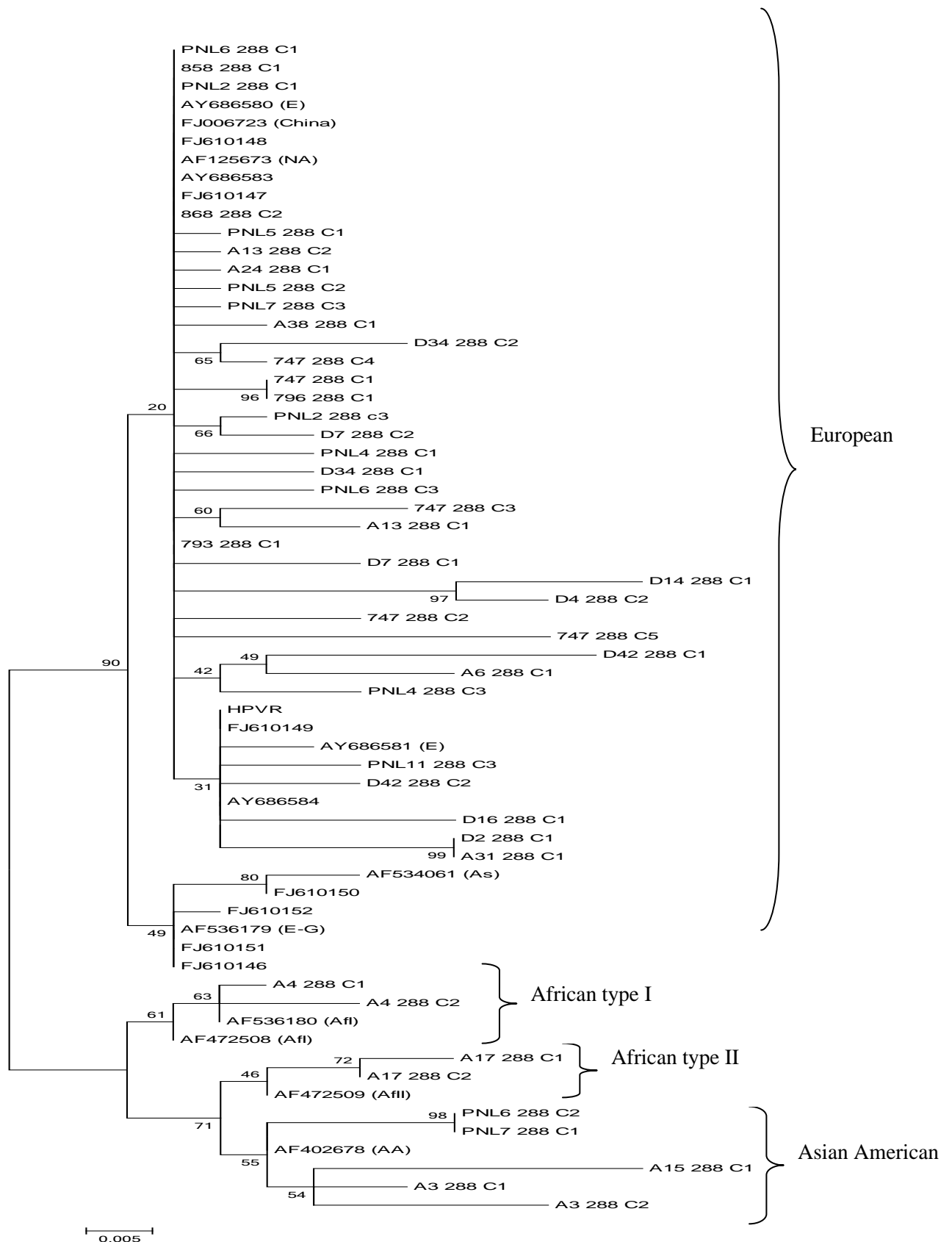
3.1.4.1 Φυλογενετική ανάλυση του HPV16 DNA με βάση την αλληλουχία του E4 γονιδίου

Για να συνδεθούν οι μεταλλάξεις που αναφέρθηκαν στη παρούσα εργασία με τις ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16 πραγματοποιήθηκε συγκριτική μελέτη της αλληλουχίας του E4 γονιδίου με την αλληλουχία των E6 και E7 γονιδίων στα ίδια κλινικά δείγματα. Μετά από συγκριτική μελέτη μεταξύ της αλληλουχίας των E4, E6 και E7 γονιδίων, παρατηρήθηκε ότι οι μεταλλάξεις A3362G, C3377G, G3449A, C3516A, T3566G, συνδέονται με τις ενδοτυπικές παραλλαγές HPV16 African type I, African type II και Asian American (ενδοτυπικοί πολυμορφισμοί E6, E7: G145T, T286A, A289G, C335T, T789C, T795G εντός των γονιδίων E6 και E7). Επιπρόσθετα, οι μεταλλάξεις T3517C και A3538C συνδέθηκαν με τις ενδοτυπικές παραλλαγές HPV16 African type II και Asian American, ενώ η μετάλλαξη C3556T εντοπίζεται σε δείγματα που σχετίζονται με τις ενδοτυπικές παραλλαγές HPV16 Asian American και African type I. Η μετάλλαξη C3556T δεν εντοπίζεται στις πρότυπες αλληλουχίες των αντίστοιχων ενδοτυπικών παραλλαγών. Επιπρόσθετα οι μεταλλάξεις T3387C και G3413A συνδέθηκαν με την ενδοτυπική παραλλαγή HPV16 Asian American (ενδοτυπικοί πολυμορφισμοί E6, E7: A532G και T732C).

Για να προσδιοριστεί εάν οι μεταλλάξεις που εντοπίζονται στο E4 γονίδιο μπορούν να ταξινομήσουν το γονιδίωμα του ιού στα τέσσερα διαφορετικά φυλογενετικά κλαδιά, που οργανώνονται οι πέντε διαφορετικές ενδοτυπικές

παραλλαγές του γονιδιώματος του ιού, κατασκευάστηκε ένα φυλογενετικό δέντρο με τη μέθοδο της Μέγιστης Πιθανοφάνειας (1000 bootstrap values, Μοντέλο Υποκατάστασης: GTR), μετά από πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του E4 γονιδίου, τις αλληλουχίες που αναλύθηκαν στη παρούσα εργασία και τις πρότυπες αλληλουχίες των HPV16 ενδοτυπικών παραλλαγών. Από τη κατασκευή του φυλογενετικού δέντρου παρατηρήθηκε ότι, το E4 γονίδιο ταξινομεί το γονιδίωμα του HPV16 σε τέσσερα διαφορετικά φυλογενετικά κλαδιά, με τιμές bootstrap > 70% (Εικόνα 3.4). Η νουκλεοτιδική και φυλογενετική μελέτη που πραγματοποιήθηκε έδειξε ότι τα δείγματα A3, A15, PNL6, PNL7, τα οποία έχουν χαρακτηριστεί ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου ομαδοποιούνται στο ίδιο φυλογενετικό κλαδί με την ενδοτυπική παραλλαγή Asian American. Επίσης, το δείγμα A4 (υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου) ομαδοποιείται με την παραλλαγή African type I, ενώ το δείγμα A17 (χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου) ομαδοποιείται με την παραλλαγή African type II (Εικόνα 3.4).

Συμπερασματικά, είναι η πρώτη εργασία που ασχολείται με την νουκλεοτιδική ανάλυση του E4 γονιδίου του HPV16. Με βάση τα αποτελέσματα προέκυψε ότι το E4 γονίδιο είναι μια μικρή περιοχή του γονιδιώματος του ιού μεγέθους 288bp, αρκετά πολυμορφική, η οποία μπορεί να ταξινομήσει το DNA του HPV16 σε τέσσερα διαφορετικά φυλογενετικά κλαδιά και να ταυτοποιήσει τις διαφορετικές παραλλαγές του γονιδιώματος του ιού.



Εικόνα 3.4. Φυλογενετικό δέντρο (Μέγιστη Πιθανοφάνεια, 1000 bootstrap values) στο E4 γονίδιο. Το E4 γονίδιο μεγέθους 288bp, μπορεί να ταξινομήσει το DNA των ιών HPV16 σε τέσσερα διαφορετικά φυλογενετικά κλαδιά που οργανώνονται οι διαφορετικές παραλλαγές του HPV16.

		E4															Ενδοτοπικές Παραλλαγές	Αριθμοί καταχώρησης στη Genbank
Αλλοίωση	Κλώνοι	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3		
		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3		
		6	6	6	7	8	1	1	4	7	7	1	1	3	5	5		
		0	1	2	7	7	0	3	9	6	6	6	7	8	6	6		
		G	C	A	C	T	C	G	G	C	C	C	T	A	C	T		
	A3 288 C1	-	-	G	G	C	T	A	A	-	-	A	C	C	-	G	Asian American	HQ702334
	A3 288 C2	-	T	G	G	C	T	A	A	-	-	A	C	C	-	G	Asian American	HQ702335
	A4 288 C1	-	-	G	G	-	T	-	A	-	-	A	-	-	-	G	African type I	HQ702336
	A4 288 C2	-	-	G	G	-	T	-	A	-	-	A	-	-	T	G	African type I	HQ702337
	A6 288 C1	-	-	-	-	-	T	-	-	-	-	A	-	-	-	-	European	HQ702338
	A13 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702339
	A13 288 C2	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702340
	A15 288 C1	-	-	-	G	C	-	A	A	T	-	A	C	C	-	G	Asian American	HQ702341
	A31 288 C1	-	-	-	-	-	-	-	-	-	A	A	-	-	-	-	European	HQ702342
	D2 288 C1	-	-	-	-	-	T	-	-	-	-	A	-	-	-	-	European	HQ693893
	D4 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ693894
Υψηλού βαθμού	D14 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693891
	D16 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ693892
	PNL4 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032673
	PNL4 288 C3	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032674
	PNL5 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032668
	PNL5 288 C2	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032669
	PNL6 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032665
	PNL6 288 C2	C	-	G	G	C	T	-	A	-	-	A	C	C	T	G	Asian American	JN032666
	PNL6 288 C3	-	-	-	-	-	T	A	-	-	-	-	-	-	-	-	European	JN032667
	PNL7 288 C1	C	-	G	G	C	T	-	A	-	-	A	C	C	T	G	Asian American	JN032670
	PNL7 288 C3	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032671
	PNL11 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Prototype	JN032672
	E4 πρωτεΐνη	A7P	A7V	A7A	L12L	L16L	T23T	P24P	P33P	D45E	D45E	L59I	L59T	Q66P	S72F	H75Q		
	E2 πρωτεΐνη	S202S	S202S	N2023D	P208A	I211T	P219S	A220T	E232K	P241S	P241T	T254N	T254N	S261S	I267I	F271V		

Πίνακας 3.6. Νουκλεοτιδική και αμινοξική ανάλυση του E4 γονιδίου σε δείγματα υψηλού βαθμού αλλοίωσης. Ο αριθμός 288 αντιστοιχεί στο μέγεθος των προϊόντων και το τελευταίο γράμμα (C1 ή C2) αντιστοιχεί στον αριθμό του κλώνου. Επιπλέον αναγράφονται οι αντίστοιχες αμινοξικές μεταλλάξεις της E2 πρωτεΐνης

		E4															Ενδοτοπικές Παραλλαγές	Αριθμοί καταχώρησης στη Genbank
Αλλοίωση	Κλώνοι	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3		
		3	3	3	3	3	4	4	4	4	5	5	5	5	5	3		
		6	6	6	7	8	1	1	4	6	7	1	1	3	5	6		
		0	1	2	7	7	0	3	9	2	6	6	7	8	6	6		
		G	C	A	C	T	C	G	G	C	C	C	T	A	C	T		
	PNL2 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032663
	PNL2 288 C3	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	JN032664
	D7 288 C1	-	-	-	-	-	T	-	-	T	-	-	-	-	-	-	European	HQ693895
	D7 288 C2	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693896
	D34 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693897
	D34 288 C2	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ683998
	D42 288 C1	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ693899
	D42 288 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HQ693900
	747 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693905
	747 288 C2	-	-	-	-	-	T	-	-	T	-	-	-	-	-	-	European	HQ693906
Χαμηλού βαθμού	747 288 C3	-	-	-	-	A	T	-	-	-	-	-	-	-	-	-	European	HQ702327
	747 288 C4	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702328
	747 288 C5	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702329
	796 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	Prototype	HQ702330
	858 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702331
	868 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702332
	793 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ702333
	A17 288 C1	-	T	G	G	A	T	-	A	-	-	A	C	C	-	G	African type II	HQ693901
	A17 288 C2	-	T	G	G	-	T	-	A	-	-	A	C	C	-	G	African type II	HQ693902
	A24 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693904
	A38 288 C1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	European	HQ693903
	European	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-		
	Asian	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-		
	African type I	-	-	G	G	-	T	-	A	-	-	A	-	C	-	G		
African type II	-	-	G	G	-	T	-	A	-	-	A	C	C	-	G			
Asian American	-	-	G	G	C	T	A	A	-	-	A	C	C	-	G			
E4 protein	A7P	A7V	A7A	L12L	L16I	T23T	P24P	P33P	R40R	D45E	L59I	L59T	Q66P	S72F	H75Q			
E2 protein	S202S	S202S	N2023D	P208A	I211V	P219S	A220T	E232K	T236M	P241T	T254N	T254N	S261S	I267I	F271V			

Πίνακας 3.7. Νουκλεοτιδική και αμινοξική ανάλυση του E4 γονιδίου σε δείγματα χαμηλού βαθμού αλλοίωσης. Ο αριθμός 288 αντιστοιχεί στο μέγεθος των προϊόντων και το τελευταίο γράμμα (C1 ή C2) αντιστοιχεί στον αριθμό του κλώνου. Επιπλέον αναγράφονται οι αντίστοιχες αμινοξικές μεταλλάξεις της E2 πρωτεΐνης.

3.1.4.2 Εξελικτική μελέτη του αλληλεπικαλυπτόμενου πλαισίου ανάγνωσης των E4 και E2 γονιδίων

Όπως αναφέρθηκε σε προηγούμενη ενότητα το ανοιχτό πλαίσιο ανάγνωσης του E4 γονιδίου τοποθετείται εντός του ανοιχτού πλαισίου ανάγνωσης του E2 γονιδίου, που κωδικοποιεί για τη περιοχή "αρμό" της E2 πρωτεΐνης. Για το σκοπό αυτό πραγματοποιήθηκε υπολογισμός της επιλεκτικής πίεσης, που ασκείται στην συγκεκριμένη περιοχή του γονιδιώματος του ιού στα δύο διαφορετικά πλαίσια ανάγνωσης των γονιδίων E4 και E2, αντίστοιχα. Από την εξελικτική μελέτη που πραγματοποιήθηκε για το E4 γονίδιο παρατηρήθηκε ότι λόγος $dN/dS=0.97$ και ο μέσος ρυθμός υποκατάστασης για τις τρεις θέσεις του κωδικονίου είναι $CP1=0.929$, $CP2=1.023$, $CP3=1.048$, μετά από πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του E4 γονιδίου, τις αλληλουχίες της παρούσας διατριβής και όλες τις διαθέσιμες αλληλουχίες του E4 γονιδίου στη βάση δεδομένων GenBank με αριθμούς καταχώρησης: AF472509, AF402678, AF472508, AF536180, AF534061, AF536179, FJ610150, FJ610151, FJ610147, AY686583, FJ610149, FJ610146, AY686581, AF125673, FJ610152, FJ610148, FJ006723, AY686584, AY686580, AF193429, DQ191463, AF067034, AF193425, AF193427, AF115832, DQ191465, AY686579, NC001526, AY686582, AF067035, AF115831, EU430675, EU430667, EU430683, EU118173, AF515620, AF067029, DQ191462, AF193430, AF193426, AF067032, AF067028, AF067030, K02718, DQ191461, AF193428, AF067031, EU918764, και HM057182. Επιπλέον, από την εξελικτική μελέτη παρατηρήθηκε θετική επιλεκτική πίεση στα κωδικόνια 59 (L59I/T) και 66 (Q66P) με τη μέθοδο της Μέγιστης Πιθανοφάνειας και τη μέθοδο της Μπεϋζιανής Συμπερασματολογίας. Όπως προέκυψε από τη νουκλεοτιδική και αμινοξική μελέτη του E4 γονιδίου, οι θέσεις 59 και 66 διαχωρίζουν τις HPV16 ενδοτυπικές παραλλαγές African type I, African type II και Asian American από τις ενδοτυπικές παραλλαγές Asian και European (Πίνακας 3.6, Πίνακας 3.7).

Για το πλαίσιο ανάγνωσης του E2 γονιδίου στο αλληλεπικαλυπτόμενο τμήμα των E2 και E4 γονιδίων ο λόγος $dN/dS=1$ και ο μέσος ρυθμός υποκατάστασης για τις τρεις θέσεις του κωδικονίου είναι $CP1=1.056$, $CP2=0.927$ και $CP3=1.018$, μετά από πολλαπλή ομοπαράθεση των αλληλουχιών που περιγράφηκαν παραπάνω, αλλά με το πλαίσιο ανάγνωσης του E2 γονιδίου. Ακόμη, δεν ανιχνεύτηκε θέση με θετική επιλεκτική πίεση στο ανοιχτό πλαίσιο ανάγνωσης του E2 γονιδίου. Με βάση τα

αποτελέσματα από τον λόγο των τιμών dN/dS, παρατηρήθηκε ότι ουδέτερη επιλεκτική πίεση ασκείται ταυτόχρονα στα δύο αλληλεπικαλυπτόμενα τμήματα του γονιδιώματος του HPV16.

3.1.5 Νουκλεοτιδική ανάλυση του E1 γονιδίου

Για την ανάλυση του E1 γονιδίου χρησιμοποιήθηκαν τριάντα κλινικά δείγματα (πέντε τραχηλικοί ιστοί μονιμοποιημένοι σε παραφίνη, τρία δείγματα από τραχηλικά επιχρίσματα και είκοσι δύο δείγματα ThinPrep). Από τα τριάντα δείγματα, δεκαέξι δείγματα χαρακτηρίστηκαν από τους κυτταροπαθολόγους ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και δεκατρείς περιπτώσεις χαρακτηρίστηκαν ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Επιπρόσθετα μια περίπτωση χαρακτηρίστηκε ως καρκίνος του τραχήλου της μήτρας. Η ενίσχυση του E1 γονιδίου πραγματοποιήθηκε με τη μέθοδο της PCR, χρησιμοποιώντας τους εσωτερικούς αλληλεπικαλυπτόμενους εκκινητές E1F1-E1R3, E1F4-E1R5, E1F6-E1R7, E1F8-E1R9, οι οποίοι καλύπτουν σε επιμέρους τμήματα ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου (Πίνακας 2.4).

Η περιοχή μεταξύ των νουκλεοτιδίων 878 και 1523 δεν ενισχύθηκε στα δείγματα υψηλού βαθμού αλλοίωσης ATT9, D2, D13, D14, και D15, ενώ η περιοχή αυτή δεν ενισχύθηκε σε δύο δείγματα χαμηλού βαθμού αλλοίωσης 747 και D7. Η ενίσχυση τμήματος του γονιδίου της β-ακτίνης μεγέθους 498 bp (Li et al., 2008) σε όλα τα δείγματα επιβεβαίωσε ότι η ρήξη του E1 γονιδίου δεν προέρχεται από κατακερματισμό του γονιδιώματος του ιού αλλά από την ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα.

Μέχρι στιγμής νουκλεοτιδική ανάλυση του E1 γονιδίου δεν είχε πραγματοποιηθεί σε προηγούμενες μελέτες και εντοπίστηκαν μεταλλάξεις οι οποίες αναφέρθηκαν για πρώτη φορά στη βιβλιογραφία. Πιο συγκεκριμένα, η ανάλυση της αλληλουχίας του E1 γονιδίου μεταξύ των νουκλεοτιδίων 858 και 1523 έδειξε την ύπαρξη των νουκλεοτιδικών μεταλλάξεων T921C, C1096G, G1163A, T1200C, T1366A, C1377T, C1426G, T1486C, οι οποίες ανιχνεύτηκαν στα δείγματα A17 (χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση) και PNL7 (καρκίνος) (Πίνακας 3.8). Οι νουκλεοτιδικές αλλαγές T921C, T1200C, C1377T και T1486C, δε προκαλούν αμινοξικές υποκαταστάσεις στην E1 πρωτεΐνη, ενώ οι μεταλλάξεις C1096G, G1163A, T1366A και C1426G προκαλούν τις αμινοξικές μεταλλάξεις Q78E, G100E, C168S και Q188E, αντίστοιχα. Οι μεταλλάξεις αυτές εντοπίζονται στο N-τελικό άκρο

της E1 πρωτεΐνης. Επιπλέον, η σιωπηλή μετάλλαξη A1041G ανιχνεύτηκε μόνο στο δείγμα PNL7, ενώ η σιωπηλή μετάλλαξη C1416T ανιχνεύτηκε μόνο στο δείγμα A17 (Πίνακας 3.8).

Μεταξύ των νουκλεοτιδίων 1503 και 1991 παρατηρήθηκαν οι νουκλεοτιδικές μεταλλάξεις C1624T, C1744A, A1842G, οι οποίες εντοπίστηκαν στα δείγματα PNL7, A17 (Πίνακας 3.9). Η νουκλεοτιδική μετάλλαξη C1624T δεν προκαλεί αμινοξική αλλαγή, ενώ οι μεταλλάξεις C1744A και A1842G προκαλούν τις αμινοξικές αλλαγές L294M και I326M, αντίστοιχα. Οι μεταλλάξεις αυτές εντοπίζονται στη κεντρική περιοχή της E1 πρωτεΐνης, η οποία εμπλέκεται στην πρόσδεση της πρωτεΐνης με το DNA του ιού. Επιπλέον, στη περιοχή αυτή εντοπίστηκε η νουκλεοτιδική μετάλλαξη A1668G, η οποία χαρακτηρίστηκε ως σιωπηλή και ανιχνεύτηκε στο δείγμα PNL7 (Πίνακας 3.9).

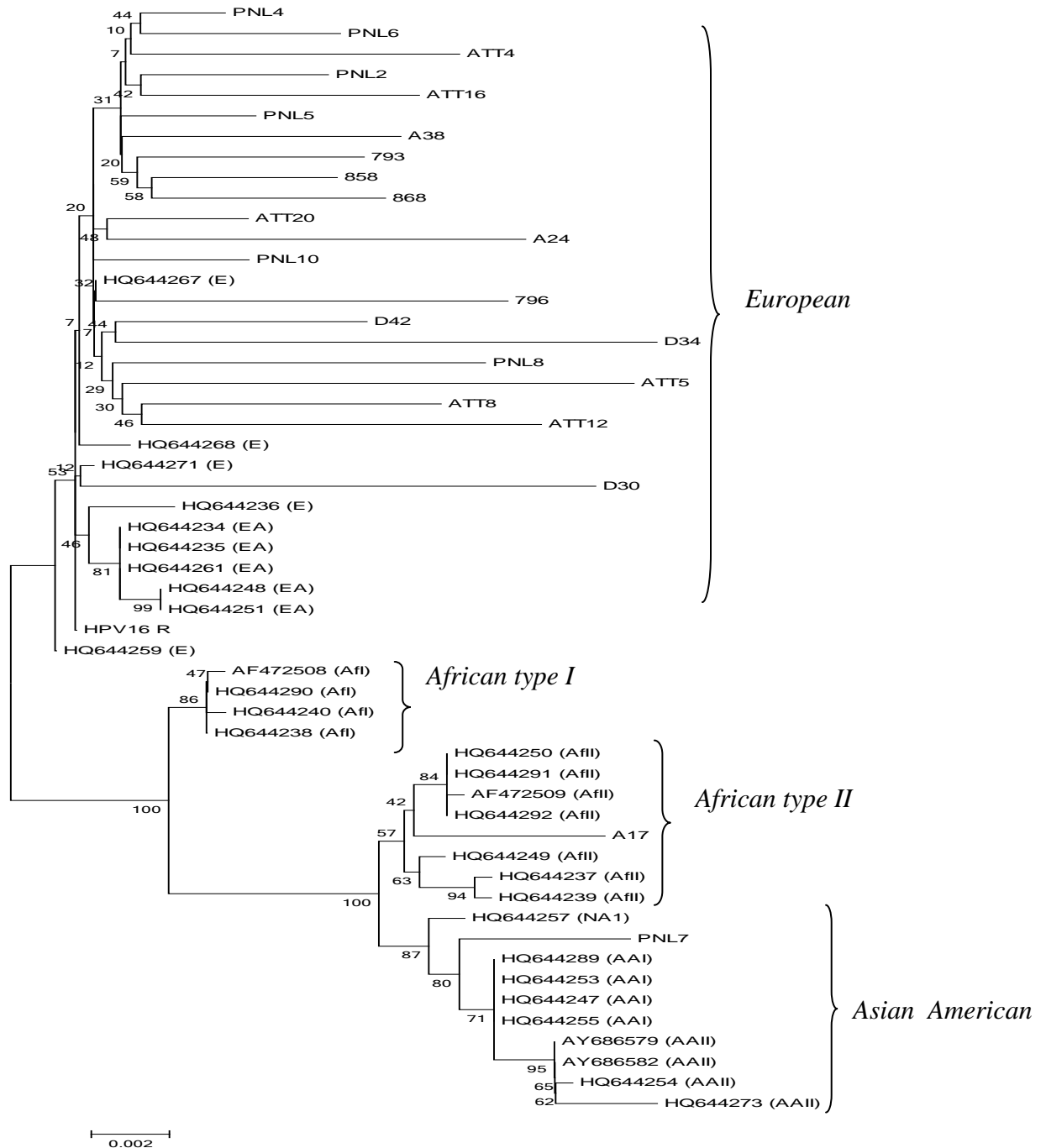
Επιπρόσθετα, με την ανάλυση του E1 γονιδίου μεταξύ των νουκλεοτιδίων 1952 και 2485 εντοπίστηκαν οι νουκλεοτιδικές μεταλλάξεις G2073A, T2169C, T2189C, A2453T και C2454T, οι οποίες ανιχνεύτηκαν σε υψηλότερου βαθμού δυσπλασίες (D2, D13, D15) (Πίνακας 3.10). Ακόμα, ανιχνεύτηκαν οι νουκλεοτιδικές μεταλλάξεις C2041T, G2220C, C2237G, G2249A, C2262T, C2287T και C2344T, εκ των οποίων οι C2041T, G2220C, C2237G και G2249A προκαλούν τις αμινοξικές μεταλλάξεις L393I, E425D, T458S, και R462K, ενώ οι C2262T, C2287T και C2344T χαρακτηρίστηκαν ως σιωπηλές μεταλλάξεις. Αυτές οι μεταλλάξεις αναφέρθηκαν στα δείγματα PNL7 και A17 (Πίνακας 3.10). Επιπρόσθετα η μετάλλαξη T2355C ανιχνεύτηκε μόνο στο δείγμα A17 (Πίνακας 3.10). Τέλος, η περιοχή μεταξύ των νουκλεοτιδίων 2443 and 2848 φαίνεται να είναι πιο συντηρημένη. Οι μεταλλάξεις A2587T (R575W) και G2650A (E596K) εντοπίστηκαν σε υψηλότερου βαθμού δυσπλασίες και καρκίνο (PNL7, D13, D14, and D30). Οι σιωπηλές μεταλλάξεις T2586C, A2608C και T2631T ανιχνεύτηκαν στα δείγματα A17 και PNL7 (Πίνακας 3.11), ενώ η νουκλεοτιδική μετάλλαξη G2650A ανιχνεύτηκε μόνο στο δείγμα PNL7 (Πίνακας 3.11). Οι μεταλλάξεις αυτές εντοπίστηκαν στη περιοχή του E1 γονιδίου που κωδικοποιεί για το C-τελικό άκρο της E1 πρωτεΐνης.

3.1.5.1 Φυλογενετική ανάλυση του HPV16 DNA με βάση την αλληλουχία του E1 γονιδίου

Για να συνδεθούν οι μεταλλάξεις που εντοπίστηκαν στο E1 γονίδιο με τις ενδοτυπικές παραλλαγές του ιού πραγματοποιήθηκε συγκριτική μελέτη της αλληλουχίας του E1 γονιδίου με την αλληλουχία των E6 και E7 γονιδίων στα ίδια κλινικά δείγματα. Πιο συγκεκριμένα, οι μεταλλάξεις T921C, C1096G, G1163A, T1200C, T1366A, C1377T, C1426G, T1486C, C1624T, C1744A, A1842G, C2041T, G2220C, C2237G, G2249A, C2262T, C2287T, C2344T, T2586C, A2608C, T2631T, συνδέονται με τους ενδοτυπικούς πολυμορφισμούς G145T, T286A, A289G, C335T, T789C, T795G που εντοπίζονται εντός της αλληλουχίας των E6 και E7 γονιδίων των ενδοτυπικών παραλλαγών Asian American και African type II. Επιπλέον, οι νουκλεοτιδικές αλλαγές A1041G, C1587T, A1668G, G2650A συνδέθηκαν μόνο με την ενδοτυπική παραλλαγή Asian American (ενδοτυπικοί πολυμορφισμοί E6, E7: A532G και T732C), ενώ οι μεταλλάξεις C1416T και T2355C συσχετίστηκαν μόνο με την ενδοτυπική παραλλαγή African type II (ενδοτυπικός πολυμορφισμός E6: A403G).

Στη συνέχεια, για να προσδιοριστεί ότι οι ενδοτυπικές μεταλλάξεις που εντοπίζονται στο E1 γονίδιο μπορούν να διαχωρίσουν τις παραλλαγές του γονιδιώματος του ιού στα τέσσερα μεγάλα φυλογενετικά κλαδιά κατασκευάστηκε ένα φυλογενετικό δέντρο με τη μέθοδο της Μέγιστης Πιθανοφάνειας (1000 bootstrap values, Μοντέλο Υποκατάστασης GTR + G), χρησιμοποιώντας τις αλληλουχίες της παρούσας διατριβής, τη πρότυπη αλληλουχία του E1 γονιδίου και τις αλληλουχίες των ενδοτυπικών παραλλαγών που είναι διαθέσιμες στη βάση δεδομένων GenBank με αριθμούς καταχώρησης: European: HQ644236, HQ644259, HQ644267, HQ644268, HQ644271, Asian: HQ644234, HQ644235, HQ644248, HQ644251, HQ644261, African type I: HQ644238, HQ644240, AF472508, HQ644290, African type II: HQ644237, HQ644239, HQ644249, HQ644250, AF472509, HQ644291, HQ644292, Asian American: HQ644289, HQ644247, HQ644253, HQ644255, HQ644254, AY686579, AY686582, HQ644273, HQ644257). Με βάση τα αποτελέσματα από το φυλογενετικό δέντρο παρατηρήθηκε ότι η αλληλουχία του E1 γονιδίου μπορεί να ταξινομήσει το γονιδίωμα του ιού στα τέσσερα διαφορετικά φυλογενετικά κλαδιά, που οργανώνονται οι πέντε διαφορετικές παραλλαγές του HPV16 DNA με τιμές bootstrap > 80% (Εικόνα 3.5). Συμπερασματικά, είναι η πρώτη εργασία που ασχολείται με την νουκλεοτιδική και αμινοξική ανάλυση ολόκληρου του E1 γονιδίου και αποδεικνύει ότι συγκεκριμένες μεταλλάξεις του E1 γονιδίου μπορούν

να ταυτοποιήσουν τις διαφορετικές παραλλαγές του γονιδιώματος του HPV16 (Εικόνα 3.5).



Εικόνα 3.5. Φυλογενετικό δέντρο κατασκευασμένο με τη μέθοδο της Μέγιστης Πιθανοφάνειας (Maximum Likelihood, 1000 bootstrap values). Για τη κατασκευή του φυλογενετικού δέντρου τα αλληλεπικαλυπτόμενα τμήματα του E1 γονιδίου κάθε δείγματος ενώθηκαν για να σχηματιστεί ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του γονιδίου. Το φυλογενετικό δέντρο κατασκευάστηκε μετά από πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του E1 γονιδίου, με τις αλληλουχίες των δειγμάτων και τις αλληλουχίες των ενδοτυπικών παραλλαγών του HPV16 γονιδιώματος.

878-1523	9	1	1	1	1	1	1	1	1	1	Ενδοτυπικές παραλλαγές	Αριθμοί Καταχώρησης στη Genbank
	2	0	0	1	2	3	3	4	4	4		
	1	1	6	3	0	6	7	1	2	8		
	T	A	C	G	T	T	C	C	C	T		
Υψηλού βαθμού αλλοίωση												
PNL4 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791086
PNL5 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791089
PNL6 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791087
PNL6 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791088
PNL7 645 C1	C	G	G	A	C	A	T	-	G	C	Asian American	JQ791066
PNL8 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791067
PNL10 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791069
ATT4 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791090
ATT5 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791071
ATT12 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791075
ATT16 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791091
ATT20 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791073
D30 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791079
Χαμηλού βαθμού αλλοίωση												
PNL2 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791084
ATT8 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791070
A17 645 C1	C	-	G	A	C	A	T	T	G	C	African type II	JQ791076
A24 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791080
A38 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791092
793 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791095
796 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791082
858 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791094
868 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791097
D34 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791078
D42 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791077
Ενδοτυπικές Παραλλαγές												
European	-	-	-	-	-	-	-	-	-	-		
African type I	-	-	G	-	-	A	T	-	-	C		
African type II	C	-	G	A	C	A	T	T	G	C		
Asian American	C	G	G	A	C	A	T	-	G	C		
Asian	-	-	-	-	-	-	-	-	-	-		
Αμινοξέα	F19F	L59L	Q78E	G100E	A112A	C168S	Y171Y	H184H	Q188E	L208L		

Πίνακας 3.8. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων στην περιοχή του E1 γονιδίου μεταξύ των νουκλεοτιδίων 878 έως 1523. Ο αριθμός 645 αντιστοιχεί στο μέγεθος του προϊόντος της PCR και το τελευταίο γράμμα (C1) αντιστοιχεί στον αριθμό του κλώνου. Με βάση τη νουκλεοτιδική ανάλυση οκτώ μεταλλάξεις ανιχνεύθηκαν και συσχετίστηκαν με τις παραλλαγές Asian American και African type II. Επίσης η νουκλεοτιδική μετάλλαξη A1041G ανιχνεύτηκε μόνο στη παραλλαγή Asian American, ενώ η μετάταξη C1416T ανιχνεύτηκε μόνο στη παραλλαγή African type II.

	1	1	1	1	1		Αριθμοί καταχώρησης στη Genbank
1503-1991	5	6	6	7	8	Ενδοτυπικές παραλλαγές	
	8	2	6	4	4		
	7	4	8	4	2		
	T	C	A	C	A		
Υψηλού βαθμού αλλοίωση							
PNL4 488 C1	-	-	-	-	-	European	JQ768483
PNL5 488 C1	-	-	-	-	-	European	JQ768485
PNL6 488 C1	-	-	-	-	-	European	JQ768487
PNL7 488 C1	-	T	G	A	G	Asian American	JQ768488
PNL8 488 C1	C	-	-	-	-	European	JQ768490
PNL10 488 C1	-	T	G	A	G	European	JQ768491
ATT4 488 C1	-	-	-	-	-	European	JQ768492
ATT5 488 C1	C	-	-	-	-	European	JQ768504
ATT9 488 C1	C	-	-	-	-	European	JQ768496
ATT12 488 C1	C	-	-	-	-	European	JQ768498
ATT16 488 C1	-	-	-	-	-	European	JQ768500
ATT20 488 C1	-	-	-	-	-	European	JQ768502
D2 488 C1	-	-	-	-	-	European	JQ768462
D13 488 C1	-	-	-	-	-	European	JQ768476
D14 488 C1	-	-	-	-	-	European	JQ768477
D15 488 C1	-	-	-	-	-	European	JQ768478
D30 488 C1	-	-	-	-	-	European	JQ768463
Χαμηλού βαθμού αλλοίωση							
PNL2 488 C1	-	-	-	-	-	European	JQ768481
ATT8 488 C1	C	-	-	-	-	European	JQ768494
A17 488 C1	-	T	-	A	G	African type II	JQ768467
A24 488 C1	-	-	-	-	-	European	JX297818
A38 488 C1	-	-	-	-	-	European	JQ768468
747 488 C1	-	-	-	-	-	European	JQ768479
793 488 C1	-	-	-	-	-	European	JQ768470
796 488 C1	-	-	-	-	-	European	JQ768474
858 488 C1	-	-	-	-	-	European	JQ768472
868 488 C1	-	-	-	-	-	European	JQ768480
D7 488 C1	-	-	-	-	-	European	JQ768465
D34 488 C1	-	-	-	-	-	European	JQ768473
D42 488 C1	-	-	-	-	-	European	JQ768475
Ενδοτυπικές παραλλαγές							
European	-	-	-	-	-		
African type I	-	T	-	T	G		
African type II	-	T	-	A	G		
Asian American	C	T	G	A	G		
Asian	-	-	-	-	G		
Αμινοξέα	F241F	L254L	A268A	L294M	I326M		

Πίνακας 3.9. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων στην περιοχή μεταξύ των νουκλεοτιδίων 1503 έως 1991. Ο αριθμός 488 αντιστοιχεί στο μέγεθος του προϊόντος της PCR και το τελευταίο γράμμα (C1) αντιστοιχεί στον αριθμό του κλώνου. Με βάση τη νουκλεοτιδική ανάλυση τρεις μεταλλάξεις ανιχνεύθηκαν και συνδέθηκαν με τις παραλλαγές Asian American και African type II. Οι νουκλεοτιδικές μεταλλάξεις T1587C και A1668G ανιχνεύθηκαν μόνο στη παραλλαγή Asian American.

	2	2	2	2	2	2	2	2	2	2	2	2	2	2	Ενδοτυπικές παραλλαγές	Αριθμοί καταχώρισης στη Genbank
1952-2485	0	0	1	1	2	2	2	2	2	3	4	3	4			
	4	7	6	8	2	3	4	6	8	4	5	5	5			
	1	3	9	9	0	7	9	2	7	4	3	5	4			
	C	G	T	T	G	C	G	C	C	C	A	T	C			
Υψηλού βαθμού αλλοίωση																
PNL4 533 C1	-	A	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791030
PNL5 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791032
PNL6 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791034
PNL7 533 C1	T	-	-	-	C	G	A	T	T	T	-	-	-	Asian American	JQ791035	
PNL8 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791037
PNL10 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791036
ATT4 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791017
ATT5 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791019
ATT9 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791021
ATT12 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791022
ATT16 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791024
ATT20 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791026
D2 533 C1	-	A	C	C	-	-	-	-	-	-	T	-	T	-	European	JQ791038
D13 533 C1	-	A	C	C	-	-	-	-	-	-	T	-	T	-	European	JQ791053
D14 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791044
D15 533 C1	-	A	C	C	-	-	-	-	-	-	T	-	T	-	European	JQ791052
D30 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791045
Χαμηλού βαθμού αλλοίωση																
PNL2 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791028
ATT8 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791020
A17 533 C1	T	-	-	-	C	G	A	T	T	T	-	C	-	African type II	JQ791062	
A24 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791061
A38 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791063
747 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791059
793 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791050
796 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791048
858 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791057
868 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791055
D7 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791043
D34 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791040
D42 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791047
Ενδοτυπικές παραλλαγές																
European	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
African type I	T	-	-	-	C	-	-	-	-	-	-	-	-			
African type II	T	-	-	-	C	G	A	T	T	T	-	C	-			
Asian American	T	-	-	-	C	G	A	T	T	T	-	-	-			
Asian	-	-	-	-	-	-	-	-	-	-	-	-	-			
Αμινοξέα	L393I	K403K	D435D	I442T	E425D	T458S	R462K	G466G	L475L	L494L	N530I	S497S	N530I			

Πίνακας 3.10. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων στην περιοχή μεταξύ των νουκλεοτιδίων 1952 έως 2485. Ο αριθμός 533 αντιστοιχεί στο μέγεθος του προϊόντος της PCR και το τελευταίο γράμμα (C1) αντιστοιχεί στον αριθμό του κλώνου. Με βάση τη νουκλεοτιδική ανάλυση επτά μεταλλάξεις ανιχνεύθηκαν και συνδέθηκαν με τις παραλλαγές Asian American και African type II. Η μετάλλαξη T2355C ανιχνεύτηκε μόνο στη παραλλαγή African type II.

	2	2	2	2	2		Αριθμοί καταχώρησης στη Genbank
2423-2848	5	5	6	6	6	Ενδοτυπικές παραλλαγές	
	8	8	0	3	5		
	6	7	8	1	0		
	T	A	A	T	G		
Υψηλού βαθμού αλλοίωση							
PNL4 405 C1	-	-	-	-	-	European	JQ779917
PNL5 405 C1	-	-	-	-	-	European	JQ779919
PNL6 405 C1	-	-	-	-	-	European	JQ779921
PNL7 405 C1	C	-	C	A	A	Asian American	JQ779923
PNL8 405 C1	-	-	-	-	-	European	JQ779925
PNL10 405 C1	-	-	-	-	-	European	JQ779927
ATT4 405 C1	-	-	-	-	-	European	JQ779933
ATT5 405 C1	-	-	-	-	-	European	JQ779934
ATT9 405 C1	-	-	-	-	-	European	JQ779937
ATT12 405 C1	-	-	-	-	-	European	JQ779939
ATT16 405 C1	-	-	-	-	-	European	JQ779941
ATT20 405 C1	-	-	-	-	-	European	JQ779942
D2 405 C1	-	-	-	-	-	European	JQ779961
D13 405 C2	-	T	-	-	-	European	JQ779962
D14 405 C1	-	T	-	-	-	European	JQ779963
D15 405 C1	-	-	-	-	-	European	JQ779960
D30 405 C1	-	T	-	-	-	European	JQ779951
Χαμηλού βαθμού αλλοίωση							
PNL2 405 C1	-	-	-	-	-	European	JQ779915
ATT8 405 C1	-	-	-	-	-	European	JQ779936
A17 405 C1	C	-	C	A	-	African type II	JQ779944
A24 405 C2	-	-	-	-	-	European	JQ779946
A38 405 C1	-	-	-	-	-	European	JQ779947
747 405 C1	-	-	-	-	-	European	JQ779967
793 405 C1	-	-	-	-	-	European	JQ779929
796 405 C1	-	-	-	-	-	European	JQ779931
858 405 C1	-	-	-	-	-	European	JQ779965
868 405 C1	-	-	-	-	-	European	JQ779968
D7 405 C1	-	-	-	-	-	European	JQ779957
D34 405 C1	-	-	-	-	-	European	JQ779955
D42 405 C1	-	-	-	-	-	European	JQ779949
Ενδοτυπικές παραλλαγές							
European	-	-	-	-	-		
African type I	C	-	-	A	-		
African type II	C	-	C	A	-		
Asian American	C	-	C	A	A		
Asian	-	-	-	-	-		
Αμινοξέα	S574S	R575W	R582R	P589P	E596K		

Πίνακας 3.11. Παρουσίαση των νουκλεοτιδικών και αμινοξικών μεταλλάξεων στην περιοχή μεταξύ των νουκλεοτιδίων 2423 έως 2848. Ο αριθμός 405 αντιστοιχεί στο μέγεθος του προϊόντος της PCR και το τελευταίο γράμμα (C1) αντιστοιχεί στον αριθμό του κλώνου. Με βάση τη νουκλεοτιδική ανάλυση τρεις μεταλλάξεις ανιχνεύθηκαν και συνδέθηκαν με τις παραλλαγές Asian American και African type II. Η μετάλλαξη G2650A ανιχνεύτηκε μόνο στη παραλλαγή Asian American.

3.1.5.2 Εξελικτική μελέτη του E1 γονιδίου

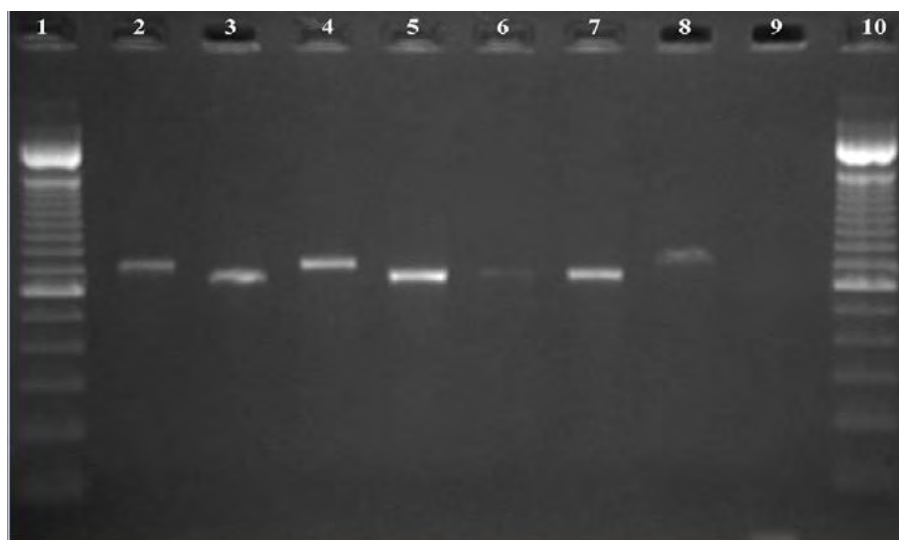
Η επιλεκτική πίεση που ασκείται στα τέσσερα αλληλοεπικαλυπτόμενα τμήματα του E1 γονιδίου μελετήθηκε μέσω του λόγου dN/dS μετά από πολλαπλή ομοπαράθεση της πρότυπης E1 αλληλουχίας, τις αλληλουχίες της παρούσας διατριβής και τις αλληλουχίες των ενδοτυπικών παραλλαγών που χρησιμοποιήθηκαν για τη κατασκευή του φυλογενετικού δέντρου. Στη περιοχή μεταξύ των νουκλεοτιδίων 878 και 1523 παρατηρήθηκε ότι ο λόγος dN/dS=0.68 και ο μέσος ρυθμός υποκατάστασης των τριών θέσεων ενός κωδικονίου ήταν CP1=0.823, CP2=1.027, CP3=1.15. Επιπρόσθετα, στη περιοχή μεταξύ των νουκλεοτιδίων 1503 και 1991 ο λόγος dN/dS=0.67 και ο μέσος ρυθμός υποκατάστασης ήταν CP1=1.075, CP2=0.813, CP3=1.13. Στη περιοχή μεταξύ των νουκλεοτιδίων 1952 και 2485 ο λόγος dN/dS=0.57 με CP1=0.77, CP2=0.881 and CP3=1.349, ενώ στη περιοχή μεταξύ των νουκλεοτιδίων 2443 και 2848 ο λόγος dN/dS=0.9 με CP1=0.878, CP2=0.916 και CP3=1.006. Επιπλέον, σύμφωνα με την ανάλυση που πραγματοποιήθηκε μέσω της μεθόδου Μέγιστης Πιθανοφάνειας και η Μπεϋζιανής Συμπερασματολογίας, δεν ανιχνεύτηκαν θέσεις που επιδρούν με θετική επιλεκτική πίεση.

Η ανάλυση της επιλεκτικής πίεσης που ασκείται στο E1 γονίδιο, έδειξε ότι αρνητική επιλογή ασκείται στο E1 γονίδιο, καθώς ο λόγος dN/dS < 1 και ο μέσος ρυθμός υποκατάστασης του τρίτου νουκλεοτιδίου του κωδικονίου είναι υψηλότερος από το πρώτο και το δεύτερο. Το γεγονός αυτό αποδεικνύει ότι επικρατούν οι συνώνυμες μεταλλάξεις και το E1 γονίδιο είναι μια εξαιρετικά συντηρημένη περιοχή του γονιδιώματος του ιού. Το φαινόμενο αυτό ίσως να οφείλεται στο κρίσιμο ρόλο που διαδραματίζει η E1 πρωτεΐνη στην αντιγραφή του HPV16 DNA.

3.1.6 Ανάλυση διπλασιασμού στο γονίδιο E1.

Η ανάλυση του E1 γονιδίου οδήγησε στην ανίχνευση ενός ενδογονιδιακού διπλασιασμού 63 νουκλεοτιδίων (E1-1374^Δ63nt) στο 5' άκρο του γονιδίου. Πιο συγκεκριμένα, η περιοχή μεταξύ των νουκλεοτιδίων 1311 και 1374 βρέθηκε διαδοχικά διπλασιασμένη χωρίς να καταστρέφει το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου. Επιπλέον, ο ενδογονιδιακός διπλασιασμός προκαλεί την επανάληψη της θέσης μεταξύ των αμινοξέων 150 έως 170 στο N-τελικό άκρο της πρωτεΐνης. Η ανίχνευση του διπλασιασμού οδήγησε στη μελέτη της επικράτησης του σε πενήντα τραχηλικά κλινικά δείγματα που προήρθαν από τον Ελληνικό πληθυσμό και στον προσδιορισμό της σημασίας του διπλασιασμού στη λειτουργία και τη δομή της E1 πρωτεΐνης. Για το σκοπό αυτό μελετήθηκαν συνολικά πενήντα κλινικά δείγματα, τα οποία προήρθαν από ThinPrep και τραχηλικά επιχρίσματα. Από τα πενήντα δείγματα, είκοσι επτά χαρακτηρίστηκαν ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου, και είκοσι τρία δείγματα ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Η ανίχνευση του διπλασιασμού έγινε με τη μέθοδο της PCR και τη χρήση των εκκινητών E1F1-E1R3 (Πίνακας 2.4) Οι εκκινητές υβριδίζονται σε θέσεις που περιλαμβάνουν τη περιοχή του διπλασιασμού. Το προϊόν της αντίδρασης έχει μέγεθος 645bp, ενώ το προϊόν που περιλαμβάνει το διπλασιασμό έχει μέγεθος 708bp.

Η ενίσχυση του 5' άκρου του E1 γονιδίου με τους εκκινητές E1F1-E1R1, οδήγησε στην ανίχνευση της ζώνης ενίσχυσης μεγέθους 708bp στο 24% των εξεταζόμενων δειγμάτων (12/50), ενώ το 76% των δειγμάτων έφεραν την κανονική ζώνη ενίσχυσης μεγέθους 645bp (Εικόνα 3.6). Η ανάλυση της αλληλουχίας της ζώνης μεγέθους 708bp σε όλα τα κλινικά δείγματα έδειξε την ύπαρξη διαδοχικού διπλασιασμού μεταξύ των νουκλεοτιδίων 1311 και 1374. Ο διπλασιασμός των 63 νουκλεοτιδίων ανιχνεύτηκε στο 30.4% των δειγμάτων που έχουν χαρακτηριστεί ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση (7/23) και στο 18% των δειγμάτων υψηλού βαθμού αλλοίωσης (5/27) (Πίνακας 3.12). Επιπλέον καμιά στατιστικά σημαντική συσχέτιση δεν παρατηρήθηκε μεταξύ της παρουσίας του διπλασιασμού και της σοβαρότητας της ενδοεπιθηλιακής αλλοίωσης



Εικόνα 3.6. Οι θέσεις 2, 4 και 8 αντιστοιχούν στα ενισχυμένα προϊόντα που φέρουν το διπλασιασμό και έχουν μέγεθος 708bp, ενώ οι θέσεις 3,5, και 7 αντιστοιχούν στα ενισχυμένα προϊόντα χωρίς τον διπλασιασμό και έχουν μέγεθος 645bp. Στη θέση 1 και 10 βρίσκεται ο μάρτυρας μοριακού βάρους (100 bp DNA Ladder, Invitrogen UK). Στη θέση 9 βρίσκεται ο αρνητικός μάρτυρας (dd H₂O)

Προηγούμενες μελέτες έχουν συνδέσει τον E1-1374^{63nt} διπλασιασμό με την νουκλεοτιδική μετάλλαξη T350G (E6 γονίδιο) και με την Ευρωπαϊκή ενδοτυπική παραλλαγή του γονιδιώματος του HPV16 (Sabol et al. 2008, Bogovac et al. 2011, Sabol et al. 2012). Για να εξακριβωθεί η σύνδεση του διπλασιασμού με τη μετάλλαξη T350G πραγματοποιήθηκε νουκλεοτιδική ανάλυση του E6 γονιδίου σε όλα τα δείγματα που έφεραν το συγκεκριμένο διπλασιασμό. Με βάση τη νουκλεοτιδική ανάλυση του E6 γονιδίου προτάθηκε, ότι όλα τα δείγματα που φέρουν το διπλασιασμό φέρουν επιπλέον τη μετάλλαξη T350G και έχουν νουκλεοτιδική ομοιότητα με την Ευρωπαϊκή παραλλαγή του γονιδιώματος του HPV16. Επιπρόσθετα, για να μελετηθεί η σχέση μεταξύ της μετάλλαξης T350G και του E1-1374^{63nt} διπλασιασμού, μελετήθηκε η αλληλουχία του E6 γονιδίου σε είκοσι ένα δείγματα που δεν έφεραν τον διπλασιασμό. Από την ανάλυση των αποτελεσμάτων παρατηρήθηκε ότι δώδεκα δείγματα έφεραν ομοιότητα με την Ευρωπαϊκή παραλλαγή του γονιδιώματος του HPV16 και έφεραν τη μετάλλαξη T350G. Η μετάλλαξη T350G δεν ανιχνεύτηκε σε επτά δείγματα που παρουσίασαν νουκλεοτιδική ομοιότητα με τη πρότυπη αλληλουχία του HPV16 (JQ422087, JQ422091, JQ422070, JQ422065, JQ422094, JQ422108, JQ422117) και δεν εντοπίστηκε σε δύο δείγματα τα οποία χαρακτηρίστηκαν ως ενδοτυπικές παραλλαγές African type I και African type II (JQ422059, JQ422056). Η στατιστική ανάλυση των αποτελεσμάτων έδειξε ότι ο E1-

1374⁶³nt διπλασιασμός συνδέεται με τη μετάλλαξη T350G και την Ευρωπαϊκή παραλλαγή του HPV16, όταν συγκρίνεται με τη παρουσία της μετάλλαξης T350G σε δείγματα που δε φέρουν το διπλασιασμό ($\chi^2 = 5.076$, $P = 0.024$). Παρά το γεγονός ότι η μετάλλαξη T350G συνδέεται με υψηλότερου βαθμού αλλοιώσεις και σχετίζεται με την παρουσία του διπλασιασμού E1-1374⁶³nt δεν παρατηρήθηκε σύνδεση του διπλασιασμού με υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις.

Επιπλέον από τη νουκλεοτιδική ανάλυση του E1 γονιδίου που φέρει τον ενδογονιδιακό διπλασιασμό ανιχνεύτηκε η μετάλλαξη A1053C. Η μετάλλαξη αυτή εντοπίστηκε σε όλες τις αλληλουχίες με τον διπλασιασμό και προκαλεί την αμινοξική αλλαγή E63D στο N τελικό άκρο της E1 πρωτεΐνης. Για να εξακριβωθεί η σχέση της μετάλλαξης A1053C με τον E1-1374⁶³nt διπλασιασμό αλληλουχίστηκαν δεκαπέντε δείγματα που δεν έφεραν τον ενδογονιδιακό διπλασιασμό (JQ791078, JQ791077, JQ791079, JQ791076, JQ791080, JQ791082, JQ791066, JQ791067, JQ791069, JQ791071, JQ791070, JQ791075, JQ791074, JX297816, and JX297817). Η μετάλλαξη A1053C δεν εντοπίστηκε σε δείγματα χωρίς τον διπλασιασμό E1-1374⁶³nt, και παρουσιάστηκε σημαντική σύνδεση της συγκεκριμένης μετάλλαξης με τον E1-1374⁶³nt διπλασιασμό ($\chi^2 = 23.102$, $P < 0.0001$). Ακόμη, για να εξακριβωθεί εάν η μετάλλαξη A1053C εντοπίζεται στις ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16, πραγματοποιήθηκε πολλαπλή ομοπαράθεση της πρότυπης αλληλουχίας του HPV16 με τις αλληλουχίες των ενδοτυπικών παραλλαγών. Από την νουκλεοτιδική ανάλυση του E1 γονιδίου των HPV16 ενδοτυπικών παραλλαγών δεν εντοπίστηκε η μετάλλαξη A1053C. Συμπερασματικά, με βάση τα αποτελέσματα της ανάλυσης του E1 γονιδίου παρατηρήθηκε ότι, ο διπλασιασμός των 63 νουκλεοτιδίων παρουσιάζει σημαντική σύνδεση με τις νουκλεοτιδικές μεταλλάξεις T350G και A1053C.

	E1 γονίδιο	E6 γονίδιο	
Τραχηλικά δείγματα (n=50)	E1-1374 ^{63nt} διπλασιασμός / A1053C (E63D)	T350G (L83V)	Ενδοτυπικές παραλλαγές
	Αριθμοί καταχώρησης στη GenBank	Αριθμοί καταχώρησης στη GenBank	
Υψηλού βαθμού αλλοίωση (n=27)			
ATT4	JQ791090	JQ422080	European
ATT16	JQ791091	JQ422082	European
PNL4	JQ791086	JQ422083	European
PNL5	JQ791065	JX291526	European
PNL6	JQ791087	JX291528	European
Χαμηλού βαθμού αλλοίωση (n=23)			
A38	JQ791092	JN205112	European
858	JQ791094	JQ422101	European
793	JQ791095	JQ422103	European
868	JQ791097	JQ422106	European
PNL2	JQ791084	JQ422078	European
ATT6	JX291529	JQ422061	European
ATT21	JX291531	JQ422068	European

Πίνακας 3.12. Η κατανομή του E1-1374^{63nt} διπλασιασμού μεταξύ δειγμάτων υψηλού βαθμού αλλοίωσης (18.5%) και χαμηλού βαθμού αλλοίωσης (30.4%) και η συσχέτιση με τη μετάλλαξη T350G ($\chi^2 = 5.076$, $P < 0.024$).

3.1.6.1 Ανάλυση της λειτουργίας της E1 πρωτεΐνης

Λαμβάνοντας υπόψη ότι ο διπλασιασμός E1-1374^{63nt} προκαλεί την επανάληψη της περιοχής μεταξύ των αμινοξέων 150 έως 170 στο N τελικό άκρο της E1 πρωτεΐνης, πραγματοποιήθηκε εξελικτική μελέτη της περιοχής του διπλασιασμού μέσω του λόγου dN/dS, έτσι ώστε να προσδιοριστεί μέσω επιλεκτικής πίεσης εάν η λειτουργία και η δομή του N-τελικού άκρου της πρωτεΐνης επηρεάζεται λόγω του διπλασιασμού. Ο προσδιορισμός της επιλεκτικής πίεσης που ασκείται στο γονίδιο έδειξε ότι dN/dS=0.56. Επιπλέον ο μέσος ρυθμός υποκατάστασης του πρώτου, δεύτερου και τρίτου νουκλεοτιδίου του κωδικονίου έδειξε ότι CP1=0.918, CP2=0.722, CP3=1.36. Από τα αποτελέσματα αυτά αποδείχθηκε ότι στη περιοχή του διπλασιασμού ασκείται αρνητική επιλεκτική πίεση γεγονός που δηλώνει ότι η περιοχή του E1 γονιδίου είναι αρκετά συντηρημένη με αποτέλεσμα η λειτουργία και η δομή του N-τελικού άκρου της πρωτεΐνης να μην επηρεάζεται αρνητικά λόγω διπλασιασμού.

Σε προηγούμενη μελέτη στην οποία προβλέφθηκε η δομή της E1 πρωτεΐνης με το διπλασιασμό των 63 νουκλεοτιδίων μέσω τρισδιάστατης απεικόνισης, παρατηρήθηκε ότι ο E1-1374^{63nt} διπλασιασμός και η αμινοξική μετάλλαξη E63D,

τοποθετούνται στην επιφάνεια της πρωτεΐνης και προτάθηκε ότι δεν υπάρχουν σοβαρές αλλαγές στην δομή και λειτουργία της πρωτεΐνης λόγω του διπλασιασμού. Η υπόθεση αυτή σε συνδυασμό με το αποτέλεσμα της επιλεκτικής πίεσης ενισχύει το γεγονός ότι το N-τελικό άκρο της πρωτεΐνης δεν επηρεάζεται αρνητικά λόγω του διπλασιασμού των 63 νουκλεοτιδίων (Sabol et al., 2012).

3.2 Προσδιορισμός της μορφής του γονιδιώματος του HPV16

Η μορφή του γονιδιώματος του HPV16 πραγματοποιήθηκε με τη βοήθεια της Ποσοτικής PCR Πραγματικού Χρόνου και το προσδιορισμό του λόγου των τιμών E1/E6 και E2/E6. Η βασική αρχή της μεθόδου βασίζεται στην υπόθεση ότι τα γονίδια E1, E2 και E6 παρουσιάζουν ίση ποσότητα αντιγράφων όταν ο ιός βρίσκεται στην επισωματική του μορφή, ενώ όταν ο ιός ενσωματώνεται στο γονιδίωμα του κυττάρου ξενιστή, καταστρέφεται το ανοιχτό πλαίσιο ανάγνωσης των E1 ή/και E2 γονιδίων. Για να καθοριστούν οι θέσεις - στόχοι για τον υπολογισμό του αριθμού των αντιγράφων των γονιδίων E1 και E2, πραγματοποιήθηκε χαρτογράφηση των γονιδίων, με σκοπό να προσδιοριστούν οι κοινές θέσεις ρήξης του γονιδιώματος του HPV16.

3.2.1 Χαρτογράφηση των γονιδίων E1 και E2

Η νουκλεοτιδική ανάλυση των E1 και E2 γονιδίων που περιγράφηκε σε προηγούμενες ενότητες (ενότητα 3.1.3 και 3.1.5) έδειξε σε αρκετές περιπτώσεις φαινόμενα ρήξης στο ανοιχτό πλαίσιο ανάγνωσης των E1 και E2 γονιδίων. Ο προσδιορισμός των θέσεων ρήξης πραγματοποιήθηκε με PCR με τη χρήση εννέα ζευγαριών αλληλεπικαλυπτόμενων εκκινητικών μορίων που ενισχύουν σε επιμέρους τμήματα ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου και με τρία ζευγάρια αλληλεπικαλυπτόμενων εκκινητών που ενισχύουν σε επιμέρους τμήματα ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του E2 γονιδίου (Πίνακας 2.5). Για την διαδικασία αυτή χρησιμοποιήθηκαν συνολικά ογδόντα τέσσερα κλινικά δείγματα (δεκαεπτά τραχηλικοί ιστοί μονιμοποιημένοι σε παραφίνη, οκτώ δείγματα από τραχηλικά επιχρίσματα και πενήντα εννέα δείγματα ThinPrep). Από τα ογδόντα τέσσερα δείγματα που εξετάστηκαν τριάντα τρία δείγματα είχαν χαρακτηριστεί από τους κυτταροπαθολόγους ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και σαράντα δείγματα είχαν χαρακτηριστεί ως χαμηλού βαθμού

ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Επιπρόσθετα, έντεκα δείγματα είχαν χαρακτηριστεί ως καρκίνος του τραχήλου της μήτρας (9 περιπτώσεις τραχηλικού ιστού μονιμοποιημένου σε παραφίνη και δύο περιπτώσεις ThinPrep).

Στη παρούσα διατριβή χαρακτηρίστηκαν ως γεγονότα ρήξης οι περιπτώσεις των δειγμάτων που δεν έδωσαν ζώνη ενίσχυσης μετά από τη PCR, αλλά ενισχύθηκε το τμήμα του γονιδίου της β-ακτίνης και του E6 γονιδίου. Η ενίσχυση του E6 γονιδίου πραγματοποιήθηκε με σκοπό να εξεταστεί η ακεραιότητα του γονιδιώματος του HPV16, έτσι ώστε να προσδιορισθεί ότι η ρήξη εντός των γονιδίων E1 και E2 οφείλεται στην ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα και όχι από τον κατακερματισμό του γονιδιώματος του ιού. Η ευαισθησία της μεθόδου για την ενίσχυση των επιμέρους τμημάτων των E1, E2 και τμήματος του E6 γονιδίου, καθορίστηκε στα 10 αντίγραφα/αντίδραση για κάθε ζευγάρι εκκινητικών μορίων.

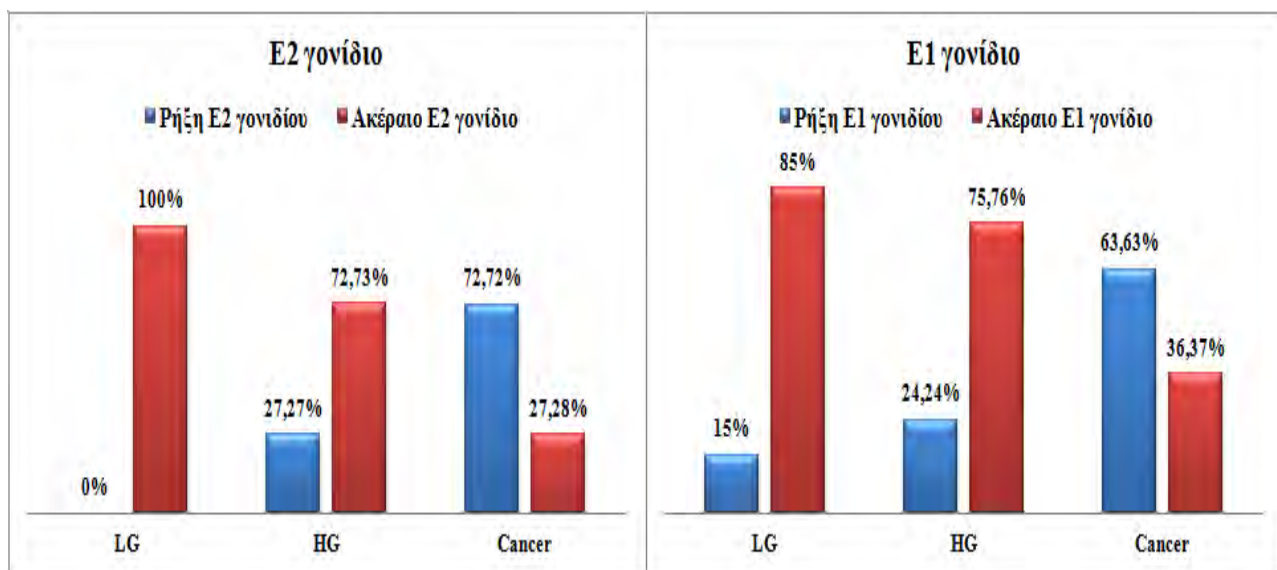
Η ενίσχυση των E1 και E2 γονιδίων πραγματοποιήθηκε παράλληλα σε όλα τα κλινικά δείγματα. Από την ενίσχυση του E2 γονιδίου ρήξη παρατηρήθηκε συνολικά σε δεκαεπτά κλινικά δείγματα (20.23%). Πιο συγκεκριμένα, εννέα περιπτώσεις (9/33) υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (27.27%), παρουσίασαν ρήξη εντός του E2 γονιδίου, ενώ είκοσι τέσσερις περιπτώσεις υψηλού βαθμού αλλοίωσης (72.72%) περιείχαν ακέραιο το ανοιχτό πλαίσιο ανάγνωσης του E2 γονιδίου (Εικόνα 3.7). Όσον αφορά τα δείγματα χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης δεν καταγράφηκε ρήξη εντός του E2 γονιδίου. Με βάση τα αποτελέσματα της παρούσας διατριβής παρατηρήθηκε ότι η ρήξη εντός του ανοιχτού πλαισίου ανάγνωσης του E2 γονιδίου συνδέεται σημαντικά με δείγματα υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, όταν συγκρίνεται με δείγματα χαμηλού βαθμού δυσπλασίας ($\chi^2 = 10.05$, $P < 0.01$). Επιπλέον, από τις έντεκα περιπτώσεις καρκίνου του τραχήλου της μήτρας παρατηρήθηκε ότι οκτώ δείγματα (72.72%) φέρουν ρήξη στο E2 γονίδιο και αναφέρθηκε σημαντική σχέση μεταξύ της ρήξης του γονιδίου E2 με τις περιπτώσεις καρκίνου, όταν συγκρίνεται με χαμηλού βαθμού δυσπλασίες ($\chi^2 = 22.2$, $P < 0.001$) και υψηλού βαθμού δυσπλασίες ($\chi^2 = 5.4$, $P < 0.05$) (Εικόνα 3.7).

Από την ενίσχυση του E2 γονιδίου παρατηρήθηκε ρήξη μεταξύ των νουκλεοτιδίων 2735 και 3210 σε ένα δείγμα υψηλού βαθμού αλλοίωσης (Εικόνα 3.8, Πίνακας 3.13). Επιπλέον ρήξη μεταξύ των νουκλεοτιδίων 3127 και 3649 αναφέρθηκε σε έξι δείγματα υψηλού βαθμού αλλοίωσης και σε επτά περιπτώσεις καρκίνου, ενώ ρήξη μεταξύ των νουκλεοτιδίων 3597 και 3873 ανιχνεύτηκε σε δύο δείγματα υψηλού βαθμού αλλοίωσης και σε μία περίπτωση καρκίνου (Εικόνα 3.8, Πίνακας 3.13). Με

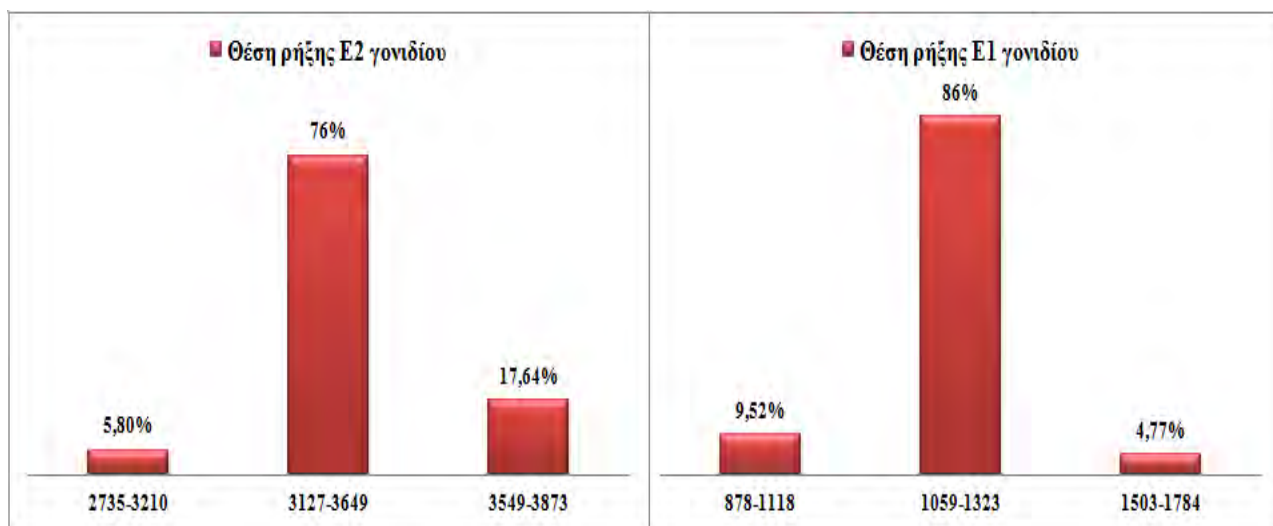
βάση τα αποτελέσματα παρατηρήθηκε ότι η περιοχή μεταξύ των νουκλεοτιδίων 3172 και 3649, η οποία κωδικοποιεί για την περιοχή "αρμό" της E2 πρωτεΐνης, αποτελεί τη πιο συχνή θέση ρήξης του E2 γονιδίου (Εικόνα 3.8, Πίνακας 3.13).

Η ενίσχυση του E1 γονιδίου πραγματοποιήθηκε παράλληλα στα ίδια κλινικά δείγματα. Ρήξη εντός του E1 γονιδίου παρατηρήθηκε σε είκοσι ένα δείγματα (25%). Η πιο συχνή θέση ρήξης του E1 γονιδίου παρατηρήθηκε στο 5' άκρο του γονιδίου μεταξύ των νουκλεοτιδίων 1059 και 1323 (Εικόνα 3.8, Πίνακας 3.13). Πιο συγκεκριμένα, οκτώ από τα τριάντα τρία δείγματα υψηλού βαθμού αλλοίωσης (24.24%) παρουσίασαν ρήξη εντός του ανοιχτού πλαισίου ανάγνωσης του E1 γονιδίου, ενώ ακέραιο το E1 γονίδιο παρατηρήθηκε σε είκοσι τρεις περιπτώσεις υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (75.76%) (Εικόνα 3.7). Επιπρόσθετα, τριάντα τέσσερα από τα σαράντα δείγματα χαμηλού βαθμού αλλοίωσης (85%), περιέχουν ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου, ενώ έξι δείγματα (15%) παρουσίασαν ρήξη στο E1 γονίδιο (Εικόνα 3.7). Επιπρόσθετα, από τις έντεκα περιπτώσεις καρκίνου παρατηρήθηκε ρήξη σε επτά δείγματα (63.63%) μεταξύ των νουκλεοτιδίων 1059 και 1323 (Εικόνα 3.8, Πίνακας 3.13). Η χαρτογράφηση των E1 και E2 γονιδίων στα ίδια κλινικά δείγματα, έδειξε ότι σε δείγματα χαμηλού βαθμού αλλοίωσης παρατηρείται ρήξη μόνο εντός του γονιδίου E1 και υπάρχει σημαντική σχέση μεταξύ της ρήξης του γονιδίου E1 με χαμηλού βαθμού αλλοιώσεις όταν αυτές συγκρίνονται με τις ρήξεις των γονιδίων E1 και E2 σε δείγματα χαμηλού και υψηλού βαθμού ενδοεπιθηλιακών αλλοιώσεων (Fisher's exact test, $P < 0.05$). Επιπλέον παρατηρήθηκε σύνδεση της ρήξης του E1 γονιδίου με τις περιπτώσεις καρκίνου όταν συγκρίνονται με χαμηλού βαθμού δυσπλασίες ($\chi^2 = 14.75$, $P < 0.001$) και υψηλού βαθμού δυσπλασίες ($\chi^2 = 8.44$, $P < 0.01$).

Τέλος, από τη χαρτογράφηση των θέσεων ρήξης του γονιδιώματος του ιού παρατηρήθηκε ταυτόχρονη ρήξη εντός των γονιδίων E1 και E2 στο 10.7% των εξεταζόμενων δειγμάτων. Πιο συγκεκριμένα, ρήξη και στα δύο γονίδια παρατηρήθηκε σε τέσσερα δείγματα υψηλού βαθμού αλλοίωσης (D2, D13, D14, D15) και σε τέσσερις περιπτώσεις καρκίνου (D58, D62, D64, D65, D66) (Πίνακας 3.13).



Εικόνα 3.7. Παρουσίαση των ποσοστών ρήξης των γονιδίων E1 και E2 σε τριάντα τρία δείγματα υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (HG), σαράντα δείγματα χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (LG) και έντεκα περιπτώσεις καρκίνου του τραχήλου της μήτρας (cancer). Σε χαμηλού βαθμού αλλοιώσεις είναι δυνατόν να παρατηρηθεί ρήξη μόνο του γονιδίου E1. Επιπλέον υπάρχει στατιστικά σημαντική σχέση μεταξύ της ρήξης των γονιδίων E1 και E2 με τις περιπτώσεις καρκίνου του τραχήλου της μήτρας όταν συγκρίνονται με την ρήξη των E1 και E2 γονιδίων σε χαμηλού βαθμού και υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις (χ^2 , $P < 0.05$).



Εικόνα 3.8. Παρουσίαση των θέσεων ρήξης των E1 και E2 γονιδίων. Στο σχεδιάγραμμα παρουσιάζονται τα ποσοστά των θέσεων ρήξης για τα γονίδια E1 και E2. Η πιο κοινή θέση ρήξης του E2 γονιδίου είναι η περιοχή μεταξύ των νουκλεοτιδίων 3127 και 3649, η οποία παρατηρείται στο 76% των δειγμάτων που φέρουν ρήξη στο E2 γονίδιο. Επιπλέον, η πιο κοινή θέση ρήξης του E1 γονιδίου είναι η περιοχή μεταξύ των νουκλεοτιδίων 1059 και 1323, η οποία απαντάται στο 86% των εξεταζόμενων δειγμάτων που φέρουν ρήξη στο E1 γονίδιο.

			E1 γονίδιο							E2 γονίδιο				
			EIF1-E1R1	EIF2-E1R2	EIF3-E1R3	EIF4-E1R4	EIF5-E1R5	EIF6-E1R6	EIF7-E1R7	EIF8-E1R8	EIF9-E1R9	a1-a2	b1-b2	c1- c2
			878-1118	1059-1323	1312-1523	1503-1784	1760-1991	1952-2186	2170-2485	2443-2743	2662-2848	2735-3210	3172-3649	3597-3873
Αλλοίωση	Πρόελευση	Θέση Δείγμα												
LG	paraffin	D7	-											
LG	ThinPrep	747		-										
LG	ThinPrep	ATT6				-								
LG	ThinPrep	ATT78		-										
LG	ThinPrep	ATT91		-										
LG	ThinPrep	ATT98		-										
CIN														
HG	paraffin	D2		-								-		
HG	paraffin	D4									-			
HG	paraffin	D13		-								-		
HG	paraffin	D14		-								-		
HG	paraffin	D15		-									-	
HG	paraffin	D30		-									-	
HG	ThinPrep	ATT9		-										
HG	ThinPrep	ATT21										-		
HG	ThinPrep	ATT24										-		
HG	ThinPrep	PNL9		-										
HG	ThinPrep	PNL1		-										
HG	ThinPrep	PNL12	-											
HG	ThinPrep	ATT79										-		
cancer	ThinPrep	PNL7										-		
cancer	paraffin	D58		-								-		
cancer	paraffin	D59												-
cancer	paraffin	D60		-										
cancer	paraffin	D61		-										
cancer	paraffin	D62		-								-		
cancer	paraffin	D63										-		
cancer	paraffin	D64		-								-		
cancer	paraffin	D65		-								-		
cancer	paraffin	D66		-								-		

Πίνακας 3.13. Στον πίνακα παρουσιάζονται οι θέσεις ρήξης των γονιδίων E1 και E2. Από την ανάλυση προέκυψε ότι στα δείγματα χαμηλού βαθμού αλλοίωσης παρατηρείται μόνο ρήξη του E1 γονιδίου (Fisher's exact test, $P < 0.05$). Επιπλέον η συνηθέστερη θέση ρήξης του E1 γονιδίου βρίσκεται μεταξύ των νουκλεοτιδίων 1059 και 1323, ενώ η συνηθέστερη θέση ρήξης του E2 γονιδίου βρίσκεται μεταξύ των νουκλεοτιδίων 3172 και 3649.

3.2.2 Προσδιορισμός της μορφής του γονιδιώματος του HPV16 με Ποσοτική PCR Πραγματικού Χρόνου (quantitative - Real Time PCR)

Η επιλογή των εκκινητικών μορίων για τον υπολογισμό των αντιγράφων των E1, E2 και E6 γονιδίων πραγματοποιήθηκε λαμβάνοντας υπόψη δύο κριτήρια, α) τις πιο συχνές θέσεις ρήξης των E1 και E2 γονιδίων και β) να υβρίζονται σε συντηρημένες περιοχές του γονιδιώματος του ιού. Πιο συγκεκριμένα, με βάση τα αποτελέσματα που προέκυψαν από τη χαρτογράφηση των E1 και E2 γονιδίων, παρατηρήθηκε ότι η συνηθέστερη θέση ρήξης του E1 γονιδίου είναι η περιοχή μεταξύ των νουκλεοτιδίων 1059 έως 1323 και για το E2 γονίδιο είναι η περιοχή μεταξύ των νουκλεοτιδίων 3172 έως 3649 (Εικόνα 3.8, Πίνακας 3.13). Επιπλέον, πραγματοποιήθηκε πολλαπλή ομοπαράθεση, μεταξύ των αλληλουχιών που μελετήθηκαν στη παρούσα διατριβή, της πρότυπης αλληλουχίας του HPV16 γονιδιώματος και τις πρότυπες αλληλουχίες των ενδοτυπικών παραλλαγών που υπάρχουν διαθέσιμες στη τράπεζα δεδομένων GenBank με αριθμό καταχώρησης: European: HQ644236, HQ644259, HQ644267, HQ644268, HQ644271, Asian: HQ644234, HQ644235, HQ644248, HQ644251, HQ644261, African type I: HQ644238, HQ644240, AF472508, HQ644290, African type II: HQ644237, HQ644239, HQ644249, HQ644250, AF472509, HQ644291, HQ644292, Asian American: HQ644289, HQ644247, HQ644253, HQ644255, HQ644254, AY686579, AY686582, HQ644273, HQ644257, για κάθε γονίδιο ξεχωριστά (E1, E2 και E6). Λαμβάνοντας υπόψη τα δύο αυτά κριτήρια ο υπολογισμός των αντιγράφων των E1 και E2 γονιδίων για τον προσδιορισμό του λόγου των τιμών E1/E6 και E2/E6 πραγματοποιήθηκε με την χρήση των εκκινητών E1F12-E1R12 (E2 γονίδιο, θέση 3243- 3539) και E1F2-E1R2 (E1 γονίδιο, θέση 1059-1323) (Πίνακας 3.13), ενώ για το γονίδιο E6 χρησιμοποιήθηκαν οι εκκινητές HPV16 1- HPV16 353 (Πίνακας 2.6).

Για τον προσδιορισμό της μορφής του γονιδιώματος του HPV16 χρησιμοποιήθηκαν συνολικά εβδομήντα έξι κλινικά δείγματα (δεκαεπτά τραχηλικοί ιστοί μονιμοποιημένοι σε παραφίνη και πενήντα εννέα δείγματα ThinPrep). Από τα εβδομήντα έξι δείγματα που εξετάστηκαν, είκοσι οκτώ δείγματα είχαν χαρακτηριστεί από τους κυτταροπαθολόγους ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου και τριάντα επτά δείγματα είχαν χαρακτηριστεί ως χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση πλακώδους επιθηλίου. Επιπρόσθετα, έντεκα δείγματα είχαν χαρακτηριστεί ως καρκίνος του τραχήλου της μήτρας (εννέα περιπτώσεις τραχηλικού ιστού μονιμοποιημένου σε παραφίνη και δύο περιπτώσεις ThinPrep).

3.2.2 1 Καθορισμός του εύρους των τιμών E2/E6 και E1/E6

Για να διευκρινιστούν τα όρια της επισωματικής, ενσωματωμένης και μεικτής μορφής του γονιδιώματος του ιού χρησιμοποιήθηκαν διαλύματα πλασμιδίων που περιέχουν 100 αντίγραφα του πλασμιδίου pE6-E2 (επισωματική μορφή του γονιδιώματος του ιού) και ποσότητες από 0 έως 1×10^5 αντίγραφα του πλασμιδίου pE6 (ενσωματωμένη μορφή του γονιδιώματος του ιού) (Πίνακας 2.7). Στη συνέχεια, πραγματοποιήθηκε προσδιορισμός των αντιγράφων των E1, E2 και E6 γονιδίων με τη μέθοδο της Ποσοτικής PCR Πραγματικού Χρόνου, χρησιμοποιώντας τις πρότυπες καμπύλες που αναλύθηκαν στην ενότητα 2.6.2. Από την ανάλυση του λόγου των τιμών E1/E6 βρέθηκε ότι τιμές κοντά στο μηδέν (< 0.003) υπολογίζονται, όταν το πλασμίδιο pE6 βρίσκεται σε εκατονταπλάσια περίσσεια του πλασμιδίου pE6-E2. Η τιμή E1/E6 = 0.03 παρατηρείται, όταν στο διάλυμα των πλασμιδίων το πλασμίδιο pE6 βρίσκεται σε δεκαπλάσια περίσσεια ενώ η τιμή E1/E6 = 0.52 παρατηρείται όταν το διάλυμα των πλασμιδίων περιείχε ίση ποσότητα των pE6-E2 και pE6 πλασμιδίων (100 αντίγραφα). Τέλος, όταν η συγκέντρωση του pE6-E2 είναι σε δεκαπλάσια περίσσεια στο διάλυμα τότε η τιμή E1/E6 = 0.7. Η τιμή E1/E6 = 0.98 παρατηρείται, όταν το διάλυμα των πλασμιδίων περιείχε μόνο το πλασμίδιο pE6-E2 (επισωματική μορφή του ιού) (Πίνακας 3.14).

Όσον αφορά το εύρος των τιμών E2/E6 παρατηρείται αντίστοιχα, ότι τιμές κοντά στο μηδέν (< 0.001) υπολογίζονται, όταν το πλασμίδιο pE6 βρίσκεται σε εκατονταπλάσια περίσσεια του πλασμιδίου pE2-E6. Όταν το πλασμίδιο pE6 βρίσκεται σε δεκαπλάσια περίσσεια η τιμή E2/E6 = 0.01, ενώ όταν η ποσότητα των δύο πλασμιδίων είναι ίση στο διάλυμα τότε E2/E6 = 0.62. Όταν η συγκέντρωση του pE6-E2 είναι σε δεκαπλάσια περίσσεια στο διάλυμα τότε η τιμή E2/E6 = 0.82. Τέλος E2/E6 = 0.96 όταν το διάλυμα των πλασμιδίων περιείχε μόνο το πλασμίδιο pE6-E2 (επισωματική μορφή του ιού) (Πίνακας 3.14).

Συμπερασματικά, η ενσωμάτωση του ιού συμβαίνει όταν ο λόγος E1/E6 < 0.003 . Σε τιμές μεταξύ 0.003 έως 0.98 ο ιός βρίσκεται και στις δύο μορφές, ενώ σε τιμές E1/E6 μεγαλύτερες ή ίσες με 0.98 ο ιός βρίσκεται υπό την επισωματική του μορφή (Πίνακας 3.14). Επίσης, όσον αφορά τις τιμές E2/E6 παρατηρείται ότι η ενσωμάτωση του ιού συμβαίνει σε τιμές μικρότερες από 0.001 ενώ μεταξύ 0.001 έως 0.96 παρουσιάζεται η μεικτή μορφή του γονιδιώματος του ιού, ενώ σε τιμές μεγαλύτερες ή ίσες με 0.96 παρατηρείται η επισωματική μορφή του γονιδιώματος του HPV16 (Πίνακας 3.14).

	Πλασμιδία		Λόγος τιμών		HPV16 DNA
	pE6-E2 (επίσωμα)	pE6 (ενσωμάτωση)	E1/E6	E2/E6	
Διάλυμα 1	100 αντίγραφα	0 αντίγραφα	0.98	0.96	επίσωμα
Διάλυμα 2	100 αντίγραφα	10 αντίγραφα	0.7	0.82	μείγμα
Διάλυμα 3	100 αντίγραφα	100 αντίγραφα	0.52	0.62	μείγμα
Διάλυμα 4	100 αντίγραφα	1000 αντίγραφα	0.03	0.01	μείγμα
Διάλυμα 5	100 αντίγραφα	10000 αντίγραφα	0.003	0.001	ενσωματωμένο
Διάλυμα 6	100 αντίγραφα	100000 αντίγραφα	0	0	ενσωματωμένο

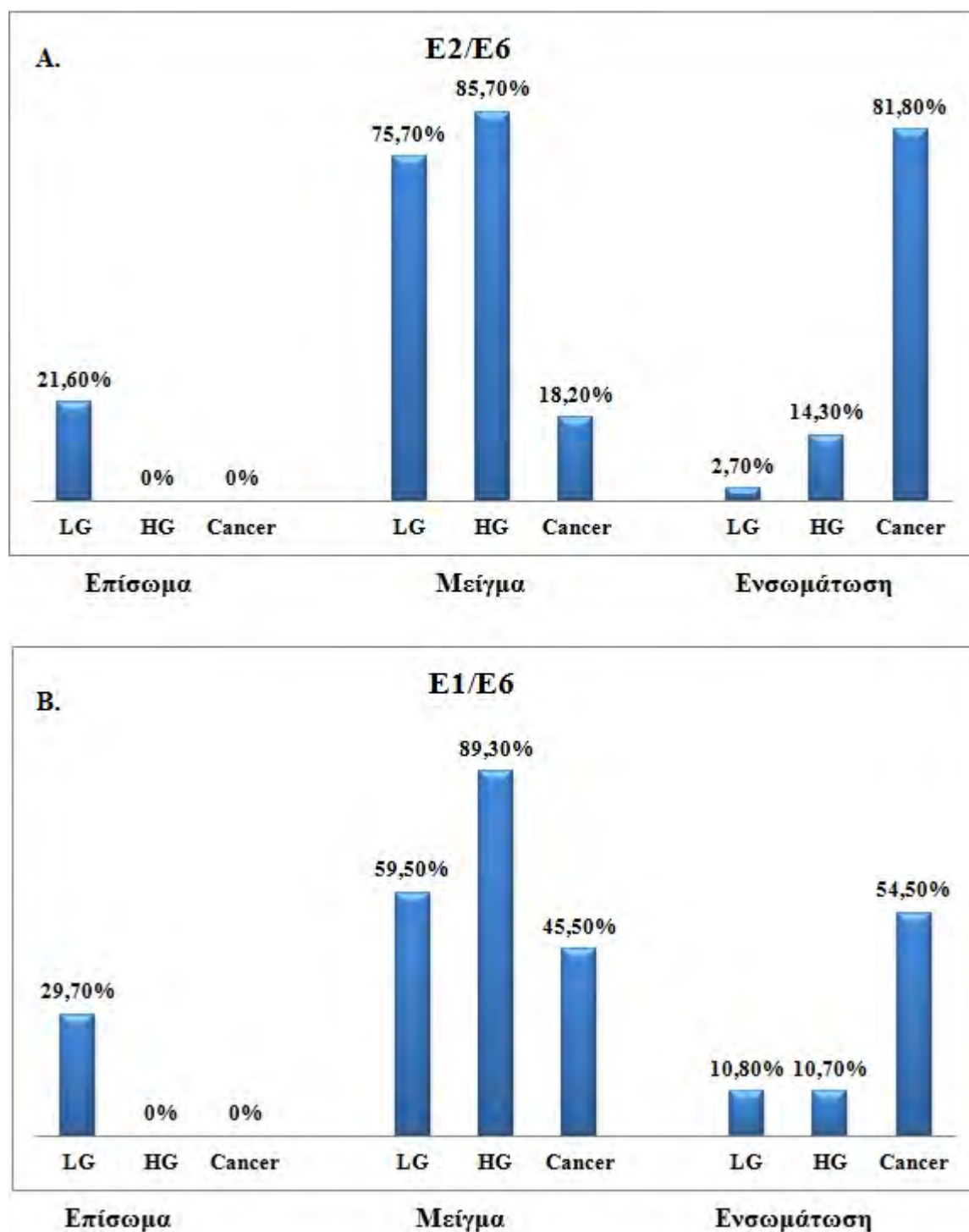
Πίνακας 3.14. Στον πίνακα αναγράφεται η σύσταση των διαλυμάτων των πλασμιδίων, και ο λόγος των τιμών E1/E6 και E2/E6, όπως υπολογίστηκε με τη μέθοδο Ποσοτικής PCR Πραγματικού Χρόνου. Επιπλέον με βάση τις τιμές καθορίστηκαν τα όρια που προσδιορίζουν την επισωματική, μεικτή και ενσωματωμένη μορφή του γονιδιώματος του HPV16.

3.2.2.2 Προσδιορισμός των τιμών E1/E6, E2/E6

Από την ανάλυση του λόγου E2/E6 σε όλα τα κλινικά δείγματα παρατηρήθηκε ότι η καθαρά επισωματική μορφή του ιού εντοπίστηκε σε οκτώ από τα τριάντα επτά δείγματα χαμηλού βαθμού αλλοίωσης (21.6%) και σε κανένα δείγμα υψηλού βαθμού αλλοίωσης και καρκίνου (Εικόνα 3.9 Α). Η μεικτή μορφή του HPV16 DNA εντοπίστηκε σε είκοσι τέσσερα από τα είκοσι οκτώ δείγματα υψηλού βαθμού αλλοίωσης (85.7%) και σε είκοσι οκτώ από τα τριάντα επτά δείγματα χαμηλού βαθμού αλλοίωσης (75.7%). Όσον αφορά τις περιπτώσεις καρκίνου η μεικτή μορφή του HPV16 DNA εντοπίστηκε σε δύο από τα έντεκα δείγματα καρκίνου (18.2%) (Πίνακας 3.15, 3.16). Τέλος, η καθαρά ενσωματωμένη μορφή του ιού ανιχνεύτηκε σε ένα δείγμα χαμηλού βαθμού αλλοίωσης (2.7%), σε τέσσερα δείγματα υψηλού βαθμού αλλοίωσης (14.3%) και σε εννέα δείγματα καρκίνου του τραχήλου της μήτρας (81.8%) (Πίνακας 3.15, 3.16). Πιο συγκεκριμένα, η καθαρά επισωματική μορφή του ιού βρίσκεται μόνο σε χαμηλού βαθμού αλλοιώσεις (21.6%). Παρόλα αυτά, στις περιπτώσεις χαμηλού βαθμού δυσπλασίας επικρατεί η μεικτή μορφή του HPV16 DNA, σε σύγκριση με την επισωματική (21.6%) και την ενσωματωμένη μορφή (2.7%) του ιού ($\chi^2 = 9.757$ $P = 0.0018$). Επίσης, στις περιπτώσεις υψηλού βαθμού αλλοίωσης παρατηρήθηκε σημαντική επικράτηση της μεικτής μορφής του HPV16 DNA (85.7%), σε σύγκριση με την ενσωματωμένη και την επισωματική μορφή του ιού ($\chi^2 = 14.286$, $P = 0.0002$). Τέλος, στις περιπτώσεις καρκίνου επικρατεί η ενσωματωμένη μορφή του ιού όταν συγκρίνεται με την επισωματική και τη μεικτή μορφή του HPV16 DNA ($\chi^2 = 4.5$, $P < 0.05$) (Εικόνα 3.9 Α).

Όσον αφορά τον λόγο των τιμών E1/E6 υπολογίστηκε ότι η καθαρά επισωματική μορφή του γονιδιώματος του ιού εντοπίστηκε σε έντεκα από τα τριάντα επτά δείγματα χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (29.7%) και σε κανένα δείγμα υψηλού βαθμού

αλλοίωσης και καρκίνου. Η μεικτή μορφή του γονιδιώματος του HPV16 εντοπίστηκε σε είκοσι δύο από τα τριάντα επτά δείγματα χαμηλού βαθμού αλλοίωσης (59.5%), σε είκοσι πέντε από τα είκοσι οκτώ δείγματα υψηλού βαθμού αλλοίωσης (89,3%) και σε πέντε από τα 11 δείγματα καρκίνου του τραχήλου της μήτρας (45.5%) (Πίνακας 3.15, 3.16). Τέλος η καθαρά ενσωματωμένη μορφή του γονιδιώματος του HPV16 εντοπίστηκε σε τρία δείγματα υψηλού βαθμού αλλοίωσης (10.7%) και σε έξι δείγματα καρκίνου του τραχήλου της μήτρας (54.5%). Επιπλέον, η καθαρά επισωματική μορφή του γονιδιώματος του ιού εντοπίστηκε σε τέσσερα δείγματα χαμηλού βαθμού αλλοίωσης (10.8%) (Πίνακας 3.15, 3.16). Πιο συγκεκριμένα, η καθαρά επισωματική μορφή του γονιδιώματος του ιού εντοπίζεται με μεγαλύτερη συχνότητα σε χαμηλού βαθμού αλλοιώσεις (29.7%), ενώ δεν παρατηρήθηκε στατιστικά σημαντική σύνδεση της μεικτής μορφής του HPV16 DNA με χαμηλού βαθμού αλλοιώσεις. Επιπλέον, στα δείγματα υψηλού βαθμού αλλοίωσης επικρατεί η μεικτή μορφή του γονιδιώματος του HPV16 (89.3%), σε σύγκριση με την επισωματική και ενσωματωμένη μορφή του γονιδιώματος του ιού (10.7%) ($\chi^2 = 17.286$, $P < 0.0001$). Επίσης, η ενσωματωμένη μορφή του γονιδιώματος του ιού βρίσκεται με μεγαλύτερη συχνότητα στις περιπτώσεις καρκίνου (54.5%) (Εικόνα 3.9 B).



Εικόνα 3.9. Α. Κατανομή της μορφής του γονιδιώματος του HPV16, μεταξύ χαμηλού και υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης και καρκίνου του τραχήλου της μήτρας, μέσω υπολογισμού του λόγου των τιμών E2/E6. Β. Κατανομή της μορφής του γονιδιώματος του HPV16 μεταξύ δειγμάτων χαμηλού και υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης και καρκίνου, μέσω προσδιορισμού του λόγου E1/E6.

Δείγμα	Βαθμός αλλοίωσης	E2	E1	E6	E2/E6	Μορφή (E2/E6)	E1/E6	Μορφή (E1/E6)
PNL2	Χαμηλού βαθμού αλλοίωση	800	721	820	0,98	επίσωμα	0,88	μείγμα
PNL3	Χαμηλού βαθμού αλλοίωση	110	240	240	0,46	μείγμα	1	επίσωμα
ATT8	Χαμηλού βαθμού αλλοίωση	60	260	265	0,23	μείγμα	0,98	επίσωμα
ATT10	Χαμηλού βαθμού αλλοίωση	13	12	13	1	επίσωμα	0,98	επίσωμα
ATT18	Χαμηλού βαθμού αλλοίωση	12	172	175	0,07	μείγμα	0,98	επίσωμα
ATT23	Χαμηλού βαθμού αλλοίωση	1944	1960	1970	0,98	επίσωμα	0,99	επίσωμα
ATT26	Χαμηλού βαθμού αλλοίωση	4390	4320	4400	1	επίσωμα	0,98	επίσωμα
858	Χαμηλού βαθμού αλλοίωση	1565	1580	1600	0,98	επίσωμα	0,99	επίσωμα
2279	Χαμηλού βαθμού αλλοίωση	1350	1365	1380	0,98	επίσωμα	0,99	επίσωμα
2290	Χαμηλού βαθμού αλλοίωση	110500	113650	114300	0,97	επίσωμα	0,99	επίσωμα
ATT12	Χαμηλού βαθμού αλλοίωση	150	4400	4500	0,03	μείγμα	0,98	επίσωμα
ATT78	Χαμηλού βαθμού αλλοίωση	80	0	110	0,72	μείγμα	0	ενσωμάτωση
ATT91	Χαμηλού βαθμού αλλοίωση	100	0	200	0,5	μείγμα	0	ενσωμάτωση
ATT46	Χαμηλού βαθμού αλλοίωση	11	27	27	0,4	μείγμα	1	επίσωμα
ATT54	Χαμηλού βαθμού αλλοίωση	208	163	214	0,97	επίσωμα	0,76	μείγμα

Πίνακας 3.15. Παρουσίαση των δειγμάτων που φέρουν την επισωματική μορφή του γονιδιώματος του ιού μέσω του προσδιορισμού του λόγου των τιμών E1/E6 και E2/E6

Δείγμα	Βαθμός αλλοίωσης	E2	E1	E6	E2/E6	Μορφή (E2/E6)	E1/E6	Μορφή (E1/E6)
747	Χαμηλού βαθμού αλλοίωση	12	7	86475	0	ενσωμάτωση	0	ενσωμάτωση
D7	Χαμηλού βαθμού αλλοίωση	24	0	154	0,15	μείγμα	0	ενσωμάτωση
PNL1	Υψηλού βαθμού αλλοίωση	22	5	42	0,52	μείγμα	0,1	μείγμα
PNL9	Υψηλού βαθμού αλλοίωση	15	7	80	0,18	μείγμα	0,08	μείγμα
ATT21	Υψηλού βαθμού αλλοίωση	8	7362	52000	0	ενσωμάτωση	0,14	μείγμα
D13	Υψηλού βαθμού αλλοίωση	0	0	83	0	ενσωμάτωση	0	ενσωμάτωση
D14	Υψηλού βαθμού αλλοίωση	0	0	40	0	ενσωμάτωση	0	ενσωμάτωση
D15	Υψηλού βαθμού αλλοίωση	0	0	54	0	ενσωμάτωση	0	ενσωμάτωση
PNL7	Καρκίνος	5	13820	214440	0	ενσωμάτωση	0,064	μείγμα
PNL16	Καρκίνος	84450	276200	950000	0,08	μείγμα	0,3	μείγμα
D58	Καρκίνος	0	0	370	0	ενσωμάτωση	0	ενσωμάτωση
D59	Καρκίνος	15	10	16200	0	ενσωμάτωση	0	ενσωμάτωση
D60	Καρκίνος	12	7	15700	0	ενσωμάτωση	0	ενσωμάτωση
D61	Καρκίνος	70	6	1600	0,043	μείγμα	0,004	μείγμα
D62	Καρκίνος	0	7	120	0	ενσωμάτωση	0,05	μείγμα
D63	Καρκίνος	0	12	1000	0	ενσωμάτωση	0,012	μείγμα
D64	Καρκίνος	0	1	470	0	ενσωμάτωση	0,002	ενσωμάτωση
D65	Καρκίνος	0	0	1250	0	ενσωμάτωση	0	ενσωμάτωση
D66	Καρκίνος	0	2	1300	0	ενσωμάτωση	0,001	ενσωμάτωση

Πίνακας 3.16. Παρουσίαση των δειγμάτων που φέρουν την ενσωματωμένη μορφή του γονιδιώματος του ιού μέσω του προσδιορισμού του λόγου των τιμών E1/E6 και E2/E6.

3.2.2.3 Συγκριτική μελέτη των λόγων E1/E6, E2/E6

Με τον προσδιορισμό του λόγου των τιμών E1/E6 και E2/E6 στα ίδια δείγματα παρατηρήθηκε ότι είκοσι τέσσερα δείγματα υψηλού βαθμού αλλοίωσης, είκοσι δείγματα χαμηλού βαθμού αλλοίωσης και δύο περιπτώσεις καρκίνου του τραχήλου της μήτρας φέρουν τη μεικτή μορφή του γονιδιώματος του ιού. Αυτό οδήγησε στο συμπέρασμα ότι συνυπάρχουν τρεις διαφορετικοί "πληθυσμοί" HPV16 στα συγκεκριμένα δείγματα. Πιο συγκεκριμένα, ένας πληθυσμός που φέρει την επισωματική μορφή του ιού, ένας πληθυσμός που έχει υποστεί ενσωμάτωση μέσω ρήξης στο E1 γονίδιο και ένας πληθυσμός που έχει υποστεί ρήξη στο E2 γονίδιο, αντίστοιχα. Επιπλέον σε έξι δείγματα χαμηλού βαθμού αλλοίωσης (ATT10, ATT23, ATT26, 858, 2279, 2290) παρατηρήθηκε ότι το γονιδίωμα του ιού βρίσκεται στη καθαρά επισωματική μορφή, μέσω υπολογισμού του λόγου των τιμών E1/E6 και E2/E6 (Πίνακας 3.15, 3.16). Ακόμη, σε ένα δείγμα χαμηλού βαθμού αλλοίωσης (747), τρία δείγματα υψηλού βαθμού αλλοίωσης (D13, D14, D15) και σε έξι δείγματα καρκίνου (D58, D59, D60, D64, D65, D66) παρατηρήθηκε ότι οι δύο λόγοι E1/E6 και E2/E6 δείχνουν ότι το γονιδίωμα του ιού είναι ενσωματωμένο στο DNA του κυττάρου ξενιστή (Πίνακας 3.15, 3.16). Αυτό ίσως να οφείλεται στο γεγονός ότι η ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα προκαλεί την έλλειψη ολόκληρης της περιοχής μεταξύ των γονιδίων E1 και E2.

Επιπλέον, σε πέντε περιπτώσεις χαμηλού βαθμού αλλοίωσης (PNL3, ATT8, ATT18, ATT12, ATT46) ο λόγος E2/E6 δείχνει ότι το γονιδίωμα του ιού βρίσκεται σε μεικτή μορφή, ενώ σε ο λόγος E1/E6 δείχνει ότι ο ιός βρίσκεται σε μορφή επισώματος (Πίνακας 3.15, 3.16). Αυτό ίσως να οφείλεται στο γεγονός ότι στα δείγματα αυτά η ενσωμάτωση του HPV16 γονιδιώματος έχει γίνει αποκλειστικά μόνο μέσω ρήξης του E2 γονιδίου. Αντιθέτως, σε δύο περιπτώσεις χαμηλού βαθμού αλλοίωσης (PNL2, ATT54) ο λόγος E2/E6 δείχνει, ότι το γονιδίωμα του ιού είναι σε επισωματική μορφή, ενώ ο λόγος E1/E6 δείχνει ότι ο ιός βρίσκεται στη μεικτή μορφή του (Πίνακας 3.15, 3.16). Το γεγονός αυτό ίσως να οφείλεται στην ενσωμάτωση του γονιδιώματος του HPV16 αποκλειστικά μέσω ρήξης του E1 γονιδίου.

Ακόμη σε τρεις περιπτώσεις καρκίνου του τραχήλου της μήτρας (PNL7, D62, D63) και μία περίπτωση υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (ATT21) παρατηρείται, ότι ο λόγος E2/E6 δείχνει ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα, ενώ στα ίδια δείγματα ο λόγος E1/E6 προβλέπει την ύπαρξη της μεικτής μορφής του γονιδιώματος του ιού. Αντιθέτως, σε τρεις περιπτώσεις χαμηλού βαθμού αλλοίωσης (D7, ATT78, ATT91) ο λόγος E1/E6 προβλέπει ότι ο ιός είναι ενσωματωμένος, ενώ ο λόγος E2/E6 δείχνει ότι ο ιός βρίσκεται σε μεικτή μορφή (Πίνακας 3.15, 3.16). Οι διαφορές αυτές που υπάρχουν στα επτά

δείγματα, ίσως να οφείλονται σε διαφορετικά πρότυπα ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα, σε αναδιάταξη του γονιδιώματος του ιού ή σε συσσώρευση μεταλλάξεων στη περιοχή των γονιδίων E1 και E2 των HPV16 στελεχών που κυκλοφορούν στον Ελληνικό πληθυσμό, με αποτέλεσμα να προκύπτουν διαφορετικά δεδομένα με τους δύο λόγους των τιμών E1/E6 και E2/E6. Παρόλα αυτά, παρατηρείται ότι ο συνδυασμός των αποτελεσμάτων των λόγων των τιμών E1/E6 και E2/E6 μπορεί να δώσει σημαντικές πληροφορίες για τη μορφή του γονιδιώματος του HPV16 σε κλινικά δείγματα, καθώς επίσης και να δώσει πληροφορίες για τον τρόπο που ο ιός έχει ενσωματωθεί στο κυτταρικό χρωμόσωμα.

3.3 Προσδιορισμός θέσεων ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα

Στο τελευταίο μέρος της διατριβής πραγματοποιήθηκε ο προσδιορισμός των θέσεων ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα με τη βοήθεια της μεθόδου RS - PCR (Restriction Site PCR). Η συγκεκριμένη μέθοδος βασίζεται στη PCR και τη χρήση ειδικών εκκινητικών μορίων που υβριδίζονται σε συγκεκριμένες θέσεις του γονιδιώματος του ιού και συνδυάζονται στην αντίδραση με εκκινητές που αποκαλούνται εκκινητές αναγνώρισης θέσεων περιορισμού (Restriction site oligonucleotide primers, RSO) (Εικόνα 2.2). Σύμφωνα με την ανάλυση των αποτελεσμάτων αναμένονται τρεις διαφορετικές κατηγορίες αλληλουχιών. α) αλληλουχίες οι οποίες περιλαμβάνουν μόνο κυτταρικό DNA, β) αλληλουχίες που περιλαμβάνουν μόνο ιικό DNA και γ) αλληλουχίες που περιλαμβάνουν τη σύνδεση ιικού και κυτταρικού DNA, όπως έχει προκύψει μετά από την ενσωμάτωση του ιού στο κυτταρικό χρωμόσωμα.

Από την ανάλυση των αλληλουχιών που περιλαμβάνουν μόνο τμήματα του γονιδιώματος του HPV16, εντοπίστηκαν σε τρεις περιπτώσεις καρκίνου του τραχήλου της μήτρας, μια περίπτωση υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης και σε τέσσερις περιπτώσεις χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, φαινόμενα αναδιάταξης του γονιδιώματος του ιού με αναστροφή. Πιο συγκεκριμένα, στο δείγμα καρκίνου PNL7 με τον συνδυασμό των εκκινητών HPV16-2964-25D (υβριδίζεται στη θέση 2964 του HPV16 DNA E2 γονίδιο) και RSO-Bam (υβριδίζεται σε θέση αναγνώρισης του ενζύμου περιορισμού BamHI), εντοπίστηκε μια ενιαία αλληλουχία, η οποία περιλαμβάνει τμήμα του γονιδίου E2 και τμήμα της αλληλουχίας του L1 γονιδίου με αντίθετο προσανατολισμό. Πιο συγκεκριμένα, η θέση 1 έως τη θέση 373 αντιστοιχεί στη θέση 2965 έως 3337 του E2 γονιδίου με προσανατολισμό 5' - 3', ενώ στη συνέχεια η αλληλουχία από τη θέση 372 έως τη θέση 659 αντιστοιχεί στη περιοχή μεταξύ των νουκλεοτιδίων 6418 έως 6151 του L1

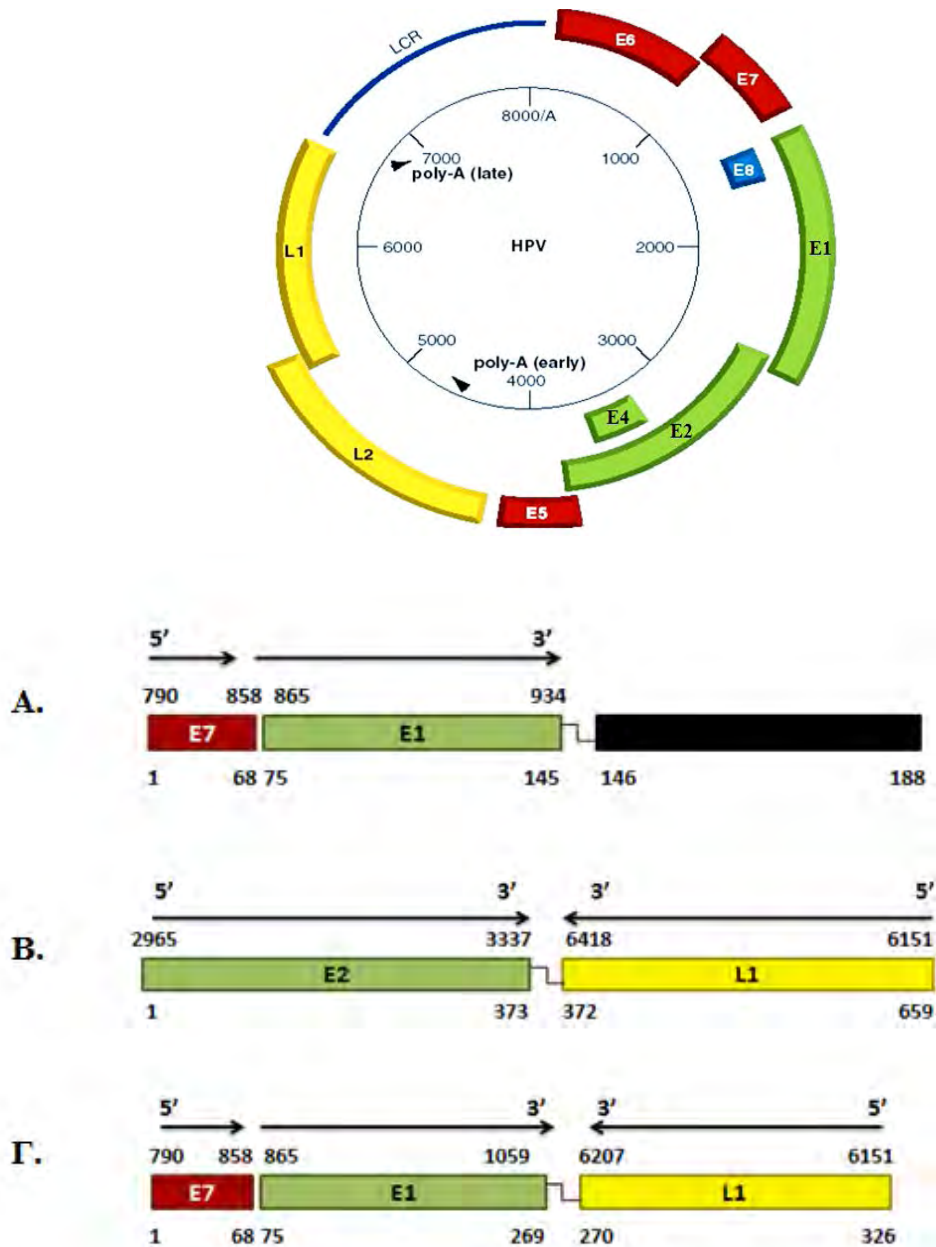
γονιδίου με αντίθετο προσανατολισμό (3' - 5') (Εικόνα 3.10 Α). Όπως φαίνεται στην εικόνα 3.10 Α η περιοχή μεταξύ των νουκλεοτιδίων 3337 έως 6418, η οποία περιλαμβάνει τη συνέχεια του E2 γονιδίου, το E4, E5, L2 και τμήμα του L1 γονιδίου έχει απαλειφθεί και συνεχίζει τμήμα του L1 γονιδίου με αντίθετο προσανατολισμό (3' - 5').

Επιπλέον σε δύο περιπτώσεις καρκίνου του τραχήλου της μήτρας (D59, PNL16), σε μία περίπτωση υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (ATT21) και σε τέσσερις περιπτώσεις χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης (ATT6, ATT78, ATT91 και ATT98) με τον συνδυασμό των εκκινητών HPV16-790-25D (υβριδίζεται στη θέση 790 του HPV16 DNA - E7 γονίδιο) και RSO-Bam (υβριδίζεται σε θέση αναγνώρισης του ενζύμου περιορισμού BamHI), εντοπίστηκε ένα διαφορετικό πρότυπο αναδιάταξης του γονιδιώματος του ιού με αναστροφή. Πιο συγκεκριμένα, εντοπίστηκε μια αλληλουχία ή οποία περιλαμβάνει τμήμα του E7 γονιδίου, τμήμα του E1 γονιδίου και τμήμα του L1 γονιδίου, με αντίθετο προσανατολισμό (3' - 5') (Εικόνα 3.10 Β). Η αλληλουχία μεταξύ των θέσεων 1 έως 269, περιλαμβάνει τη περιοχή μεταξύ των νουκλεοτιδίων 790 έως 858, που αντιστοιχεί στο τέλος του E7 γονιδίου και τη περιοχή των νουκλεοτιδίων 865 έως 1059 που περιλαμβάνει την αρχή του E1 γονιδίου με προσανατολισμό 5' - 3'. Στη συνέχεια μεταξύ των θέσεων 270 έως 326 ακολουθεί η περιοχή του L1 γονιδίου μεταξύ των νουκλεοτιδίων 6207 έως 6151 με αντίθετο προσανατολισμό (3' - 5'). Όπως φαίνεται στην εικόνα 3.10 η περιοχή μεταξύ των νουκλεοτιδίων 1059 έως 6207, η οποία περιλαμβάνει τη συνέχεια του E1 γονιδίου, και ολόκληρες τις περιοχές των γονιδίων E2, E4, E5, L2 και τμήματος του L1 γονιδίου έχει απαλειφθεί και συνεχίζει τμήμα του L1 γονιδίου με αντίθετο προσανατολισμό (3' - 5').

Για να επιβεβαιωθεί η ύπαρξη της αναδιάταξης του γονιδιώματος του HPV16, σχεδιάστηκαν εκκινητικά μόρια, τα οποία στοχεύουν σε θέσεις εκατέρωθεν του σημείου σύνδεσης των δύο ανεστραμμένων τμημάτων του HPV16 DNA (E7, E1-L1 και E2-L2). Οι εκκινητές σχεδιάστηκαν με βάση την αλληλουχία της αναδιάταξης για κάθε δείγμα ξεχωριστά. Επομένως, για την επιβεβαίωση της αναδιάταξης E7, E1-L1, το ζεύγος των εκκινητών RSE1F1-RSE1R1 εφαρμόστηκε στα δείγματα ATT6, ATT21, ATT78, ATT91 ATT98, ενώ το ζεύγος RSE1F1-RSE1R2 εφαρμόστηκε στα δείγματα Δ59, PNL16. Επιπλέον, για την αναδιάταξη E2-L1 σχεδιάστηκε και εφαρμόστηκε το ζευγάρι εκκινητών RSE2F1- RSE2R1 για τη περίπτωση καρκίνου PNL7. Από την ανάλυση των αποτελεσμάτων επιβεβαιώθηκε σε όλα τα δείγματα το φαινόμενο της αναδιάταξης του DNA του ιού. Επιπλέον, πρόσφατη μελέτη πρότεινε την ύπαρξη αντίστοιχων φαινομένων αναδιάταξης σε ενσωματωμένα μόρια HPV16 DNA που προέρχονται από τη καρκινική κυτταρική σειρά

Caski (Akagi et al., 2014). Ο ρόλος των νέων αλληλουχιών στην παθογένεια του ιού και ο τρόπος με τον οποίο έχουν σχηματιστεί συνεχίζεται να διερευνάται.

Τέλος, στη περίπτωση καρκίνου του τραχήλου της μήτρας D59 εκτός από την αλληλουχία αναδιάταξης του ιικού DNA εντοπίστηκε με το συνδυασμό των εκκινητών HPV16-790-25D (υβριδίζεται στη θέση 790 του HPV16 DNA - E7 γονίδιο) και RSO-Bam (υβριδίζεται σε θέση αναγνώρισης του ενζύμου περιορισμού BamHI), μια αλληλουχία που περιλαμβάνει τον συνδυασμό ιικού DNA και άγνωστης μέχρι στιγμής αλληλουχίας (Εικόνα 3.10 Γ). Πιο συγκεκριμένα, η αλληλουχία του δείγματος D59, περιλαμβάνει μεταξύ των θέσεων 1 έως 145 τη περιοχή μεταξύ των νουκλεοτιδίων 790 έως 858, που αντιστοιχεί στο τέλος του E7 γονιδίου και τη περιοχή των νουκλεοτιδίων 865 έως 934 που αντιστοιχεί στο 5' άκρο του E1 γονιδίου με προσανατολισμό 5' - 3'. Επιπλέον, από τη θέση 146 έως 188 υπάρχει νουκλεοτιδική αλληλουχία, η οποία δεν αντιστοιχεί σε αλληλουχία του ιού, ενώ δε μπορεί να διευκρινιστεί ακριβώς η ταυτότητα της συγκεκριμένης αλληλουχίας με τη βοήθεια του προγράμματος NCBI BLAST (Εικόνα 3.10 Γ). Για το σκοπό αυτό σχεδιάστηκε ένας εκκινητής, ο οποίος υβριδίζεται ακριβώς στη θέση σύνδεσης της ιικής και άγνωστης αλληλουχίας, έτσι ώστε να συνδυαστεί με τους εκκινητές αναγνώρισης θέσεων περιορισμού (Restriction site oligonucleotide primers, RSO). Η διαδικασία αυτή βρίσκεται σε εξέλιξη με στόχο, να ενισχυθεί μεγαλύτερο τμήμα της αλληλουχίας και να ταυτοποιηθεί με ακρίβεια η θέση της άγνωστης αλληλουχίας στο κυτταρικό χρωμόσωμα.



Εικόνα 3.10. Α) Αναδιάταξη του γονιδιώματος του HPV16 που περιλαμβάνει από τη θέση 1 έως τη θέση 373 τη περιοχή του E2 γονιδίου (πράσινο) μεταξύ των νουκλεοτιδίων 2965 έως 3337, και από τη θέση 372 έως 659 περιλαμβάνει την περιοχή του L1 γονιδίου (κίτρινο) μεταξύ των νουκλεοτιδίων 6418 έως 6151 με αντίθετο προσανατολισμό (3' - 5'). Β) Αναδιάταξη του γονιδιώματος του HPV16 που περιλαμβάνει από τη θέση 1 έως 269 τη περιοχή του E7 γονιδίου (κόκκινο) μεταξύ των νουκλεοτιδίων 790 έως 858, και το 5' άκρο του E1 γονιδίου (πράσινο) μεταξύ των νουκλεοτιδίων 865 έως 1059 και από τη θέση 270 έως 326 ακολουθεί η περιοχή του L1 γονιδίου μεταξύ των νουκλεοτιδίων 6207 έως 6151 με αντίθετο προσανατολισμό (3' - 5'). Γ) Στο δείγμα Δ59 εντοπίστηκε η σύνδεση ιικού DNA με άγνωστη αλληλουχία. Μεταξύ των θέσεων 1 έως 145 περιέχεται το τέλος του E7 γονιδίου (790 - 858) (κόκκινο) και το 5' άκρο του E1 γονιδίου (865 - 934) (πράσινο), ενώ από τη θέση 146 έως 188 υπάρχει άγνωστη νουκλεοτιδική αλληλουχία (μαύρο).

Επιπλέον, πραγματοποιήθηκε συγκριτική μελέτη μεταξύ του λόγου των τιμών E1/E6, E2/E6 και των περιπτώσεων αναδιάταξης. Πιο συγκεκριμένα, στα δείγματα που παρουσιάζουν το ίδιο μοτίβο αναδιάταξης του γονιδιώματος του HPV16, μεταξύ της περιοχής E7, E1-L1 παρατηρήθηκε, ότι στις περιπτώσεις PNL16 (καρκίνος), ATT6 και ATT98 (χαμηλού βαθμού αλλοίωση) επικρατεί μεικτή μορφή του γονιδιώματος του ιού, ενώ το δείγμα D59 (καρκίνος) φέρει μόνο την ενσωματωμένη μορφή του γονιδιώματος του HPV16 (Πίνακας 3.17). Αντιθέτως, στο δείγμα ATT21 (υψηλού βαθμού αλλοίωση) ο λόγος E2/E6 δείχνει ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα, ενώ ο λόγος E1/E6 προβλέπει την ύπαρξη της μεικτής μορφής του γονιδιώματος του ιού. Επιπλέον, στα δείγματα ATT78, ATT91 (χαμηλού βαθμού αλλοίωση) ο λόγος E1/E6 προβλέπει ότι ο ιός είναι ενσωματωμένος, ενώ ο λόγος E2/E6 δείχνει ότι ο ιός βρίσκεται σε μεικτή μορφή (Πίνακας 3.17). Τέλος, στο δείγμα PNL7 (καρκίνος) το οποίο παρουσιάζει διαφορετικό μοτίβο αναδιάταξης του HPV16 DNA (E2-L1), ο λόγος E2/E6 δείχνει ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα, ενώ ο λόγος E1/E6 προβλέπει την ύπαρξη της μεικτής μορφής του γονιδιώματος του ιού (Πίνακας 3.17). Τα αποτελέσματα αυτά ενισχύουν την υπόθεση ότι η ασυμφωνία των αποτελεσμάτων των λόγων E1/E6 και E2/E6 σε ορισμένες περιπτώσεις (ATT21, ATT78, ATT91, PNL7), ίσως να οφείλεται σε φαινόμενα αναδιάταξης του HPV16 DNA. Για τον λόγο αυτό πειράματα βρίσκονται ακόμα σε εξέλιξη, έτσι ώστε να διευκρινιστούν τα αίτια ασυμφωνίας των αποτελεσμάτων και στις υπόλοιπες περιπτώσεις.

Δείγμα	Βαθμός ενδοεπιθηλιακής αλλοίωσης	E2/E6	Μορφή HPV16 DNA	E1/E6	Μορφή HPV16 DNA	Αναδιάταξη HPV16 DNA
ATT6	Χαμηλού βαθμού αλλοίωση	0,33	μείγμα	0,57	μείγμα	E7, E1-L1, (790-1059, 6207-6151)
ATT78	Χαμηλού βαθμού αλλοίωση	0,72	μείγμα	0	ενσωμάτωση	E7, E1-L1, (790-1059, 6207-6151)
ATT98	Χαμηλού βαθμού αλλοίωση	0,45	μείγμα	0,07	μείγμα	E7, E1-L1, (790-1059, 6207-6151)
ATT91	Χαμηλού βαθμού αλλοίωση	0,5	μείγμα	0	ενσωμάτωση	E7, E1-L1, (790-1059, 6207-6151)
ATT21	Υψηλού βαθμού αλλοίωση	0	ενσωμάτωση	0,14	μείγμα	E7, E1-L1, (790-1059, 6207-6151)
PNL7	Καρκίνος	0	ενσωμάτωση	0,064	μείγμα	E2-L1, (2965-3337, 6418-6151)
PNL16	Καρκίνος	0,08	μείγμα	0,3	μείγμα	E7, E1-L1, (790-1059, 6207-6151)
D59	Καρκίνος	0	ενσωμάτωση	0	ενσωμάτωση	E7, E1-L1, (790-1059, 6207-6151)

Πίνακας 3.17. Συγκεντρωτικά αποτελέσματα σε οκτώ κλινικά δείγματα από τον προσδιορισμό της μορφής του γονιδιώματος του HPV16, μέσω του λόγου των τιμών E1/E6, E2/E6 και της αναδιάταξης του γονιδιώματος του HPV16.

Κεφάλαιο 4

4. Συζήτηση

Οι ιοί των θηλωμάτων εξελίχθηκαν για εκατομμύρια χρόνια για να επιβιώσουν σε ένα ευρύ φάσμα ζωικών ειδών, μεταξύ των οποίων και ο άνθρωπος. Ο Ιός των Ανθρώπινων Θηλωμάτων (HPV) αποτελεί μια ετερογενή ομάδα ιών με διαφορετικό επιθηλιακό τροπισμό και στρατηγικές κύκλου ζωής. Οι Ιοί των Ανθρώπινων Θηλωμάτων διαχωρίζονται στους τύπους χαμηλού (low-risk) και υψηλού κινδύνου (high-risk). Οι πιο συχνά εμφανιζόμενοι HPV τύποι υψηλού κινδύνου είναι οι HPV16 και HPV18, οι οποίοι εντοπίζονται στο 56,6% και 16% των περιπτώσεων καρκίνου του τραχήλου της μήτρας παγκοσμίως (Li et al., 2011, de Sanjose et al., 2010). Οι HPV τύποι υψηλού κινδύνου προκαλούν διάφορους τύπους καρκίνου, συμπεριλαμβάνοντας όλες τις περιπτώσεις καρκίνου του τραχήλου της μήτρας, ένα μεγάλο ποσοστό διαφορετικών τύπου καρκίνου του ουρογενετικού συστήματος, καθώς επίσης και όγκους στο στοματοφάρυγγα.(Cubie, 2013). Η εμφάνιση καρκίνου συνδέεται με "επίμονη" μόλυνση με HPV τύπους υψηλού κινδύνου και με την απορύθμιση της έκφρασης των ιικών ογκογονιδίων E6 και E7. Η ανεξέλεγκτη έκφραση αυτών των ογκογονιδίων οδηγεί στον εκτεταμένο πολλαπλασιασμό των μολυσμένων κυττάρων, στην ατελή επιδιόρθωση του DNA και στη συσσώρευση γενετικών λαθών στα μολυσμένα κύτταρα (Doorbar, 2012).

Παράλληλα με τη δράση των ογκοπρωτεϊνών του ιού η "επίμονη" μόλυνση από HPV τύπους υψηλού κινδύνου συνδέεται με αυξανόμενο κίνδυνο ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα και την εξέλιξη της αλλοίωσης σε καρκίνο. Έχει παρατηρηθεί ότι η ενσωμάτωση του HPV16 στο κυτταρικό γονιδίωμα γίνεται μέσω ρήξης της συνέχειας των E1 ή και E2 γονιδίων (Kalantari et al., 1998, Arias-Pulido et al., 2006, Cricca et al., 2009). Το γεγονός αυτό δίνει την δυνατότητα στον ιό να διεγείρει τον κυτταρικό κύκλο με διαρκή είσοδο του κυττάρου στη φάση της μίτωσης, δίνοντας στα μολυσμένα κύτταρα ισχυρό αναπτυξιακό πλεονέκτημα. (Jeon et al., 1995a, Doorbar, 2012). Ο συνεχής πολλαπλασιασμός των μολυσμένων κυττάρων προκαλεί συσσώρευση γενετικών λαθών και αποσταθεροποίηση του γονιδιώματος του ξενιστή με αποτέλεσμα την ανάπτυξη καρκίνου (Duensing et al., 2009).

Στη παρούσα διδακτορική διατριβή λαμβάνοντας υπόψη ότι ο HPV16 εντοπίζεται σε πάνω από το 50% των περιπτώσεων καρκίνου του τραχήλου της μήτρας πραγματοποιήθηκε

νουκλεοτιδική, αμινοξική και φυλογενετική ανάλυση σε ολόκληρη σχεδόν τη πρώιμη περιοχή του γονιδιώματος του HPV16, συμπεριλαμβάνοντας τα γονίδια E1, E2, E4, E6 και E7 σε κλινικά τραχηλικά δείγματα που συλλέχθηκαν από τον Ελληνικό γυναικείο πληθυσμό. Η μελέτη αυτή πραγματοποιήθηκε με σκοπό να προσδιοριστούν γονιδιωματικές τροποποιήσεις, όπως μεταλλάξεις, διπλασιασμοί, ελλείψεις και ενθέσεις στη πρώιμη περιοχή του HPV16 γονιδιώματος και να συσχετιστούν με το βαθμό της ενδοεπιθηλιακής αλλοίωσης, ενώ παράλληλα χαρακτηρίστηκαν οι διαφορετικές παραλλαγές του γονιδιώματος του HPV16 που κυκλοφορούν στον Ελληνικό πληθυσμό. Επιπλέον, πραγματοποιήθηκε ανάλυση της επιλεκτικής πίεσης που ασκείται στην πρώιμη περιοχή του γονιδιώματος του ιού με στόχο να μελετηθεί η εξέλιξη των πρώιμων γονιδίων του HPV16. Επιπρόσθετα, λαμβάνοντας υπόψη ότι η "επίμονη" μόλυνση από HPV16 σχετίζεται με αυξανόμενο κίνδυνο ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα, πραγματοποιήθηκε προσδιορισμός της μορφής του γονιδιώματος του ιού σε κλινικά τραχηλικά δείγματα και συσχετίστηκε με τον βαθμό της ενδοεπιθηλιακής αλλοίωσης, με σκοπό να χρησιμοποιηθεί ως δείκτης εξέλιξης της ενδοεπιθηλιακής αλλοίωσης σε καρκίνο του τραχήλου της μήτρας. Τέλος, προσδιορίστηκαν οι θέσεις ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα.

4.1 Νουκλεοτιδική ανάλυση των γονιδίων E6 και E7

Η ανάλυση των γονιδίων E6 και E7 πραγματοποιήθηκε με σκοπό να μελετηθούν οι ενδοτυπικοί πολυμορφισμοί που χαρακτηρίζουν τις πέντε διαφορετικές παραλλαγές του γονιδιώματος του HPV16 (Yamada et al., 1995, 1997, Wheeler et al., 1997, Swan et al., 2005, Pande et al., 2008). Παράλληλα εντοπίστηκαν νουκλεοτιδικές μεταλλάξεις στα HPV16 στελέχη που κυκλοφορούν στον Ελληνικό πληθυσμό. Από τη ανάλυση της αλληλουχίας των E6 και E7 γονιδίων παρατηρήθηκε ότι η πιο συχνά εμφανιζόμενη παραλλαγή του γονιδιώματος του HPV16 είναι η Ευρωπαϊκή παραλλαγή (T350G) και αμέσως μετά σε επίπεδο συχνότητας ακολουθεί το πρότυπο στέλεχος (Andersson et al., 2000, Zuna et al., 2009). Πιο συγκεκριμένα, η Ευρωπαϊκή παραλλαγή εντοπίστηκε στο 65.2% των δειγμάτων υψηλού βαθμού αλλοίωσης και στο 64.2% των περιπτώσεων χαμηλού βαθμού αλλοίωσης, ενώ το πρότυπο στέλεχος (350T) εντοπίστηκε στο 26.1% των δειγμάτων υψηλού βαθμού αλλοίωσης και στο 35.7% των δειγμάτων χαμηλού βαθμού αλλοίωσης. Προηγούμενες μελέτες έχουν προτείνει, ότι η μόλυνση από μη Ευρωπαϊκές παραλλαγές του γονιδιώματος του HPV16 επηρεάζουν την εξέλιξη της ενδοεπιθηλιακής αλλοίωσης σε καρκίνο, αναφέροντας σημαντική σύνδεση της HPV16 ενδοτυπικής παραλλαγής Asian

American με την εμφάνιση καρκίνου του τραχήλου της μήτρας (Junes-Gill et al., 2008, Quint et al., 2010, Sanchez et al., 2011, Tornesello et al., 2011). Σύμφωνα με την ανάλυση των E6 και E7 αλληλουχιών μη Ευρωπαϊκές παραλλαγές του γονιδιώματος του ιού χαρακτηρίστηκαν σε τέσσερα δείγματα (PNL7, A3, A4, A17). Πιο συγκεκριμένα, η ενδοτυπική παραλλαγή του HPV16 γονιδιώματος Asian American, εντοπίστηκε στο δείγμα PNL7, το οποίο χαρακτηρίστηκε ως καρκίνος του τραχήλου της μήτρας και στο δείγμα υψηλού βαθμού αλλοίωσης A3. Επιπλέον, στο δείγμα υψηλού βαθμού αλλοίωσης A4 εντοπίστηκε η ενδοτυπική παραλλαγή African type I, ενώ στο δείγμα A17 εντοπίστηκαν παράλληλα οι ενδοτυπικές παραλλαγές African type II και European. Το δείγμα A17 είναι η μοναδική περίπτωση στην οποία αναφέρθηκε μεικτή μόλυνση με περισσότερες από μία παραλλαγές του γονιδιώματος του ιού. Συμπερασματικά μη Ευρωπαϊκές παραλλαγές του HPV16 κυκλοφορούν στον Ελληνικό πληθυσμό και εντοπίζονται σε χαμηλού και υψηλού βαθμού δυσπλασίες. Η συγκεκριμένη μελέτη είναι η πρώτη προσπάθεια, που επιχειρεί τον χαρακτηρισμό ενδοτυπικών παραλλαγών που κυκλοφορούν στον Ελληνικό πληθυσμό. Παρόλα αυτά απαιτούνται περισσότερες επιδημιολογικές μελέτες έτσι ώστε να προσδιοριστεί η συχνότητα εμφάνισης των ενδοτυπικών παραλλαγών στον Ελληνικό πληθυσμό και να συνδεθούν με το βαθμό της ενδοεπιθηλιακής αλλοίωσης.

Επιπλέον, ο πιο συχνά εμφανιζόμενος πολυμορφισμός που ανιχνεύτηκε στα δείγματα που αναλύθηκαν στη παρούσα εργασία είναι ο T350G, ο οποίος εντοπίζεται στην πρότυπη αλληλουχία των HPV16 ενδοτυπικών παραλλαγών European και Asian American. (Yamada et al., 1995, 1997, Wheeler et al., 1997, Swan et al., 2005, Wu et al., 2006). Ο πολυμορφισμός T350G προκαλεί την αμινοξική αλλαγή L83V και έχει προταθεί ότι επικρατεί σε υψηλού βαθμού δυσπλασίες, ενώ παράλληλα έχει αναφερθεί ότι συνδέεται με την εξέλιξη της ενδοεπιθηλιακής αλλοίωσης σε καρκίνο του τραχήλου της μήτρας (Yamada et al., 1997, Andersson et al., 2000, Grodzki et al., 2006). Στη παρούσα εργασία η αμινοξική αλλαγή L83V παρουσίασε ίση κατανομή μεταξύ των δειγμάτων υψηλού βαθμού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Πιο συγκεκριμένα η μετάλλαξη L83V εντοπίστηκε στο 69.5% των περιπτώσεων υψηλού βαθμού αλλοίωσης και στο 64.28% των περιπτώσεων χαμηλού βαθμού αλλοίωσης, ενώ παράλληλα εντοπίστηκε σε δύο δείγματα που χαρακτηρίστηκαν ως καρκίνος του τραχήλου της μήτρας (PNL7, PNL16). Η παρουσία της μετάλλαξης L83V σε περιπτώσεις χαμηλού βαθμού αλλοίωσης, αποτελεί ένδειξη εξέλιξης της ενδοεπιθηλιακής αλλοίωσης από χαμηλού βαθμού αλλοίωσης σε υψηλού βαθμού αλλοίωση και κατ' επέκταση σε καρκίνο του τραχήλου της μήτρας. Επομένως, η

ανίχνευση της μετάλλαξης L83V σε κλινικά τραχηλικά δείγματα σε πρώιμο στάδιο μπορεί να δώσει σημαντικές πληροφορίες για τον τρόπο με τον οποίο ο κλινικός γιατρός θα διαχειριστεί την ενδοεπιθηλιακή αλλοίωση.

Παράλληλα με την ανίχνευση συγκεκριμένων ενδοτυπικών πολυμορφισμών που χαρακτηρίζουν τις πέντε διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16, εντοπίστηκαν εννέα νουκλεοτιδικές μεταλλάξεις στη περιοχή του γονιδίου E6 και επτά μεταλλάξεις στη περιοχή του γονιδίου E7, οι οποίες δεν έχουν αναφερθεί ξανά στη βιβλιογραφία. Όλες οι καινούργιες μεταλλάξεις που αναφέρθηκαν εντοπίζονται σε δείγματα τα οποία ομαδοποιήθηκαν στο ίδιο φυλογενετικό κλαδί με την Ευρωπαϊκή παραλλαγή του γονιδιώματος του ιού. Οι νουκλεοτιδικές και αμινοξικές μεταλλάξεις που εντοπίζονται στα γονίδια E6 και E7 της Ευρωπαϊκής παραλλαγής του HPV16 που κυκλοφορεί στον Ελληνικό πληθυσμό δηλώνουν, ότι τα ογκογονίδια E6 και E7 παρουσιάζουν σημαντική νουκλεοτιδική και αμινοξική μεταβλητότητα. Η συσσώρευση μεταλλάξεων ίσως να οφείλεται στην επιλεκτική πίεση που ασκείται στα γονίδια E6 και E7, με αποτέλεσμα να επικρατούν στο γονιδίωμα των στελεχών μεταλλάξεις, που ευνοούν την αλληλεπίδραση των ογκοπρωτεϊνών με διαφορετικές παραλλαγές κυτταρικών πρωτεϊνικών στόχων (DeFilippis, 2002, Chen et al., 2005, Carvajal-Rodriguez, 2008). Οι δυνάμεις που οδηγούν στην εξέλιξη των γονιδίων του HPV16 και ο τρόπος με τον οποίο η επιλογή επιδρά σε συγκεκριμένες θέσεις του γονιδιώματος του ιού θα συζητηθούν στη συνέχεια.

Επιπλέον, από τη νουκλεοτιδική ανάλυση των γονιδίων E6 και E7 παρατηρήθηκε η ύπαρξη δύο νέων E6-E7 αλληλουχιών σε τρία κλινικά δείγματα. Πιο συγκεκριμένα, στο δείγμα υψηλού βαθμού αλλοίωσης A15 μεταξύ των νουκλεοτιδίων 123 έως 350 παρατηρείται νουκλεοτιδική ομοιότητα με την ενδοτυπική παραλλαγή Asian American, ενώ μεταξύ των νουκλεοτιδίων 350 έως 839 απουσιάζει ένας ενδοτυπικός πολυμορφισμός στο γονίδιο E6 (A532G) και ένας ενδοτυπικός πολυμορφισμός στο γονίδιο E7 (T732C). Η απουσία των συγκεκριμένων πολυμορφισμών παρατηρείται στη πρότυπη αλληλουχία της HPV16 ενδοτυπικής παραλλαγής African type I. Παράλληλα με τη νουκλεοτιδική ανάλυση, το πρόγραμμα SBP έδειξε ένδειξη ανασυνδυασμού στη θέση 406, ενώ η ανάλυση Bootscan απέδειξε ότι το δείγμα A15 ομαδοποιείται με τις ενδοτυπικές παραλλαγές Asian American και African type I σε διαφορετικές περιοχές της αλληλουχίας E6-E7. Παρόλα αυτά, πρόσφατα αποδείχτηκε ότι συγκεκριμένοι ενδοτυπικοί πολυμορφισμοί εντός του γονιδίου E6 και της ρυθμιστικής περιοχής LCR, επιτρέπουν τη ταξινόμηση των πέντε κύριων γενών των ενδοτυπικών παραλλαγών του γονιδιώματος του HPV16, σε υψηλότερου βαθμού

ταξινόμησης χαρακτηρίζοντας τις "υπογενιές" (Cornet et al., 2012). Πιο συγκεκριμένα αποδείχθηκε ότι η ενδοτυπική παραλλαγή Asian American διαχωρίζεται στις "υπογενιές" Asian American 1, Asian American 2, και North American. Η απουσία του πολυμορφισμού A532G στο E6 γονίδιο που παρατηρήθηκε στο δείγμα A15 εντοπίζεται στην "υπογενιά" North American. Επομένως, η συγκεκριμένη αλληλουχία μπορεί να αποτελεί ένδειξη ότι στον Ελληνικό πληθυσμό κυκλοφορεί η "υπογενιά" North American. Ωστόσο, η ομοιότητα της υπογενιές North American με την ενδοτυπική παραλλαγή African type I, μεταξύ των νουκλεοτιδίων 350 έως 839, οδήγησε στο συμπέρασμα ότι οι υπογενιές που έχουν περιγραφεί μπορεί να είναι προϊόντα ανασυνδυασμού μεταξύ διαφορετικών ενδοτυπικών παραλλαγών του γονιδιώματος του ιού, γεγονός που δίνει σημαντικές πληροφορίες για τον τρόπο εξέλιξης του γονιδιώματος του HPV16.

Επιπλέον, στα δείγματα χαμηλού βαθμού αλλοίωσης A38 και 1143 παρατηρήθηκε νουκλεοτιδική ομοιότητα με την Ευρωπαϊκή παραλλαγή του γονιδιώματος του HPV16 (T350G). Ωστόσο, στο E6 γονίδιο των συγκεκριμένων δειγμάτων εντοπίστηκαν οι ενδοτυπικοί νουκλεοτιδικοί πολυμορφισμοί G132T και C143G. Ο συνδυασμός των συγκεκριμένων ενδοτυπικών πολυμορφισμών στην νουκλεοτιδική αλληλουχία του E6 γονιδίου εντοπίζεται στη πρότυπη αλληλουχία της ενδοτυπικής παραλλαγής HPV16 African type II. Με βάση τη Μπεϋζιανή Συμπερασματολογία παρατηρήθηκε, ότι στα κωδικόνια 10 (G132T), 14 (C143G) και 83 (T350G) ασκείται θετική επιλεκτική πίεση (Chen et al., 2005, DeFilippis, 2002, Carvajal-Rodriguez, 2008, Lee et al., 2008), ενώ παράλληλα έχει προταθεί ότι ο η θέση 10I διαχωρίζει τις ενδοτυπικές παραλλαγές HPV16 African type I και African type II και η θέση 14E διαχωρίζει τις Αφρικανικές παραλλαγές (African type I, African type II) από τις μη Αφρικανικές παραλλαγές του γονιδιώματος του ιού (Carvajal-Rodriguez, 2008). Σύμφωνα με αυτά τα δεδομένα παρατηρήθηκε, ότι στη περιοχή αυτή έχουν προηγηθεί γεγονότα θετικής επιλεκτικής πίεσης, καθιστώντας τη περιοχή αυτή του ιού εξαιρετικά πολυμορφική (Yang et al., 2000, Bush, 2001, Kosiol, 2006), ενώ παράλληλα η εξέλιξη έχει οδηγήσει στην επικράτηση των μεταλλάξεων G132T και C143G στην ενδοτυπική παραλλαγή African type II. Παρόλα αυτά, σε δύο δείγματα οι μεταλλάξεις G132T (10I) και C143G (14E) εμπεριέχονται στην αλληλουχία της Ευρωπαϊκής παραλλαγής του γονιδιώματος του HPV16. Επίσης το φυλογενετικό δέντρο που κατασκευάστηκε στη περιοχή των γονιδίων E6 και E7, έδειξε απόκλιση των συγκεκριμένων αλληλουχιών από τα τέσσερα φυλογενετικά κλαδιά που ταξινομούν τις πέντε διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16, ενώ το πρόγραμμα SBP έδειξε ύπαρξη

ανασυνδυασμού μεταξύ των διαφορετικών παραλλαγών του ιού στη θέση 143. Λαμβάνοντας υπόψη ότι η Ε6 ογκοπρωτεΐνη αλληλεπιδρά με διαφορετικές παραλλαγές της πρωτεΐνης p53 καταλήγουμε στο συμπέρασμα ότι νέες αλληλουχίες μπορεί να σχηματίζονται και να μοιράζονται κοινά χαρακτηριστικά με τις συγγενικές παραλλαγές του γονιδιώματος του HPV16. Δύο πιθανές εξηγήσεις έχουν προταθεί για την ύπαρξη των νέων αλληλουχιών. Πρώτον, καθώς οι θέσεις με θετική επιλεκτική πίεση υποδηλώνουν ότι η εξέλιξη έχει οδηγήσει σε αυξημένη μεταβλητότητα του Ε6 γονιδίου (Yang et al., 2000, Bush, 2001, Kosiol, 2006) και η Ε6 πρωτεΐνη αλληλεπιδρά με διαφορετικές παραλλαγές της πρωτεΐνης p53, οδηγούμαστε στο συμπέρασμα ότι η εξέλιξη επιδρά στην Ε6 πρωτεΐνη μέσω θετικής επιλεκτικής πίεσης, προάγοντας την επικράτηση συγκεκριμένων νουκλεοτιδικών μεταλλάξεων. Αυτές οι μεταλλάξεις μπορεί να προσδίδουν πλεονέκτημα στην Ε6 ογκοπρωτεΐνη για πιο αποτελεσματική αλληλεπίδραση με τις παραλλαγές των κυτταρικών πρωτεϊνών - στόχων καθώς επίσης και με το ανοσολογικό σύστημα του ασθενή (Chen et al., 2005, DeFilippis, 2002, Carvajal-Rodriguez, 2008, Lee et al., 2008). Δεύτερον, οι νέες αλληλουχίες μπορεί να προέρχονται από γεγονότα ανασυνδυασμού μεταξύ διαφορετικών παραλλαγών του γονιδιώματος του HPV16. Προηγούμενες μελέτες έχουν προτείνει ότι ανασυνδυασμός μπορεί να συμβεί στο γονιδίωμα των HPV (Varsani et al., 2006) Πιο συγκεκριμένα, ο Carvajal-Rodriguez το 2008 χρησιμοποιώντας αλληλουχίες από τη τράπεζα δεδομένων GenBank πρότεινε την ύπαρξη θέσης ανασυνδυασμού στη θέση 171 εντός του Ε7 γονιδίου του HPV16, ενώ σε προηγούμενη μελέτη εντοπίστηκε ο πρώτος ενδοτυπικός ανασυνδυασμός μεταξύ της ενδοτυπικής παραλλαγής HPV16 European και της ενδοτυπικής παραλλαγής HPV16 African type II σε κλινικό δείγμα (Jiang et al., 2009). Συμπερασματικά, ο ενδοτυπικός ανασυνδυασμός ίσως να είναι ένα σπάνιο φαινόμενο στην εξελικτική πορεία του HPV16, αλλά μπορεί να παίζει καθοριστικό ρόλο στον τρόπο εξέλιξης του HPV16 προς την ογκογονικότητα του.

4.2 Νουκλεοτιδική ανάλυση του Ε2 γονιδίου

Η νουκλεοτιδική ανάλυση του Ε2 γονιδίου έδειξε την ύπαρξη μεταλλάξεων σε όλη την περιοχή του Ε2 γονιδίου. Οι περισσότερες μεταλλάξεις εντοπίζονται κυρίως στις περιοχές του Ε2 γονιδίου που κωδικοποιούν για τη περιοχή "αρμός" και τη περιοχή διενεργοποίησης της Ε2 πρωτεΐνης και έχουν αναφερθεί από προηγούμενες μελέτες (Meissner et al., 1997, Veress et al., 1999, Casas et al., 1999, Eriksson et al., 1999, Graham et al., 2000, Giannoudis et al., 2001, Watts et al., 2002, Sathish et al., 2004, Swan et al.,

2005, Azizi et al., 2008, Jiang et al., 2009). Επιπρόσθετα, η νουκλεοτιδική ανάλυση του E2 γονιδίου στα κλινικά δείγματα που χρησιμοποιήθηκαν στη συγκεκριμένη διατριβή έδειξε την ύπαρξη τριών νέων μεταλλάξεων, οι οποίες αναφέρθηκαν για πρώτη φορά στα πλαίσια της παρούσας διατριβής. Πιο συγκεκριμένα, στο 5' άκρο του E2 γονιδίου που κωδικοποιεί για τη περιοχή διενεργοποίησης της E2 πρωτεΐνης, δύο νέες μεταλλάξεις αναφέρθηκαν στις θέσεις 2791 και 2823 σε δείγματα τα οποία έχουν χαρακτηριστεί ως Ευρωπαϊκές παραλλαγές του γονιδιώματος του HPV16. Επιπλέον ακόμα μια νέα μετάλλαξη (C3361T) εντοπίστηκε στη περιοχή E2 γονιδίου, που κωδικοποιεί για τη περιοχή "αρμός" της E2 πρωτεΐνης. Η μετάλλαξη αυτή χαρακτηρίστηκε ως σιωπηλή μετάλλαξη και εντοπίστηκε σε δείγματα τα οποία ομαδοποιήθηκαν με τις ενδοτυπικές παραλλαγές Asian American, African type I και African type II. Η μετάλλαξη αυτή δεν εντοπίστηκε στις πρότυπες αλληλουχίες των ενδοτυπικών παραλλαγών, γεγονός που οδήγησε στο συμπέρασμα ότι είναι μια καινούργια νουκλεοτιδική μετάλλαξη η οποία εντοπίζεται στις ενδοτυπικές παραλλαγές Asian American, African type I και African type II που κυκλοφορούν στον Ελληνικό πληθυσμό.

Επιπρόσθετα, στη περιοχή του E2 γονιδίου που κωδικοποιεί για τη περιοχή "αρμός" της E2 πρωτεΐνης, εντοπίστηκε η νουκλεοτιδική μετάλλαξη C3410T, η οποία προκαλεί την αμινοξική αλλαγή P219S. Η μετάλλαξη C3410T έχει συνδεθεί με υψηλότερου βαθμού δυσπλασίες και έχει προταθεί ότι τα στελέχη HPV16 που φέρουν τη συγκεκριμένη μετάλλαξη εμπλέκονται με μεγαλύτερη συχνότητα στην ανάπτυξη σοβαρότερων αλλοιώσεων και καρκίνου (Graham, 2000) Η μετάλλαξη αυτή παρουσίασε ίση κατανομή μεταξύ δειγμάτων υψηλού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, ενώ παράλληλα εντοπίστηκε σε δείγματα, τα οποία χαρακτηρίστηκαν ως ενδοτυπικές παραλλαγές Asian American, African type I, African type II. Η παρουσίαση της μετάλλαξης P219S στη περιοχή "αρμός" της E2 πρωτεΐνης, ίσως να επηρεάζει τη τριτοταγή διαμόρφωση της πρωτεΐνης, γεγονός που θα μπορούσε να αναστείλει τη λειτουργία της ως μεταγραφικός καταστολέας των ογκογονιδίων E6 και E7 (Graham, 2000). Η παρουσία της μετάλλαξης αυτής σε δείγματα χαμηλού βαθμού αλλοίωσης αποτελεί σημαντική ένδειξη εξέλιξης της αλλοίωσης σε υψηλότερου βαθμού αλλοιώσεις. Παράλληλα, στη συγκεκριμένη περιοχή του E2 γονιδίου που κωδικοποιεί για τη περιοχή "αρμός" της E2 πρωτεΐνης, ανιχνεύθηκαν αμινοξικές αλλαγές οι οποίες θα μπορούσαν να επηρεάσουν τη τριτοταγή διαμόρφωση της E2 πρωτεΐνης. Επομένως, παράλληλα με την ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα μέσω ρήξης της συνέχειας του E2 γονιδίου, κρίσιμο ρόλο στην εμφάνιση ενδοεπιθηλιακών

αλλοιώσεων παίζει η συσσώρευση μεταλλάξεων στο E2 γονίδιο, οι οποίες θα μπορούσαν να επηρεάσουν τη λειτουργία της E2 πρωτεΐνης ως μεταγραφικός καταστολές των ογκογονιδίων E6 και E7 (Graham, 2000, Gianoudis et al., 2001, Arias-Pulido et al., 2006, Azizi et al., 2008). Το γεγονός αυτό θα μπορούσε να οδηγήσει σε απορύθμιση της έκφρασης των ογκογονιδίων και τη πρόκληση σοβαρότερου βαθμού αλλοίωσης και καρκίνου του τραχήλου της μήτρας ακόμη και σε δείγματα που δεν παρατηρείται ενσωμάτωση του HPV16 στο κυτταρικό χρωμόσωμα. Επομένως, παράλληλα με τη μελέτη της ενσωμάτωσης του ιού μέσω ρήξης του E2 γονιδίου, κρίσιμο ρόλο παίζει η αμινοξική ανάλυση της E2 πρωτεΐνης.

4.2 Νουκλεοτιδική ανάλυση του E4 γονιδίου

Το ανοιχτό πλαίσιο ανάγνωσης του E2 γονιδίου εμπεριέχει ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του E4 γονιδίου μεταξύ των νουκλεοτιδίων 3332 έως 3619 (Doorbar, 2013). Στη παρούσα εργασία εντοπίστηκαν συνολικά δεκαπέντε νουκλεοτιδικές μεταλλάξεις στο E4 γονίδιο μεταξύ δειγμάτων χαμηλού βαθμού και υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Μέχρι στιγμής δεν έχουν πραγματοποιηθεί μελέτες για τη νουκλεοτιδική ανάλυση του E4 γονιδίου του HPV16 σε κλινικά δείγματα, με αποτέλεσμα να αναφερθούν μεταλλάξεις για πρώτη φορά. Όπως αναφέρθηκε σε προηγούμενη ενότητα (1.6.3), το προϊόν του πρώιμου γονιδίου E4 εκφράζεται από ένα ματισμένο μόριο mRNA του οποίου τα πέντε πρώτα αμινοξέα προέρχονται από το E1 γονίδιο (Middleton et al., 2003, Doorbar, 2013). Βασική λειτουργία της E1^{E4} πρωτεΐνης είναι η αλληλεπίδραση της με το δίκτυο ινιδίων κερατίνης στα διαφοροποιημένα κερατινοκύτταρα επάγοντας την καταστροφή του δικτύου και την απελευθέρωση των νεοσχηματιζόμενων ικών σωματιδίων (McIntosh et al., 2010). Πιο συγκεκριμένα, έχει παρατηρηθεί, ότι το N-τελικό άκρο της E1^{E4} πρωτεΐνης συμβάλει στη καταστροφή του δικτύου κερατίνης μέσω ενός συντηρημένου μοτίβου λευκίνης (LLXLL) (Roberts et al., 1994, McIntosh et al., 2010, Doorbar, 2013) και έχει αναφερθεί ότι η αποτελεσματική αλληλεπίδραση του δικτύου κερατίνης με την E1^{E4} πρωτεΐνη απαιτεί την ακεραιότητα των 16 πρώτων νουκλεοτιδίων του N-τελικού άκρου της πρωτεΐνης (Wang et al., 2004). Από τη νουκλεοτιδική ανάλυση του E4 γονιδίου παρατηρήθηκε ότι στη περιοχή του E4 γονιδίου που κωδικοποιεί για το N-τελικό άκρο της πρωτεΐνης, εντοπίστηκαν δύο μεταλλάξεις, οι οποίες προκαλούν αμινοξικές αλλαγές στη θέση 7 (A7V και A7P). Οι μεταλλάξεις αυτές ανιχνεύτηκαν σε δείγματα υψηλού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Η συμβολή των μεταλλάξεων αυτών στη δομή και τη λειτουργία της πρωτεΐνης θα πρέπει να μελετηθεί περαιτέρω έτσι ώστε να προκύψουν ασφαλή συμπεράσματα για τον

ρόλο που διαδραματίζουν στη δομή και τη λειτουργία της πρωτεΐνης. Επιπρόσθετα, έχει αναφερθεί ότι η E1^{E4} πρωτεΐνη μεταξύ των αμινοξέων 17 έως 45 συμβάλει στη παύση του κυτταρικού κύκλου στη φάση G2 (Davy et al., 2002, Davy et al., 2005). Πιο συγκεκριμένα, στη περιοχή αυτή υπάρχει μια θέση φωσφορυλίωσης κυκλινοεξαρτώμενων κινασών στις θέσεις T23 και T24. Με βάση τα αποτελέσματα από τη πολλαπλή ομοπαράθεση μεταξύ της πρότυπης αλληλουχίας, των αλληλουχιών που χαρακτηρίστηκαν στη παρούσα διατριβή και τις αλληλουχίες των ενδοτυπικών παραλλαγών του γονιδιώματος του ιού, παρατηρήθηκε ότι η περιοχή φωσφορυλίωσης είναι εξαιρετικά συντηρημένη. Όλες οι νουκλεοτιδικές μεταλλάξεις που ανιχνεύθηκαν στη περιοχή που κωδικοποιούν για τη περιοχή των αμινοξέων 17 έως 45 χαρακτηρίστηκαν ως σιωπηλές. Επιπρόσθετα, στη συγκεκριμένη περιοχή στη θέση 3410 παρατηρείται η νουκλεοτιδική μετάλλαξη C3410T, η οποία προκαλεί τη μετάλλαξη P219S στην E2 πρωτεΐνη που αναφέρθηκε και προηγουμένως (ενότητα 4.2), ενώ η ίδια μετάλλαξη δεν προκαλεί αμινοξική αλλαγή στο E4 γονίδιο.

Τέλος, έχει προταθεί ότι η C-τελική περιοχή της πρωτεΐνης εμπλέκεται στην καταστροφή του δικτύου κερατίνης (Roberts et al., 2008). Στη περιοχή αυτή εντοπίστηκαν τρεις αμινοξικές μεταλλάξεις Q66P, S72F και H75Q μεταξύ δειγμάτων χαμηλού και υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης. Οι αμινοξικές αυτές μεταλλάξεις εντοπίστηκαν σε δείγματα τα οποία συνδέθηκαν με μη Ευρωπαϊκές παραλλαγές του γονιδιώματος του ιού. Λαμβάνοντας υπόψη ότι η μόλυνση από μη Ευρωπαϊκές παραλλαγές του γονιδιώματος του HPV16, επηρεάζουν την εξέλιξη της ενδοεπιθηλιακής αλλοίωσης σε καρκίνο (Junes-Gill et al., 2008, Quint et al., 2010, Sanchez et al., 2011, Tornesello et al., 2011), συμπεράναμε ότι οι μεταλλάξεις αυτές ίσως να συμβάλουν σε αυξημένη ογκογονικότητα του ιού. Η ανίχνευση των μεταλλάξεων αυτών σε κλινικά δείγματα ίσως να μπορεί να δώσει χρήσιμες πληροφορίες για τον κίνδυνο ανάπτυξης σοβαρότερου βαθμού ενδοεπιθηλιακής αλλοίωσης και κατ' επέκταση καρκίνου του τραχήλου της μήτρας.

Επιπλέον με βάση τη νουκλεοτιδική ανάλυση των γονιδίων E4, E6 και E7 που πραγματοποιήθηκε στα ίδια δείγματα, παρατηρήθηκε ότι συγκεκριμένες μεταλλάξεις που εντοπίζονται στο E4 γονίδιο συνδέονται με ειδικούς ενδοτυπικούς πολυμορφισμούς στα γονίδια E6 και E7. Η συγκριτική μελέτη των E4, E6 και E7 γονιδίων στα ίδια κλινικά δείγματα σε συνδυασμό με το φυλογενετικό δέντρο που κατασκευάστηκε στη περιοχή του E4 γονιδίου, οδήγησε στο συμπέρασμα, ότι το E4 γονίδιο μπορεί να ταξινομήσει το γονιδίωμα του HPV16 στα τέσσερα διαφορετικά φυλογενετικά κλαδιά που οργανώνονται οι πέντε διαφορετικές ενδοτυπικές παραλλαγές του HPV16 γονιδιώματος. Επομένως το E4

γονίδιο του HPV16 είναι μια μικρή πολυμορφική περιοχή του γονιδιώματος του ιού, μεγέθους 288bp, ή οποία μπορεί να δώσει πληροφορίες για τη φυλογενετική προέλευση του γονιδιώματος του ιού, χωρίς να απαιτείται η ταυτόχρονη ανάλυση των γονιδίων E6 και E7.

4.3 Νουκλεοτιδική ανάλυση του E1 γονιδίου

Το E1 γονίδιο κωδικοποιεί μια ενεργή DNA ελικάση, η οποία προάγει την αντιγραφή του γονιδιώματος του ιού όταν βρίσκεται στην επισωμική του μορφή. Το ανοιχτό πλαίσιο ανάγνωσης του E1 γονιδίου είναι το μεγαλύτερο στο γονιδίωμα του HPV16 και το πιο συντηρημένο μεταξύ των διαφορετικών τύπων HPV (Garcia- Vallve et al., 2005, Bergvall, 2013). Παρόλα αυτά, όπως αναφέρθηκε και για το γονίδιο E4, μέχρι σήμερα δεν έχει πραγματοποιηθεί εκτεταμένη μελέτη της νουκλεοτιδικής αλληλουχίας του E1 γονιδίου. Από την ανάλυση του E1 γονιδίου παρατηρήθηκαν συνολικά είκοσι επτά νουκλεοτιδικές μεταλλάξεις, οι οποίες δεν έχουν αναφερθεί στη βιβλιογραφία με μόνη εξαίρεση τη μετάλλαξη στη θέση 1053 (Sabol et al., 2012). Η μετάλλαξη A1053C εντοπίστηκε σε δείγματα υψηλού και χαμηλού βαθμού αλλοίωσης και προκαλεί την αμινοξική αλλαγή E63D, η οποία εντοπίζεται στο N- τελικό άκρο της E1 πρωτεΐνης. Η περιοχή αυτή είναι υπεύθυνη για τη πρόσδεση της E1 πρωτεΐνης στην H1 ιστόνη (Swindle, 1998, Wilson et al., 2002). Επιπλέον, αναφέρθηκαν μια σειρά νουκλεοτιδικών μεταλλάξεων μεταξύ των δειγμάτων υψηλού βαθμού και χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, οι οποίες προκαλούν αμινοξικές αλλαγές στην E1 πρωτεΐνη. Αντιθέτως, οι μεταλλάξεις A1668G, G2073A, T2169C, T2189C, A2453T, C2454T, A2587T, G2650A ανιχνεύτηκαν σε δείγματα, τα οποία χαρακτηρίστηκαν ως υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση. Πιο συγκεκριμένα οι μεταλλάξεις A1668G, G2073A, T2169C χαρακτηρίστηκαν ως σιωπηλές μεταλλάξεις, ενώ οι μεταλλάξεις T2189C, A2453T, C2454T, A2587T και G2650A, προκαλούν τις αμινοξικές αλλαγές I442T, N530I, N530I, R575W και E596K, αντίστοιχα. Οι μεταλλάξεις αυτές εντοπίζονται στο C-τελικό άκρο της πρωτεΐνης, το οποίο έχει δραστηριότητα ATPασης, 3'-5' ελικάσης και συμβάλει στον ολιγομερισμό της πρωτεΐνης σε ένα εξαμερές (Titolo et al., 2000, White et al., 2001, Wilson et al., 2002). Η επικράτηση των συγκεκριμένων μεταλλάξεων σε υψηλότερου βαθμού αλλοιώσεις και ο ρόλος τους στη δομή και τη λειτουργία της E1 πρωτεΐνης θα πρέπει να μελετηθεί περαιτέρω, έτσι ώστε να διευκρινιστεί ο τρόπος με τον οποίο οι μεταλλάξεις αυτές επηρεάζουν την αντιγραφή του ιικού DNA.

Παράλληλα με την ενίσχυση και τη μελέτη του E1 γονιδίου, πραγματοποιήθηκε νουκλεοτιδική ανάλυση των γονιδίων E6 και E7 με στόχο να συνδεθούν οι μεταλλάξεις στο γονίδιο E1 με ενδοτυπικούς πολυμορφισμούς που χαρακτηρίζουν τις πέντε διαφορετικές ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16. Από τη συγκριτική μελέτη των γονιδίων παρατηρήθηκε, ότι οι μεταλλάξεις που ανιχνεύτηκαν στο E1 γονίδιο συνδέονται με τις ενδοτυπικές παραλλαγές HPV16 African type II και Asian American. Παράλληλα, το φυλογενετικό δέντρο που κατασκευάστηκε με βάση την αλληλουχία του E1 γονιδίου, έδειξε ότι το E1 γονίδιο συμβάλει στη ταξινόμηση του γονιδιώματος του HPV16 στα τέσσερα φυλογενετικά κλαδιά, που οργανώνεται το γονιδίωμα του HPV16 με τιμές bootstrap values > 80%. Συμπερασματικά, στη παρούσα διδακτορική διατριβή πραγματοποιήθηκε μια πρώτη προσέγγιση νουκλεοτιδικής και αμινοξικής ανάλυσης των γονιδίων E1 και E4 σε κλινικά τραχηλικά δείγματα. Μέχρι στιγμής οι περιοχές αυτές του γονιδιώματος του ιού δεν έχουν μελετηθεί σε κλινικά δείγματα και δεν έχουν αναφερθεί μεταλλάξεις, οι οποίες συνδέονται με το βαθμό ενδοεπιθηλιακής αλλοίωσης και με τις ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16. Στόχος επίσης ήταν να μελετηθεί η πολυμορφικότητα των γονιδίων αυτών και να εντοπιστούν μεταλλάξεις στις αντίστοιχες πρωτεΐνες που μπορεί να επηρεάσουν τη δομή και τη λειτουργία των πρωτεϊνών. Παράλληλα η ύπαρξη νουκλεοτιδικών και αμινοξικών μεταλλάξεων στις ενδοτυπικές παραλλαγές του γονιδιώματος του HPV16 που κυκλοφορούν στον Ελληνικό πληθυσμό είναι δυνατόν να δώσουν χρήσιμες πληροφορίες για την ογκογονικότητα του HPV16.

4.4 Ανάλυση του διπλασιασμού E1-1374^{63nt}

Η νουκλεοτιδική ανάλυση του E1 γονιδίου οδήγησε στην ανίχνευση ενός διπλασιασμού 63 νουκλεοτιδίων στο 5' άκρο του E1 γονιδίου (E1-1374^{63nt}) μεταξύ των νουκλεοτιδίων 1311 και 1374. Ο διπλασιασμός των 63 νουκλεοτιδίων δεν καταστρέφει το ανοιχτό πλαίσιο ανάγνωσης του γονιδίου και προκαλεί τον διαδοχικό διπλασιασμό της περιοχής των αμινοξέων 150 έως 170. Η περιοχή του διπλασιασμού βρίσκεται στο N- τελικό άκρο της E1 πρωτεΐνης, το οποίο ευθύνεται για τη πρόσδεση της E1 πρωτεΐνης στην H1 ιστόνη (Swindle, 1998, Wilson et al., 2002). Μέχρι στιγμής ο διπλασιασμός των 63 νουκλεοτιδίων του E1 γονιδίου έχει αναφερθεί στη Κίνα, στη Κροατία και στη Σλοβενία (Dong et al., 2000, Sabol et al., 2008, Bogovac et al., 2011, Sabol et al., 2012). Πιο συγκεκριμένα, ο ενδογονιδιακός διπλασιασμός, έχει εντοπιστεί στο 8% των εξεταζόμενων δειγμάτων στη Σλοβενία (Bogovac et al., 2011), στο 10% των εξεταζόμενων δειγμάτων στη

Κροατία (Sabol et al., 2008, 2012), ενώ έχει εντοπιστεί σε ορισμένες περιπτώσεις καρκίνου του τραχήλου της μήτρας στη Κίνα. Στη παρούσα μελέτη το 24% των εξεταζόμενων δειγμάτων που προήρθαν από τον Ελληνικό γυναικείο πληθυσμό φέρουν τον διπλασιασμό E1-1374⁶³nt. Η υψηλή επικράτηση του διπλασιασμού στον Ελληνικό πληθυσμό, ίσως να οφείλεται στο γεγονός ότι η παραλλαγή αυτή του γονιδιώματος του ιού παρουσιάζει αυξημένη κατανομή σε πιο ανατολικές χώρες της Ευρώπης. Περισσότερα δείγματα από τον Ελληνικό πληθυσμό και από άλλους πληθυσμούς της Ευρώπης και της Ασίας θα πρέπει να μελετηθούν, έτσι ώστε να εξακριβωθεί η παραπάνω υπόθεση και να μελετηθεί η διασπορά της καινούργιας παραλλαγής του HPV16 στον Παγκόσμιο Πληθυσμό.

Επιπλέον, παρατηρήθηκε ότι όλα τα δείγματα που φέρουν το διπλασιασμό E1-1374⁶³nt, έχουν ταυτόχρονα τον πολυμορφισμό T350G και ομαδοποιούνται με την Ευρωπαϊκή παραλλαγή του γονιδιώματος του HPV16 ($\chi^2 = 5.076$, $P = 0.024$) (Sabol et al., 2008, Bogovac et al., 2011, Sabol et al., 2012). Παρά το γεγονός ότι ο πολυμορφισμός T350G επικρατεί σε υψηλού βαθμού δυσπλασίες και συνδέεται με την εξέλιξη της ενδοεπιθηλιακής αλλοίωσης σε καρκίνο του τραχήλου της μήτρας (Yamada et al., 1997, Andersson et al., 2000, Grodzki et al., 2006), δεν παρατηρήθηκε σημαντική σύνδεση μεταξύ του ενδογονιδιακού διπλασιασμού, ο οποίος βρίσκεται πάντα σε συνδυασμό με τη μετάλλαξη T350G, και του βαθμού της ενδοεπιθηλιακής αλλοίωσης. Επιπρόσθετα, παρατηρήθηκε σημαντική σύνδεση της μετάλλαξης A1053C και του διπλασιασμού E1-1374⁶³nt ($\chi^2 = 23.102$, $P < 0.0001$). Η μετάλλαξη A1053C δεν εντοπίστηκε σε δείγματα χωρίς τον διπλασιασμό, καθώς επίσης δεν εντοπίστηκε ούτε στη πρότυπη αλληλουχία των ενδοτυπικών παραλλαγών του γονιδιώματος του HPV16. Επομένως ο διπλασιασμός E1-1374⁶³nt βρίσκεται πάντα σε συνδυασμό με τις μεταλλάξεις T350G (E6 γονίδιο) και A1053C (E1 γονίδιο).

Στη παρούσα διατριβή για να προσδιοριστεί ο τρόπος που επιδρά ο διπλασιασμός στη δομή και τη λειτουργία της E1 πρωτεΐνης, υπολογίστηκε ο λόγος dN/dS, έτσι ώστε να προσδιοριστεί ο τρόπος με τον οποίο επιδρά η εξέλιξη στη περιοχή του διπλασιασμού. Με βάση τα αποτελέσματα που προήρθαν από τη συγκεκριμένη μελέτη παρατηρήθηκε, ότι αρνητική επιλεκτική πίεση ασκείται στη περιοχή του διπλασιασμού, γεγονός που δηλώνει ότι η εξέλιξη διατηρεί τη συγκεκριμένη περιοχή σταθερή, χωρίς να επηρεάζει αρνητικά τη λειτουργία και τη δομή του N-τελικού άκρου της E1 πρωτεΐνης (Vandepoele et al., 2005). Η υπόθεση αυτή ενισχύθηκε από πρόσφατη έρευνα, στην οποία έχει προβλεφτεί μέσω βιοπληροφορικής η τρισδιάστατη διαμόρφωσης της E1 πρωτεΐνης, που φέρει τον

διπλασιασμό των 21 αμινοξέων. Η συγκεκριμένη μελέτη έχει προτείνει ότι ο διπλασιασμός και η αμινοξική μετάλλαξη E63D, εντοπίζονται στην επιφάνεια της E1 πρωτεΐνης χωρίς να προκαλούνται σημαντικές αλλαγές στη δομή και στη λειτουργία της E1 πρωτεΐνης (Sabol et al., 2012).

Από την ανάλυση των αποτελεσμάτων δύο πιθανές εξηγήσεις δόθηκαν στην παρούσα διατριβή για τον μηχανισμό ύπαρξης του διπλασιασμού E1-1374^{63nt}. Πρώτον, επειδή η E1 πρωτεΐνη είναι μια εξαιρετικά συντηρημένη περιοχή του γονιδιώματος του HPV16, μπορεί να συμβαίνει ομόλογος ανασυνδυασμός μεταξύ διαφορετικών στελεχών της Ευρωπαϊκής παραλλαγής του γονιδιώματος του ιού, προκαλώντας τον ενδογονιδιακό διπλασιασμό των 63 νουκλεοτιδίων. Δεύτερον, η σύνδεση του διπλασιασμού E1-1374^{63nt} με την Ευρωπαϊκή παραλλαγή του HPV16 ίσως να αποτελεί ένδειξη της παρουσίας μια νέας παραλλαγής του γονιδιώματος του ιού που κυκλοφορεί με υψηλή συχνότητα στον Ευρωπαϊκό πληθυσμό (Sabol et al., 2008).

4.5 Εξελικτική μελέτη της πρώιμης περιοχής του HPV16 DNA

Η εξελικτική μελέτη της πρώιμης περιοχής του γονιδιώματος του ιού πραγματοποιήθηκε, μέσω υπολογισμού του λόγου dN/dS με στόχο να προσδιοριστεί η επιλεκτική πίεση που ασκείται στα διαφορετικά γονίδια της πρώιμης περιοχής του HPV16 DNA. Η ανάλυση της πρώιμης περιοχής του γονιδιώματος του ιού έδειξε ότι αρνητική επιλεκτική πίεση ασκείται στα γονίδια E1, E2, ουδέτερη πίεση ασκείται στα γονίδια E6 και E4, ενώ θετική επιλεκτική πίεση ασκείται στο γονίδιο E7. Παρά το γεγονός ότι ασκείται διαφορετική επιλεκτική πίεση στις διαφορετικές περιοχές του γονιδιώματος του ιού, εντοπίστηκαν συγκεκριμένες θέσεις στις οποίες επιδρά θετική επιλεκτική πίεση.

Οι πιο σημαντικές θέσεις στις οποίες έχει προταθεί ότι κυριαρχεί θετική επιλεκτική πίεση στο γονιδίωμα του HPV16, είναι τα κωδικόνια 10, 14 και 83 του E6 ογκογονιδίου (DeFilippis et al., 2002, Chen et al., 2005, Carvajal-Rodríguez, 2008, Lee et al., 2008). Στα δείγματα που εξετάστηκαν στα πλαίσια της παρούσας διατριβής, θετική επιλεκτική πίεση ανιχνεύτηκε μόνο στο κωδικόνιο 83 του E6 γονιδίου. Η απουσία θετικής επιλεκτικής πίεσης στις θέσεις των κωδικονίων 10 και 14 ίσως να οφείλεται στην επικράτηση της Ευρωπαϊκής παραλλαγής (T350G) του γονιδιώματος του ιού στον Ελληνικό πληθυσμό. Αντιθέτως, θετική επιλεκτική πίεση στις θέσεις των κωδικονίων 10, 14 και 83 εντοπίστηκε στις νέες παραλλαγές των αλληλουχιών E6-E7, που ανιχνεύτηκαν μέσω της νουκλεοτιδικής ανάλυσης της περιοχής E6-E7. Πιο συγκεκριμένα, στην ίδια αλληλουχία παρατηρούνται αλλαγές στις θέσεις των

κωδικονίων 10, 14 και 83. Η επίδραση θετικής επιλεκτικής πίεσης σε συγκεκριμένες θέσεις του γονιδίου οφείλεται στη πίεση που ασκεί το ανοσολογικό σύστημα του ξενιστή, καθώς επίσης και στην αλληλεπίδραση της Ε6 πρωτεΐνης με διαφορετικές παραλλαγές πρωτεϊνών στόχων που επικρατούν στον πληθυσμό (Chen et al., 2005). Πιο συγκεκριμένα, έχει αναφερθεί ότι η αποτελεσματική παρουσίαση των ικών πεπτιδίων από τα αντιγόνα ιστοσυμβατότητας τάξης I και II (HLA), επηρεάζεται από γενετικούς πολυμορφισμούς που εντοπίζονται στο γονιδίωμα του HPV16 καθώς επίσης και από πολυμορφισμούς που εντοπίζονται στα αντιγόνα ιστοσυμβατότητας. Έχει αποδειχθεί ότι η παραλλαγή L83V (θέση που ασκείται θετική επιλεκτική πίεση πίεση), η οποία παρατηρείται στο μεγαλύτερο ποσοστό του πληθυσμού, συνδέεται με διαφορετικά αλληλόμορφα HLA τάξης I (B*44, B*51 και B*57) (Andersson et al., 2000, Zehbe et al., 2003, Chakrabarti et al., 2004). Επομένως, η ετερογένεια των διαφορετικών στελεχών HPV16 σε συνδυασμό με την ετερογένεια της ανοσολογικής απόκρισης του ξενιστή, ίσως να επηρεάζει τον κύκλο ζωής του HPV16 και την έκβαση της λοίμωξης (Andersson et al., 2000, Zehbe et al., 2003, Chakrabarti et al., 2004). Ακόμη, η αλληλεπίδραση των ικών πρωτεϊνών με διαφορετικές παραλλαγές κυτταρικών πρωτεϊνών - στόχων, μπορεί να συμβάλει στην εξέλιξη του ιικού DNA. Για παράδειγμα, έχει αναφερθεί ότι αλλαγές στα κωδικόνια 14 και 27, μπορεί να επηρεάζουν την αλληλεπίδραση της Ε6 πρωτεΐνης με την ουβικουτινική λιγάση Ε6ΑΡ (Watts et al., 2002) Επιπρόσθετα, ένας πολυμορφισμός της πρωτεΐνης p53 στο κωδικόνιο 72 Arg/Pr σε συνδυασμό με τον πολυμορφισμό L83V του Ε6 γονιδίου αποτελεί σημαντικό παράγοντα ανάπτυξης καρκίνου του τραχήλου της μήτρας (Brady et al., 1999, Liu et al., 2013). Επιπλέον θετική επιλεκτική πίεση παρατηρήθηκε στο κωδικόνιο 85 (G85S, G85D) του Ε7 γονιδίου. Η θέση αυτή δεν έχει αναφερθεί σε προηγούμενες μελέτες και εντοπίζεται στο C-τελικό άκρο της Ε7 πρωτεΐνης των HPV16 στελεχών που κυκλοφορούν στον Ελληνικό πληθυσμό. Το C-τελικό άκρο της Ε7 πρωτεΐνης εμπλέκεται στη σύνδεση της πρωτεΐνης με την pRb και άλλες κυτταρικές πρωτεΐνες (Hebner, 2006, Munger et al., 2004). Όπως και στη περίπτωση της πρωτεΐνης p53, διαφορετικές παραλλαγές της pRb ίσως να επικρατούν στον πληθυσμό, με αποτέλεσμα η θέση 85 να επηρεάζει την αλληλεπίδραση της Ε7 πρωτεΐνης με την οικογένεια των πρωτεϊνών του ρετινοβλαστώματος. Ακόμη, θετική επιλεκτική πίεση παρατηρήθηκε στα κωδικόνια 59 (L59I/T) και 66 (Q66P) του Ε4 γονιδίου. Οι θέσεις 59 και 66 διαχωρίζουν τις HPV16 ενδοτυπικές παραλλαγές African type I, African type II και Asian American από τις ενδοτυπικές παραλλαγές Asian και European και εντοπίζονται στο C-τελικό άκρο της Ε1^{Ε4} πρωτεΐνης. Η συγκεκριμένη περιοχή της πρωτεΐνης συμβάλει στη καταστροφή του δικτύου

κερατίνης και την απελευθέρωση των νεοσχηματιζόμενων ικών σωματιδίων (McIntosh et al., 2010). Η επίδραση θετικής επιλεκτικής πίεσης στις θέσεις 59 και 66 δεν έχουν αναφερθεί σε προηγούμενη εργασία και αποτελούν αντικείμενο περαιτέρω μελέτης, έτσι ώστε να διευκρινιστεί ο ρόλος των θέσεων αυτών στην αλληλεπίδραση της E1^{E4} πρωτεΐνης με το δίκτυο κερατίνης.

Συμπερασματικά, ο προσδιορισμός νέων θέσεων επιλεκτικής πίεσης στα γονίδια E7 και E4, ίσως να σχετίζεται με την επικράτηση συγκεκριμένων παραλλαγών των πρωτεϊνών - στόχων που επικρατούν στον Ελληνικό πληθυσμό, η και από την ετερογένεια της ανοσολογικής απόκρισης των ξενιστών (Andersson et al., 2000, Zehbe et al., 2003, Chakrabarti et al., 2004). Με βάση τα αποτελέσματα του λόγου των τιμών dN/dS οι συγκεκριμένες περιοχές των γονιδίων E6, E7 και E4 εξελίσσονται μέσω επίδρασης θετικής επιλεκτικής πίεσης.

Τέλος, στις περιοχές των γονιδίων E1 και E2 παρατηρήθηκε αρνητική επιλεκτική πίεση (dN/dS<1). Το φαινόμενο αυτό δηλώνει ότι στις περιοχές των γονιδίων E1 και E2 υπάρχει αναστολή της φυσικής επιλογής. Το γεγονός αυτό ίσως να οφείλεται στο ρόλο που διαδραματίζουν οι ικές πρωτεΐνες E1 και E2 στην αντιγραφή και μεταγραφή του ικού DNA, καθώς η συσσώρευση και η εγκαθίδρυση μεταλλάξεων στις πρωτεΐνες αυτές μπορεί να επηρεάζουν αρνητικά το κύκλο ζωής του ιού. Αν και στις περιοχές των γονιδίων E1 και E2 παρατηρήθηκαν μεταλλάξεις σε αρκετές θέσεις των E1 και E2 πρωτεϊνών, φαίνεται πως δεν επηρεάζουν αρνητικά τη δομή και την λειτουργία των ικών πρωτεϊνών. Επομένως η εξέλιξη διατηρεί τις θέσεις αυτές σταθερές έτσι ώστε να προάγεται αποτελεσματικά ο κύκλος ζωής του ιού.

4.6 Προσδιορισμός της μορφής του γονιδιώματος του HPV16

Στη συνέχεια της διδακτορικής διατριβής πραγματοποιήθηκε προσδιορισμός της μορφής του γονιδιώματος του HPV16 σε δείγματα που έχουν χαρακτηριστεί ως υψηλού βαθμού, χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης και καρκίνου του τραχήλου της μήτρας. Ο προσδιορισμός της μορφής του γονιδιώματος του ιού πραγματοποιήθηκε με τη μέθοδο της Ποσοτικής PCR Πραγματικού Χρόνου, μέσω του υπολογισμού του λόγου των τιμών E1/E6 και E2/E6. Η διαδικασία του προσδιορισμού της μορφής του γονιδιώματος του HPV16, πραγματοποιήθηκε σε δύο βήματα. Πρώτον, πραγματοποιήθηκε χαρτογράφηση των γονιδίων E1 και E2 με σκοπό να προσδιοριστούν οι κοινές θέσεις ρήξης του γονιδιώματος του HPV16. Δεύτερον, για τον ποσοτικό προσδιορισμό των γονιδίων E1 και E2 και τον

προσδιορισμό του λόγου E1/E6 και E2/E6, χρησιμοποιήθηκαν εκκινητές που στοχεύουν στις θέσεις ρήξης, έτσι ώστε να προσδιοριστεί με μεγαλύτερη ακρίβεια η μορφή του γονιδιώματος του ιού.

Φαινόμενα ρήξης εντός της αλληλουχίας των γονιδίων E1 και E2, έχουν συνδεθεί με την ενσωμάτωση του ιού στο κυτταρικό χρωμόσωμα (Kalantari et al., 1998, Arias-Pulido et al., 2006, Cricca et al., 2009). Στη παρούσα διατριβή, γεγονότα ρήξης καθορίστηκαν σε κλινικά δείγματα, όταν δεν εντοπίστηκε σήμα ενίσχυσης μετά από την εφαρμογή της PCR με ένα από τους δώδεκα εκκινητές που χρησιμοποιήθηκαν για τη χαρτογράφηση των E1 και E2 γονιδίων, ενώ παράλληλα τα δείγματα ήταν θετικά για την ενίσχυση του E6 γονιδίου. Η ενίσχυση του E6 γονιδίου πραγματοποιήθηκε παράλληλα στα ίδια δείγματα, έτσι ώστε να ελεγχθεί η ακεραιότητα του HPV16 DNA και να καθοριστεί ότι η ρήξη του γονιδίου E1 ή/και E2, προέρχεται από γεγονότα ενσωμάτωσης του ιού στο κυτταρικό χρωμόσωμα. Μέχρι στιγμής έχουν αναφερθεί στη βιβλιογραφία διάφορες θέσεις ρήξης εντός της αλληλουχίας του E1 γονιδίου, ενώ έχει προταθεί ότι η πιο συχνή θέση ρήξης του E2 γονιδίου, είναι η περιοχή που κωδικοποιεί για την περιοχή "αρμός" της E2 πρωτεΐνης (Kalantari et al., 1998, Arias-Pulido et al., 2006, Cricca et al., 2009). Πιο συγκεκριμένα, έχει αναφερθεί ένα πρότυπο ρήξης των γονιδίων E1 και E2, που περιλαμβάνει το 3' άκρο του E1 γονιδίου και της θέσης που κωδικοποιεί για τη περιοχή "αρμό" της E2 πρωτεΐνης (Cricca et al., 2009), ενώ αντίθετα άλλη προσέγγιση προτείνει ότι ρήξη των E1 και E2 γονιδίων συμβαίνει πιο συχνά στο 5' άκρο των E1 και E2 γονιδίων, αντίστοιχα (Kalantari et al., 1998).

Στη παρούσα διατριβή ρήξη στο E1 γονίδιο εντοπίστηκε μόνο στο 5' άκρο του E1 γονιδίου με πιο συχνή θέση να είναι η περιοχή μεταξύ των νουκλεοτιδίων 1059 και 1323. Πιο συγκεκριμένα, ρήξη μεταξύ των νουκλεοτιδίων 1059 έως 1323 (εκκινητές E1F2-E1R2) εντοπίστηκε σε δεκαοχτώ από τα ογδόντα τέσσερα δείγματα που εξετάστηκαν (21.4%), ενώ μεταξύ των νουκλεοτιδίων 878 έως 1118 και 1503 έως 1784 εντοπίστηκε ρήξη σε δύο δείγματα (2.4%) και ένα δείγμα (1.2%), αντίστοιχα. Επιπλέον θέσεις ρήξης στο E2 γονίδιο παρατηρούνται σε ολόκληρο το ανοιχτό πλαίσιο ανάγνωσης του γονιδίου με πιο συχνή θέση ρήξης να είναι η περιοχή μεταξύ των νουκλεοτιδίων 3172 έως 3649, που κωδικοποιεί για τη περιοχή "αρμός" της E2 πρωτεΐνης (Kalantari et al., 1998, Arias-Pulido et al., 2006, Cricca et al., 2009). Πιο συγκεκριμένα, η περιοχή μεταξύ των νουκλεοτιδίων 3172 έως 3649 έχει υποστεί ρήξη σε δεκατρία δείγματα από τα ογδόντα τέσσερα δείγματα που εξετάστηκαν (15.5%). Επιπλέον, μεταξύ των νουκλεοτιδίων 3597 έως 3873 και 2735 έως 3210 εντοπίστηκε ρήξη σε τρία δείγματα (3.57%), και ένα δείγμα, αντίστοιχα (1.2%).

Επιπρόσθετα, ένα διαφορετικό μοτίβο ρήξης των E1 και E2 γονιδίων εντοπίστηκε στο 10.7% των εξεταζόμενων δειγμάτων, συμπεριλαμβάνοντας τέσσερις περιπτώσεις υψηλού βαθμού αλλοίωσης (D2, D13, D14, D15) και πέντε περιπτώσεις καρκίνου του τραχήλου της μήτρας (D58, D62, D64, D65, D66). Το μοτίβο αυτό περιλαμβάνει τη ταυτόχρονη ρήξη των γονιδίων E1 και E2. Ταυτόχρονη ρήξη μεταξύ των γονιδίων E1 και E2 έχει παρατηρηθεί και σε προηγούμενες μελέτες σε υψηλού βαθμού αλλοιώσεις και καρκίνο του τραχήλου της μήτρας. Το γεγονός αυτό έχει αποδοθεί σε φαινόμενα αναδιάταξης του γονιδιώματος του HPV16 χρωμόσωμα (Arias-Pulido et al., 2006, Cricca et al., 2009), με αποτέλεσμα την ταυτόχρονη έλλειψη διαφορετικών περιοχών των γονιδίων E1 και E2. Λαμβάνοντας υπόψη ότι η ενσωμάτωση του ιού μπορεί να περιλαμβάνει ένα μοναδικό αντίγραφο του ιού ενσωματωμένο στο κυτταρικό γονιδίωμα, ή αλλυσομερή ιικών γονιδιωμάτων σε μία περιοχή του κυτταρικού γονιδιώματος (Jeon et al., 1995, Xu et al., 2013, Akagi et al., 2014), καταλήξαμε στο συμπέρασμα, ότι εκτός από τα φαινόμενα αναδιάταξης ίσως η ταυτόχρονη ρήξη στα γονίδια E1 και E2 να οφείλεται σε διαφορετικά πρότυπα ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα.

Όσον αφορά τη σύνδεση της ρήξης του E1 και E2 γονιδίου με το βαθμό της ενδοεπιθηλιακής αλλοίωσης προέκυψε ότι ρήξη του E2 γονιδίου εμπλέκεται αποκλειστικά σε δείγματα τα οποία έχουν χαρακτηριστεί ως υψηλού βαθμού ενδοεπιθηλιακή νεοπλασία (9/33, 27.27%) και σε περιπτώσεις καρκίνου του τραχήλου της μήτρας (8/11, 72.72%) (χ^2 , test, $P < 0.01$). Επιπλέον δεν καταγράφηκε καμία θέση ρήξης του E2 γονιδίου σε δείγματα χαμηλού βαθμού αλλοίωσης. Αντίθετα, η χαρτογράφηση του E1 γονιδίου έδειξε ότι ρήξη του E1 γονιδίου εντοπίζεται σε δείγματα χαμηλού βαθμού (6/40, 15%), υψηλού βαθμού (8/33, 24.24%) ενδοεπιθηλιακής αλλοίωσης και σε περιπτώσεις καρκίνου του τραχήλου της μήτρας (7/11, 63.63%). Συμπερασματικά, μόνο η ρήξη του E1 γονιδίου εμπλέκεται σε δείγματα χαμηλού βαθμού αλλοίωσης (Fisher's exact test, $P < 0.05$), ενώ σε δείγματα υψηλού βαθμού αλλοίωσης και καρκίνου παρατηρείται ρήξη των E1 και E2 γονιδίων (Cricca et al., 2009). Επομένως, λαμβάνοντας υπόψη τη συχνότητα ρήξης του E1 γονιδίου σε δείγματα χαμηλού βαθμού αλλοίωσης, προτάθηκε ότι η περιοχή του γονιδίου E1 είναι ένα σημαντικό σημείο του γονιδιώματος του ιού, για τον περαιτέρω προσδιορισμό της μορφής του γονιδιώματος του HPV16 σε χαμηλού βαθμού δυσπλασίες.

Η χαρτογράφηση του γονιδίου E1 μεταξύ δειγμάτων υψηλού βαθμού, χαμηλού βαθμού αλλοίωσης και καρκίνου, οδήγησε στο συμπέρασμα ότι ο προσδιορισμός θέσεων ρήξης εντός του E1 γονιδίου, μπορεί να αποτελέσει ένα σημαντικό δείκτη ανίχνευσης της

ρήξης και ενσωμάτωσης του HPV16 DNA σε περιπτώσεις χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, γεγονός που μπορεί να συμβάλει στην αξιολόγηση της εξέλιξης της αλλοίωσης σε σοβαρότερου βαθμού δυσπλασία. Επιπλέον, η παρουσία ρήξης στο E2 γονίδιο μόνο σε δείγματα υψηλού βαθμού αλλοίωσης και καρκίνου, ενισχύει προηγούμενες μελέτες, οι οποίες συνέδεσαν τη ρήξη του E2 γονιδίου με υψηλού βαθμού αλλοιώσεις και καρκίνο του τραχήλου της μήτρας (Tonon et al., 2001, Cricca et al., 2007, Li et al., 2008, Cricca et al., 2009). Στο σημείο αυτό είναι σημαντικό να αναφερθεί ότι η διαφορά στη συχνότητα και στη θέση ρήξης των E1 και E2 γονιδίων ίσως να οφείλεται σε διαφορές, όσον αφορά τη τεχνική προσέγγιση του θέματος, καθώς επίσης και σε μεταλλάξεις που συσσωρεύονται στη περιοχή των γονιδίων E1 και E2. (Arias-Pulido et al., 2006). Με βάση την πληροφορία αυτή, η ευαισθησία της PCR προσαρμόστηκε στα δέκα αντίγραφα/αντίδραση, ενώ η νουκλεοτιδική ανάλυση των E1 και E2 γονιδίων έδειξε ότι οι θέσεις που υβριδίζονται οι εκκινητές και χρησιμοποιούνται για την ποσοτικοποίηση των E1 και E2, είναι αρκετά συντηρημένες μεταξύ των διαφορετικών στελεχών HPV16 που κυκλοφορούν στον Ελληνικό πληθυσμό. Συμπερασματικά, η νουκλεοτιδική ανάλυση των E1 και E2 γονιδίων σε συνδυασμό με τον προσδιορισμό των θέσεων ρήξης αποτελεί αναγκαία προϋπόθεση για τη περαιτέρω μελέτη της μορφής του γονιδιώματος του HPV16 σε διαφορετικούς πληθυσμούς.

Όπως αναφέρθηκε σε προηγούμενη ενότητα το γονιδίωμα του ιού μπορεί να παραμείνει ως επίσωμα σε χαμηλού βαθμού αλλοιώσεις, ενώ έχει παρατηρηθεί ότι η "επίμονη" HPV16 μόλυνση, σχετίζεται με αυξανόμενο κίνδυνο ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα, με αποτέλεσμα την πρόοδο της αλλοίωσης σε υψηλότερου βαθμού δυσπλασίες. (Hudelist et al., 2004, Andersson et al., 2005, Pett et al., 2006, Arias-Pulido et al., 2006, Cheung et al., 2006, De Marco et al., 2007, Cricca et al., 2007, Briolat et al., 2007, Huang et al., 2008, Li et al., 2008, Matovina et al., 2009). Μέχρι στιγμής έχει παρατηρηθεί ότι η ενσωμάτωση του ιού στο κυτταρικό γονιδίωμα γίνεται πιο συχνά μέσω ρήξης του E2 γονιδίου (Chen et al., 1994, Kalantari et al., 1998, Lukaszuk et al., 2003, Arias-Pulido et al., 2006, Cricca et al., 2009) και ο προσδιορισμός της μορφής του γονιδιώματος του HPV16 έχει επικεντρωθεί κυρίως μέσω ανάλυσης του λόγου των τιμών E2/E6 (Andersson et al., 2005, Arias Pulido et al., 2006, Kulmala et al., 2006, Cricca et al., 2007, Briolat et al., 2007, Boulet et al., 2009, Li et al., 2011, Manawapat et al., 2012, Theelen et al., 2013). Το γεγονός αυτό οδηγεί στην ανίχνευση γεγονότων ενσωμάτωσης του ιού μόνο μέσω ρήξης του γονιδίου E2, χωρίς να προσδιορίζεται η συνέχεια του E1 γονιδίου. Στη παρούσα διδακτορική

διατριβή, λαμβάνοντας υπόψη τη συχνότητα ρήξης των E1 και E2 γονιδίων, πραγματοποιήθηκε ταυτόχρονος προσδιορισμός του λόγου των τιμών E1/E6 και E2/E6 στα ίδια κλινικά δείγματα, με σκοπό να συγκριθούν οι δύο τιμές και να διεξαχθούν συμπεράσματα σχετικά με τη μορφή του γονιδιώματος του ιού και για τον τρόπο ενσωμάτωσης του HPV16 στο κυτταρικό γονιδίωμα. Επομένως, με βάση τα αποτελέσματα της χαρτογράφησης των γονιδίων E1, E2 και της νουκλεοτιδικής ανάλυσης των E1, E2 και E6 γονιδίων αντίστοιχα, ο προσδιορισμός E1/E6, E2/E6 πραγματοποιήθηκε με τη χρήση εκκινητικών μορίων, που υβριδίζονται σε συντηρημένες περιοχές του ιού και στις πιο συχνά εμφανιζόμενες θέσεις ρήξης του HPV16 DNA (E1: 1059 - 1323), (E2: 3172 - 3649), όπως προσδιορίστηκε από τη μελέτη των διαφορετικών στελεχών HPV16 που κυκλοφορούν στον Ελληνικό πληθυσμό.

Η ανάλυση του λόγου των τιμών E2/E6 και E1/E6, έδειξε ότι η επισωματική μορφή του ιού βρίσκεται με μεγαλύτερη συχνότητα σε χαμηλού βαθμού δυσπλασίες, ενώ η μεικτή μορφή του γονιδιώματος του ιού βρίσκεται με μεγαλύτερη συχνότητα σε υψηλού βαθμού ενδοεπιθηλιακές αλλοιώσεις. Τέλος, οι λόγοι E1/E6 και E2/E6, έδειξαν ότι η ενσωματωμένη μορφή του γονιδιώματος του HPV16 βρίσκεται με μεγαλύτερη συχνότητα σε περιπτώσεις καρκίνου του τραχήλου της μήτρας (Andersson et al., 2005, Kulmala et al., 2006). Επιπλέον, με βάση τα αποτελέσματα τα οποία προέκυψαν από την ανάλυση του λόγου E1/E6 και E2/E6, παρατηρήθηκε ότι σε δείγματα χαμηλού βαθμού αλλοίωσης εντοπίστηκε η καθαρά ενσωματωμένη μορφή του γονιδιώματος του ιού στο κυτταρικό γονιδίωμα. Το γεγονός αυτό σε συνδυασμό με την μεγάλη συχνότητα εμφάνισης της μεικτής μορφής του γονιδιώματος του HPV16 σε περιπτώσεις χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης οδήγησε στο συμπέρασμα, ότι η ενσωμάτωση του ιού συμβαίνει σε χαμηλού βαθμού δυσπλασίες (Cricca et al., 2007, Briolat et al., 2007, Boulet et al., 2009), υποστηρίζοντας το γεγονός ότι η ενσωμάτωση του HPV16 DNA είναι ένα φαινόμενο το οποίο συμβαίνει νωρίς στο κύκλο ζωής του ιού και μπορεί να αποτελέσει ένα σημαντικό προγνωστικό δείκτη εξέλιξης της ενδοεπιθηλιακής αλλοίωσης σε υψηλότερου βαθμού δυσπλασίες και καρκίνο του τραχήλου της μήτρας (Peitsaro et al., 2002, Gallo et al., 2003, Andersson et al., 2005, Kulmala et al., 2006, Cricca et al., 2007, Briolat et al., 2007, De Marco et al., 2007, Huang et al., 2008, Azizi et al., 2008).

Η σύγκριση των δύο λόγων των τιμών E1/E6 και E2/E6 στα δείγματα που εξετάστηκαν, καθόρισαν τον τρόπο ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα. Πιο συγκεκριμένα, παρατηρήθηκαν περιπτώσεις στις οποίες προσδιορίστηκε η συνύπαρξη τριών

διαφορετικών "πληθυσμών" HPV16 DNA. Πιο συγκεκριμένα, ένας πληθυσμός που φέρει την επισωματική μορφή του ιού, ένας πληθυσμός που έχει υποστεί ενσωμάτωση μέσω ρήξης στο E1 γονίδιο και ένας πληθυσμός που έχει υποστεί ρήξη στο E2 γονίδιο, αντίστοιχα. Το γεγονός αυτό ίσως να αποτελεί σημαντική ένδειξη για την ύπαρξη πολλαπλών θέσεων ενσωμάτωσης του ιού στο κυτταρικό DNA. Επιπλέον, παρατηρήθηκαν γεγονότα ενσωμάτωσης του HPV16 DNA, μέσω έλλειψης ολόκληρης της περιοχής μεταξύ των γονιδίων E1 και E2, ενώ παρατηρήθηκαν περιπτώσεις, που ο ιός βρισκόταν σε ενσωματωμένη μορφή αποκλειστικά μέσω ρήξης του E1 ή του E2 γονιδίου. Επομένως από την ανάλυση των αποτελεσμάτων λαμβάνονται σημαντικές πληροφορίες για τον τρόπο ενσωμάτωσης του ιού στο κυτταρικό DNA, γεγονός που δίνει σημαντικότερες ενδείξεις για την ογκογονικότητα του HPV16 DNA και την εξέλιξη της δυσπλασίας. Αντιθέτως, σε επτά περιπτώσεις παρουσιάστηκε σημαντική απόκλιση μεταξύ των τιμών E1/E6 και E2/E6. Το γεγονός αυτό ίσως να οφείλεται στη συσσώρευση μεταλλάξεων στις περιοχές που υβριδίζονται οι εκκινητές, ή σε αναδιατάξεις του HPV16 DNA στα συγκεκριμένα HPV16 στελέχη, με αποτέλεσμα να παρατηρείται σημαντική απόκλιση μεταξύ των δύο λόγων των τιμών E1/E6 και E2/E6 (Tornesello et al., 2000, Arias Pulido et al., 2006). Ακόμη, η απόκλιση των τιμών E1/E6 και E2/E6 θα μπορούσε να οφείλεται στην ύπαρξη διαφορετικών πληθυσμών HPV16 DNA που παρουσιάζουν διαφορετικά πρότυπα ενσωμάτωσης στο κυτταρικό γονιδίωμα. Παρόλα αυτά, θα πρέπει να γίνουν επιπρόσθετες μελέτες για τα μοτίβα ενσωμάτωσης του ιού στο κυτταρικό γονιδίωμα, έτσι ώστε να ενισχυθεί η παραπάνω υπόθεση.

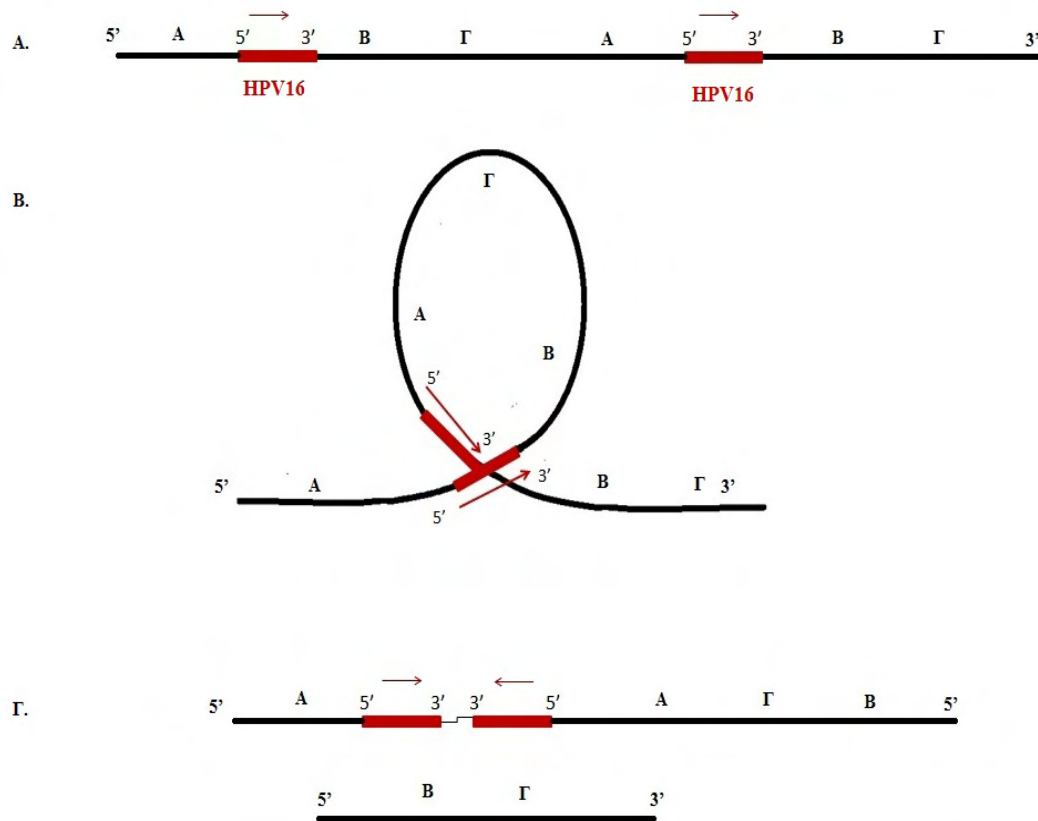
Συμπερασματικά, μέχρι στιγμής για τον προσδιορισμό της μορφής του γονιδιώματος του ιού πραγματοποιείται μόνο ανάλυση του λόγου E2/E6. Η μελέτη αυτή, είναι μια πρώτη προσέγγιση ταυτόχρονης ανάλυσης του λόγου των τιμών E1/E6 και E2/E6, έτσι ώστε να προσδιοριστεί με μεγαλύτερη ακρίβεια η μορφή του γονιδιώματος του HPV16 και ο τρόπος με τον οποίο ο ιός έχει ενσωματωθεί στο κυτταρικό DNA. Με βάση τα αποτελέσματα που προέκυψαν από τη συγκεκριμένη μελέτη, προτείνεται ότι θα πρέπει να λαμβάνεται σοβαρά υπόψη η επιλογή κατάλληλων στόχων για την ποσοτικοποίηση των E1, E2 και E6 γονιδίων, σε ένα δεδομένο πληθυσμό. Επιπρόσθετα, ο προσδιορισμός του τρόπου με τον οποίο ο ιός έχει ενσωματωθεί στο κυτταρικό γονιδίωμα, ίσως να μπορεί να δώσει σημαντικές πληροφορίες για την παθογένεια του HPV16, καθιστώντας τους λόγους των τιμών E1/E6 και E2/E6, ένα σημαντικό εργαλείο πρόγνωσης της εξέλιξης της ενδοεπιθηλιακής νεοπλασίας.

4.7 Φαινόμενα αναδιάταξης του HPV16 DNA

Στο τελευταίο μέρος της διδακτορικής διατριβής, έγινε προσπάθεια εντοπισμού των θέσεων ενσωμάτωσης του HPV16 σε κλινικά δείγματα, στα οποία προσδιορίστηκε ότι το γονιδίωμα του HPV16 βρίσκεται ενσωματωμένο στο DNA του κυττάρου ξενιστή. Η διαδικασία προσδιορισμού των θέσεων ενσωμάτωσης πραγματοποιήθηκε με τη μέθοδο RS-PCR, σύμφωνα με την οποία εκκινητές οι οποίοι αναγνωρίζουν συγκεκριμένες θέσεις στο γονιδίωμα του HPV16 (E1, E2, L1, L2 και LCR) συνδυάζονται στην αντίδραση με εκκινητές που αποκαλούνται εκκινητές αναγνώρισης θέσεων περιορισμού (Restriction site oligonucleotide primers, RSO). Με βάση τη μέθοδο αυτή ενισχύεται η περιοχή του γονιδιώματος του HPV16 που έχει ενσωματωθεί σε άγνωστη περιοχή του γονιδιώματος του κυττάρου ξενιστή (Yu et al., 2005)

Μετά την ανάλυση των αποτελεσμάτων εντοπίστηκαν σε τρεις περιπτώσεις καρκίνου, μία περίπτωση υψηλού βαθμού ενδοεπιθηλιακής αλλοίωσης και σε δύο περιπτώσεις χαμηλού βαθμού ενδοεπιθηλιακής αλλοίωσης, αναδιατάξεις του HPV16 DNA με αναστροφή. Πιο συγκεκριμένα, στο δείγμα PNL7 (καρκίνος του τραχήλου της μήτρας) εντοπίστηκε μια αλληλουχία HPV16, η οποία περιλαμβάνει τμήμα του E2 γονιδίου (2965 - 3337) με προσανατολισμό 5' - 3' και τμήμα του L1 γονιδίου (6418 έως 6151) με προσανατολισμό 3' - 5'. Επιπλέον στα δείγματα D59, PNL16 (καρκίνος του τραχήλου της μήτρας), ATT21 (υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση), ATT6, ATT78, ATT91, ATT98 (χαμηλού βαθμού ενδοεπιθηλιακή αλλοίωση) εντοπίστηκε ένα διαφορετικό πρότυπο αναδιάταξης με αναστροφή του HPV16 DNA, το οποίο περιλαμβάνει τμήμα του γονιδίου E7 (790 -858), τμήμα του E1 γονιδίου (865 - 1059) με προσανατολισμό 5' - 3' και στη συνέχεια τμήμα του L1 γονιδίου (6207 - 6151) με προσανατολισμό 3' - 5'. Φαινόμενα αναδιάταξης του γονιδιώματος του ιού έχουν αναφερθεί και από προηγούμενη εργασία του Arias Pulido και των συνεργατών του (2006), στην οποία περιγράφονται περιπτώσεις αναδιάταξης του HPV16 DNA στη περιοχή των E1 και E2 γονιδίων σε περιπτώσεις καρκίνου του τραχήλου της μήτρας. Στη συγκεκριμένη εργασία παρατηρήθηκαν ελλείψεις ή διπλασιασμοί εντός των γονιδίων, ενώ δεν καταγράφηκαν γεγονότα συνένωσης διαφορετικών τμημάτων του γονιδιώματος του ιού (Arias Pulido et al., 2006). Επιπλέον σε μια πρόσφατη εργασία, από την ανάλυση ολόκληρου του γονιδιώματος της κυτταρικής σειράς Caski, εντοπίστηκαν πολλαπλά γεγονότα αναδιάταξης σε ενσωματωμένα ιικά μόρια DNA, δημιουργώντας τμήματα με αντίθετο προσανατολισμό (Akagi et al., 2014). Μεταξύ των γεγονότων αναδιάταξης του HPV16 DNA περιγράφεται και η σύνδεση τμημάτων με αντίθετο

προσανατολισμό μεταξύ των γονιδίων E7,E1 - L1 και E2 - L2 (Akagi et al., 2014). Μέχρι στιγμής η ανάλυση τέτοιων περιπτώσεων αναδιάταξης ενσωματωμένων γονιδιωμάτων του HPV16, έχει περιγραφεί μόνο στη κυτταρική σειρά Caski, ενώ δεν έχουν αναλυθεί κλινικά δείγματα. Επιπλέον, από την ίδια μελέτη προτάθηκε ένα μοντέλο θηλιάς, σύμφωνα με το οποίο η αντιγραφή των ενσωματωμένων αντιγράφων του γονιδιώματος του ιού και ο ανασυνδυασμός, συμβάλουν στο σχηματισμό αλυσομερών ιικού και κυτταρικού DNA, με αποτέλεσμα την έλλειψη γονιδίων αλλά την έκφραση των ογκογονιδίων E6 και E7 (Akagi et al., 2014). Με βάση αυτή τη πληροφορία, ένας πιθανός τρόπος σύμφωνα με τον οποίο θα μπορούσαν να προκύψουν οι αναδιατάξεις του HPV16 DNA που εντοπίστηκαν στη παρούσα διατριβή, είναι η δημιουργία μιας δομής θηλιάς, η οποία περιλαμβάνει αλυσομερή HPV16 DNA, τα οποία έρχονται σε αντιπαράλληλο προσανατολισμό και ανασυνδυάζονται μεταξύ τους (Εικόνα 4.1). Το φαινόμενο αυτό ίσως έχει ως αποτέλεσμα τη συνένωση διαφορετικών τμημάτων HPV16 DNA με αντίθετο προσανατολισμό, την αναστροφή τμήματος του χρωμοσώματος και την απαλοιφή περιοχών του κυτταρικού DNA (Εικόνα 4.1). Εάν το φαινόμενο αυτό συμβαίνει τότε συμβάλει στην αποσταθεροποίηση του χρωμοσώματος του ξενιστή και τη πρόκληση καρκίνου. Επομένως, ακόμη και μετά την ενσωμάτωση του ιού στο κυτταρικό χρωμόσωμα ο ιός συνεχίζει να προκαλεί διαρκώς φαινόμενα αποσταθεροποίησης του κυτταρικού DNA. Κατά συνέπεια, είναι σημαντικό να ελεγχθεί η περιοχή αριστερά και δεξιά των ενσωματωμένων αλληλουχιών έτσι ώστε να επαληθευθεί η παραπάνω υπόθεση.



Εικόνα 4.1. Πιθανό μοντέλο το οποίο συμβάλει στο σχηματισμό τμημάτων HPV16 DNA με αντίθετο προσανατολισμό. Α) Αλυσομερή κυτταρικού και HPV16 DNA, όπως προτείνεται από το μοντέλο που περιγράφεται σε προηγούμενη εργασία (Akagi et al., 2014). Β) Σχηματισμός δομής θηλιάς, η οποία φέρει σε αντιπαράλληλο προσανατολισμό διαφορετικά μόρια HPV16 DNA, τα οποία στη συνέχεια ανασυνδυάζονται μεταξύ τους. Γ. Σχηματισμός τμημάτων HPV16 DNA, με αντίθετο προσανατολισμό, αναστροφή τμημάτων του χρωμοσώματος και έλλειψη περιοχών του χρωμοσώματος.

Επιπρόσθετα στη περίπτωση του δείγματος D59, εντοπίστηκε μια αλληλουχία η οποία περιλαμβάνει τμήμα των γονιδίων E7 και E1, ενώ η συνέχεια της αλληλουχίας περιλαμβάνει ακόμα σαράντα τρεις βάσεις. Παρόλα αυτά, η αλληλουχία των σαράντα τριών βάσεων δεν μπορεί να ταυτοποιηθεί στη βάση δεδομένων NCBI. Για το σκοπό αυτό το πείραμα βρίσκεται σε εξέλιξη σχεδιάζοντας έναν εκκινητή, ο οποίος υβριδίζεται ακριβώς στη θέση σύνδεσης της ικής και άγνωστης αλληλουχίας, έτσι ώστε να συνδυαστεί με τους εκκινητές αναγνώρισης θέσεων περιορισμού (Restriction site oligonucleotide primers, RSO).

Τέλος, όπως αναφέρθηκε σε προηγούμενη ενότητα σε επτά περιπτώσεις παρουσιάστηκε σημαντική απόκλιση μεταξύ των τιμών E1/E6 και E2/E6. Στα δείγματα αυτά, ανιχνεύτηκε σε τέσσερις περιπτώσεις (ATT21, ATT78, ATT98, PNL7), ανακατανομή του γονιδιώματος του ιού, ενώ τα υπόλοιπα δείγματα δεν έχουν εξεταστεί μέχρι στιγμής και το πείραμα βρίσκεται ακόμα σε εξέλιξη. Τα αποτελέσματα αυτά ενισχύουν την υπόθεση ότι

η ασυμφωνία των λόγων E1/E6 και E2/E6 ίσως να οφείλεται σε φαινόμενα αναδιάταξης του HPV16 DNA. Παρόλα αυτά ο τρόπος με τον οποίο έχουν δημιουργηθεί οι συγκεκριμένες ανακατανομές στο γονιδίωμα του ιού πρέπει να μελετηθούν περαιτέρω, γεγονός που ίσως δώσει πληροφορίες για τον τρόπο και τον μηχανισμό ενσωμάτωσης του HPV16 στο κυτταρικό χρωμόσωμα.

Συμπερασματικά, έχει προταθεί ότι ο προσδιορισμός της μορφής του γονιδιώματος του HPV16 είναι ένας σημαντικός μοριακός δείκτης, που μπορεί να προβλέψει την εξέλιξη της ενδοεπιθηλιακής αλλοίωσης από χαμηλού βαθμού σε υψηλού βαθμού ενδοεπιθηλιακή αλλοίωση και κατ' επέκταση σε καρκίνο του τραχήλου της μήτρας (Hudelist et al., 2004, Arias Pulido et al., 2006). Παρόλα αυτά είναι η πρώτη φορά που πραγματοποιείται ταυτόχρονη ανάλυση του λόγου των τιμών E1/E6 και E2/E6, μετά από προσδιορισμό των συνηθέστερων θέσεων ρήξης των HPV16 στελεχών που κυκλοφορούν σε ένα δεδομένο πληθυσμό. Ο προσδιορισμός E1/E6 και E2/E6, μπορεί να δώσει σημαντικές πληροφορίες για τη μορφή του γονιδιώματος του HPV16 σε κλινικά δείγματα. Με βάση τα αποτελέσματα που προέκυψαν από τη συγκεκριμένη διατριβή, προτείνεται ότι η χαρτογράφηση των συνηθέστερων θέσεων ρήξης των E1 και E2 γονιδίων σε συνδυασμό με τη νουκλεοτιδική ανάλυση των γονιδίων E1, E2 και E6 του HPV16 και τον προσδιορισμό γεγονότων αναδιάταξης του HPV16 DNA θα πρέπει να λαμβάνεται σοβαρά υπόψη για τον προσδιορισμό της μορφής του γονιδιώματος των διαφορετικών κυκλοφορούντων HPV16 στελεχών σε ένα δεδομένο πληθυσμό.

Βιβλιογραφία

- Akagi K, Li J, Broutian TR, Padilla-Nash H, Xiao W, Jiang B, Rocco JW, Teknos TN, Kumar B, Wangsa D, He D, Ried T, Symer DE, Gillison ML (2014) Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res*: doi 10.1101/gr.164806.113
- Akashi H, Osada N, Ohta T (2012) Weak selection and protein evolution. *Genetics*. 192:15-31
- Altschul SF, Gish W, Miller W, Myers EW, and Lipman D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Amin AA, Titolo S, Pelletier A, Fink D, Cordingley MG, Archambault J (2000) Identification of domains of the HPV11 E1 protein required for DNA replication in vitro. *Virology* 272:137-150
- Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B (2005) Type distribution, viral load and integration state of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer* 92:2195-2200
- Andersson S, Alemi M, Rylander E, Strand A, Larsson B, Sällström J, Wilander E (2000) Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. *Br J Cancer* 83:307-310
- Angulo M, Carvajal-Rodríguez A (2007) Evidence of recombination within human alpha-papillomavirus. *Virol J* 4:33
- Archambault J, Melendy T (2013) Targeting human papillomavirus genome replication for antiviral drug discovery. *Antivir Ther* 2018:271-283
- Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM (2006) Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 44:1755-1762
- Artandi SE, Attardi LD (2005) Pathways connecting telomeres and p53 in senescence, apoptosis, and cancer. *Biochem Biophys Res Commun* 331:881-890
- Ashrafi GH, Haghshenas MR, Marchetti B, O'Brien PM, Campo MS (2005) E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I *Int J Cancer* 113:276–283
- Ashrafi GH, Haghshenas M, Marchetti B, Campo MS (2006) E5 protein of human papillomavirus 16 down regulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. *Int J Cancer* 119:2105–2112
- Azizi N, Brazete J, Hankins C, Money D et al (2008) Canadian Women's HIV Study Group. Influence of human papillomavirus type 16 (HPV-16) E2 polymorphism on quantification of HPV-16 episomal and integrated DNA in cervicovaginal lavages from women with cervical intraepithelial neoplasia. *J Gen Virol* 89:1716–1728

- Badaracco G, Venuti A, Sedati A, Marcante ML (2002) HPV16 and HPV18 in genital tumors: Significantly different levels of viral integration and correlation to tumor invasiveness. *J Med Virol* 67:574–582
- Barbaresi S, Cortese MS, Quinn J, Ashrafi GH, Graham SV, Campo MS (2010) Effects of human papillomavirus type 16 E5 deletion mutants on epithelial morphology: functional characterization of each transmembrane domain. *J Gen Virol* 91:521–530
- Barrow-Laing L, Chen W, Roman A (2010) Low- and high-risk human papillomavirus E7 proteins regulate p130 differently. *Virology* 400:233–239
- Berezutskaya E, Yu B, Morozov A, Raychaudhuri P, Bagchi S (1997) Differential regulation of the pocket domains of the retinoblastoma family proteins by the HPV16 E7 oncoprotein. *Cell Growth Differ.* 8:1277-1286
- Bergant Marusic M, Ozbun MA, Campos SK, Myers MP, Banks L (2012) Human Papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17. *Traffic* 13:455–467
- Bernard HU, Calleja-Macias IE, Dunn ST (2006) Genome variation of human papillomavirus types: phylogenetic and medical implications. *Int J Cancer* 118:1071-1076
- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70-79
- Bergvall M, Melendy T, Archambault J (2013) The E1 proteins. *Virology* 445:35-56
- Blachon S, Demeret C (2003) The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic. *Biochimie* 85:813-819
- Bogovac Z, Lunar MM, Kocjan BJ, Seme K, Jančar N, Poljak M (2011) Prevalence of HPV 16 genomic variant carrying a 63 bp duplicated sequence within the E1 gene in Slovenian women. *Acta Dermatovenerol Alp Panonica Adriat* 20:135-139
- Boner W, Taylor ER, Tsirimonaki E, et al (2002) A Functional interaction between the human papillomavirus 16 transcription/replication factor E2 and the DNA damage response protein TopBP1. *J Biol Chem* 277:22297-22303
- Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, et al. (2008) Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. *Vaccine* 26:K1–16
- Bossis I, Roden RB, Gambhira R, Yang R, Tagaya M, Howley PM, Meneses PI (2005) Interaction of tSNARE syntaxin 18 with the papillomavirus minor capsid protein mediates infection. *J Virol* 79:6723-6731
- Boulet G, Horvath C, Vanden Broeck D, Sahebali S, Bogers J (2007) Human papillomavirus: E6 and E7 oncogenes. *Int J Biochem Cell Biol* 39:2006-2011

- Boulet GA, Benoy IH, Depuydt CE, Horvath CA, Aerts M, Hens N, Vereecken AJ, Bogers JJ (2009) Human papillomavirus 16 load and E2/E6 ratio in HPV16-positive women: biomarkers for cervical intraepithelial neoplasia ≥ 2 in a liquid-based cytology setting? *Cancer Epidemiol Biomarkers Prev* 18:2992-2999
- Bousarghin L, Touzé A, Sizaret PY, Coursaget P (2003) Human papillomavirus types 16, 31, and 58 use different endocytosis pathways to enter cells. *J Virol* 77:3846-3850
- Bouvard V, Storey A, Pim D, Banks L (1994) Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes *EMBO J* 13:5451-5459
- Brady CS, Duggan-Keen MF, Davidson JA, Varley JM, Stern PL (1999) Human papillomavirus type 16 E6 variants in cervical carcinoma: relationship to host genetic factors and clinical parameters. *J Gen Virol* 80:3233-3240
- Bravo IG, Alonso A (2004) Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *J Virol* 78:13613-13626
- Bravo IG, de Sanjosé S, Gottschling M (2010) The clinical importance of understanding the evolution of papillomaviruses. *Trends in microbiology* 18:432-438
- Briolat J, Dalstein V, Saunier M et al. (2007) HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. *Int J Cancer* 121:2198-2204
- Brooke B, Dasgupta J, Chen XS (2007) Structure-based engineering of papillomavirus major capsid L1: controlling particle assembly. *Virol J* 4:3
- Buck CB, Pastrana DV, Lowy DR, Schiller JT (2004) Efficient intracellular assembly of papillomaviral vectors. *J Virol*. 78:751-757
- Bush RM (2001) Predicting adaptive evolution. *Nat Rev Genet* 2:387-392
- Carvajal-Rodríguez A (2008) Detecting recombination and diversifying selection in human alpha-papillomavirus. *Infect Genet Evol* 8:689-692
- Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25-36
- Casas L, Galvan SC, Ordoñez RM, Lopez N, Guido M, Berumen J (1999) Asian-american variants of human papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. *Int J Cancer* 83:449-455
- Chakrabarti O, Veeraraghavalu K, Tergaonkar V, Liu Y, Androphy EJ, Stanley MA, Krishna S (2004) Human papillomavirus type 16 E6 amino acid 83 variants enhance E6-mediated MAPK signaling and differentially regulate tumorigenesis by Notch signaling and oncogenic Ras *J Virol* 78:5934-5945

- Chan SY, Bernard HU, Ong CK, Chan SP, Hofmann B, Delius H (1992) Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 66:5714-5725
- Chen CM, Shyu MP, Au LC, Chu HW, Cheng WT, Choo KB (1994) Analysis of deletion of the integrated human papillomavirus 16 sequence in cervical cancer: a rapid multiplex polymerase chain reaction approach. *J Med Virol* 44:206-211
- Chen SL, Huang CH, Tsai TC, Lu KY, Tsao YP, (1996) The regulation mechanism of c-jun and jun B by human papillomavirus type 16 E5 oncoprotein. *Arch Virol* 141:791–800
- Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD (2005) Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 79:7014–7023
- Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, Sahasrabudde VV, Gravitt PE, Hsing AW, Burk RD (2011) Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One* 6:e20183
- Cheung JL, Lo KW, Cheung TH, Tang JW, Chan PK (2006) Viral load, E2 gene disruption status, and lineage of human papillomavirus type 16 infection in cervical neoplasia. *J Infect Dis* 194:1706-1712
- Clower RV, Hu Y, Melendy T (2006) Papillomavirus E2 protein interacts with and stimulates human topoisomerase I. *Virology* 348:13-18
- Collins SI, Constandinou-Williams C, Wen K, Young LS, Roberts S, Murray PG, Woodman CB (2009) Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer Res* 69:3828-3832
- Conger KL, Liu JS, Kuo SR, Chow LT, Wang TS (1999) Human papillomavirus DNA replication. Interactions between the viral E1 protein and two subunits of human dna polymerase alpha/primase. *J Biol Chem* 274:2696-2705
- Cornet I, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, Tommasino M, Clifford GM, IARC HPV Variant Study Group (2012) Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol* 86:6855-6861
- Coutlée F, Trottier H, Gagnon S, Koushik A, Richardson H, Roger M, Ferenczy AS, Franco EL (2009) Low-risk human papillomavirus type 6 DNA load and integration in cervical samples from women with squamous intraepithelial lesions. *J Clin Virol* 45:96-99
- Couturier J, Sastre-Garau X, Schneider-Maunoury S, Labib A, Orth G (1991) Integration of papillomavirus DNA near myc genes in genital carcinomas and its consequences for proto-oncogene expression. *J Virol* 65:4534-4538

- Cricca M, Morselli-Labate AM, Venturoli S et al (2007) Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions. *Gynecol Oncol* 106:549-557
- Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbini M (2009) Disruption of HPV16 E1 and E2 genes in precancerous cervical lesions. *J Virol Methods* 158:180-183
- Cubie HA (2013) Diseases associated with human papillomavirus infection. *Virology* 445:21-34
- Culp TD, Budgeon LR, Marinkovich MP, Meneguzzi G, Christensen ND (2006) Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells 1. *J Virol* 80:8940-8950
- Danos O, Katinka M, Yaniv M (1982) Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. *EMBO J* 1:231-236
- Davy CE, Jackson DJ, Wang Q, Raj K, Masterson PJ, Fenner NF et al (2002) Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16. *J Virol* 76:9806-9818
- Davy CE, Jackson DJ, Raj K, Peh WL, Southern SA, Das P, Sorathia R et al (2005) Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes. *J Virol* 79:3998-4011
- Davy C, McIntosh P, Jackson DJ, Sorathia R, Miell M, Wang Q, Khan J, Soneji Y, Doorbar J (2009) A novel interaction between the human papillomavirus type 16 E2 and E1--E4 proteins leads to stabilization of E2. *Virology* 394:266-275
- Day, PM, Baker CC, Lowy DR, Schiller JT (2004) Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proc Natl Acad Sci USA* 101:14252-14257
- De Marco L, Gillio-Tos A, Bonello L, Ghisetti V, Ronco G, Merletti F (2007) Detection of human papillomavirus type 16 integration in preneoplastic cervical lesions and confirmation by DIPS-PCR and sequencing. *J Clin Virol* 38:7-13
- De S, Stanley RJ, Lu C, Long R, Antani S, Thoma G, Zuna R (2013) A fusion-based approach for uterine cervical cancer histology image classification. *Comput Med Imaging Graph* 37:475-487
- de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE et al. (2010) Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 11:1048-1056
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, Zur Hauxen H (2004) Classification of papillomaviruses. *Virology* 324:17-27

- Demarest S, Martinez-Yamout M, Chung J, Chen H, Xu W, Dyson H, Evans R, Wright P (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* 415, 549–553
- Deng W, Lin BY, Jin G, et al. (2004) Cyclin/CDK regulates the nucleocytoplasmic localization of the human papillomavirus E1 DNA helicase. *J Virol* 78: 13954-13965
- DeFilippis VR, Ayala FJ, Villarreal LP (2002) Evidence of diversifying selection in human papillomavirus type 16 E6 but not E7 oncogenes. *J Mol Evol* 55:491–499
- DiMaio D, Petti LM (2013) The E5 proteins. *Virology* 445:99-114
- Dong X, Zhou W, Pfister H (2000) Presence of genetic rearrangements in E1/E2 regions of episomal HPV 16 isolates from cervical carcinomas. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi* 14: 5–8. [In Chinese]
- Dostatni N, Lambert PF, Sousa R, Ham J, Howley PM, Yaniv M (1991) The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes Dev* 5:1657-1671
- Doorbar J (2006) Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* 110:525–541
- Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, Stanley MA (2012) The biology and life-cycle of human papillomaviruses. *Vaccine* 5:F55-70
- Doorbar J (2013) The E4 protein; structure, function and patterns of expression *Virology* 445:80-98
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
- Duensing S, Münger K (2004) Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *Int J Cancer* 109:157-162
- Duensing A, Spardy N, Chatterjee P, Zheng L, Parry J, Cuevas R, et al (2009) Centrosome overduplication, chromosomal instability, and human papillomavirus oncoproteins. *Environ Mol Mutagen* 50:741–747
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797
- Ekstrom J, Forslund O, Dillner J (2010) Three novel papillomaviruses (HPV109, HPV112 and HPV114) and their presence in cutaneous and mucosal samples *Virology* 397:331–396
- Ekström J, Bzhalava D, Svenback D, Forslund O, Dillner J (2011) High throughput sequencing reveals diversity of Human Papillomaviruses in cutaneous lesions. *Int J Cancer* 129:2643–2650
- Eriksson A, Herron JR, Yamada T et al (1999) Human papillomavirus type 16 variant lineages characterised by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* 80:595–600

- Ferber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, McGovern R, et al. (2003) Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma. *Oncogene* 22:7233–7242
- Fehrmann F, Laimins LA (2003) Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene* 22:5201–5207
- Finnen RL, Erickson KD, Chen XS, Garcea RL (2003) Interactions between papillomavirus L1 and L2 capsid proteins. *J Virol.* 77:4818-4826
- Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, Vignat J, Ferlay J, Bray F, Plummer M, et al. (2012) Global burden of human papillomavirus and related diseases. *Vaccine* 30: F12-F23
- Forslund O (2007) Genetic diversity of cutaneous human papillomaviruses. *J Gen Virol* 88:2662–2669
- Fradet-Turcotte A, Moody C, Laimins LA, Archambault J. Nuclear export of human papillomavirus type 31 E1 is regulated by Cdk2 phosphorylation and required 84:11747-11760
- Fu L, Van Doorslaer K, Chen Z, Ristriani T, Masson M, Travé G, et al (2010) Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity. *PLoS One* e12816
- Funk JO, Waga S, Harry JB, Espling E, Stillman B, Galloway DA (1997) Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes Dev* 11:2090-2100
- Gallo G, Bibbo M, Bagella L et al (2003) Study of viral integration of HPV-16 in young patients with LSIL. *J Clin Pathol* 56:532–536
- Garcia-Vallve S, Alonso A, Bravo IG (2005) Papillomaviruses: different genes have different histories. *Trends Microbiol* 13:514–521
- Ghittoni R., Accardi R., Hasan U, Gheit T, Sylla B, Tommasino M (2010) The biological properties of E6 and E7 oncoproteins from human papillomaviruses. *Virus Genes* 40:1-13
- Giannoudis A, Duin M, Snijders PJ, Herrington CS (2001) Variation in the E2-binding domain of HPV 16 is associated with high-grade squamous intraepithelial lesions of the cervix. *Br J Cancer* 84:1058–1063
- Giri I, Danos O, Yaniv M (1985) Genomic Structure of the Cottontail Rabbit (Shope) Papillomavirus *PNAS* 82: 1580–1584
- Gissmann L, Boshart M, Dürst M, Ikenberg H, Wagner D, zur Hausen H (1984) Presence of human papillomavirus in genital tumors. *J Invest Dermatol* 83:26-28
- Glover TW (2006) Common fragile sites. *Cancer Lett* 232:4-12

- Gottschling M, Göker M, Köhler A, Lehmann MD, Stockfleth E, Nindl I (2009) Cutaneotropic human beta-/gamma-papillomaviruses are rarely shared between family members. *J Invest Dermatol* 129:2427–2434
- Gottschling M, Köhler A, Stockfleth E, Nindl I (2007) Phylogenetic analysis of beta-papillomaviruses as inferred from nucleotide and amino acid sequence data. *Mol Phylogenet Evol* 42:213-222
- Gottschling M, Göker M, Stamatakis A, Bininda-Emonds OR, Nindl I, Bravo IG (2011) Quantifying the phylodynamic forces driving papillomavirus evolution. *Mol Biol Evol* 28:2101–2113
- Gradíssimo Oliveira A, Delgado C, Verdasca N, Pista A (2013) Prognostic value of human papillomavirus types 16 and 18 DNA physical status in cervical intraepithelial neoplasia. *Clin Microbiol Infect* 19:E447-450
- Graham DA, Herrington CS (2000) HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and invasive squamous cell carcinomas of the cervix: relation to numerical chromosome abnormalities. *Mol Pathol* 53:201–206
- Grassmann K, Rapp B, Maschek H, Petry KU, Iftner T (1996) Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. *J Virol*. 70:2339-2349
- Grodzki M, Besson G, Clavel C, Arslan A, Franceschi S, Birembaut P, Tommasino M, Zehbe I (2006) Increased risk for cervical disease progression of French women infected with the human papillomavirus type 16 E6-350G variant. *Cancer Epidemiol Biomarkers Prev* 15:820-822
- Ham J, Steger G, Yaniv M (1994) Cooperativity in vivo between the E2 transactivator and the TATA box binding protein depends on core promoter structure. *EMBO J* 13:147 -157
- Hamid NA, Brown C, Gaston K (2009) The regulation of cell proliferation by the papillomavirus early proteins. *Cell Mol Life Sci* 66:1700-1717
- He L, Long LR, Antani S, Thoma GR (2012) Histology image analysis for carcinoma detection and grading. *Comput Methods Programs Biomed* 107:538-556
- Hebner CM, Laimins LA (2006) Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev Med Virol* 16: 83-97
- Hegde RS (2002) The papillomavirus E2 proteins: structure, function, and biology. *Annu Rev Biophys Biomol Struct* 31: 343-360
- Heise A (2003) The clinical significance of HPV *Nurse Pract* 28:8-19
- Herrick J, Conti C, Teissier S, et al. (2005) Genomic organization of amplified MYC genes suggests distinct mechanisms of amplification in tumorigenesis. *Cancer Res* 65:1174-1179

- Hoffmann R, Hirt B, Bechtold V, Beard P, Raj K (2006) Different modes of human papillomavirus DNA replication during maintenance J Virol. 80:4431-4439
- Horvath CA, Boulet GA, Renoux VM, Delvenne PO, Bogers JP (2010) Mechanisms of cell entry by human papillomaviruses: an overview. Virol J 7:11
- Hsueh PR (2009) Human papillomavirus, genital warts, and vaccines. J Microbiol Immunol Infect 42:101–106
- Huang LW, Choa SL, Lee BH (2008) Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. J Clin Pathol 61:627-631
- Hu L ,Ceresa BP (2009) Characterization of the plasma membrane localization and orientation of HPV16 E5 for cell–cell fusion. Virology 393:135–143
- Hu L, Plafker K, Vorozhko V, Zuna RE, Hanigan MH, Gorbsky GJ, Plafker SM, Angeletti PC, Ceresa BP (2009) Human papillomavirus 16 E5 induces bi- nucleated cell formation by cell-cell fusion. Virology 384:125–134
- Hudelist, G, Manavi M, Pischinger KID, Watkins-Riedel T, Singer CF, Kubista E, and Czerwenka KF (2004) Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. Gynecol Oncol 92:873–880
- Huertas-Salgado A, Martín-Gómez DC, Moreno P, Murillo R, Bravo MM, Villa L, Molano M (2011) E6 molecular variants of human papillomavirus (HPV) type 16: an updated and unified criterion for clustering and nomenclature. Virology 410:201-215
- Hurst LD (2009) Genetics and the understanding of selection. Nat Rev Genet 10:83-93.
- Hummel M, Hudson JB, Laimins LA (1992) Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes J Virol, 66: 6070–6080
- Ilves I, Maemets K, Silla T, Janikson K, Ustav M (2006) Brd4 is involved in multiple processes of the bovine papillomavirus type 1 life cycle. J Virol 80: 3660-3665
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. CA Cancer J Clin 61:69-90
- Jeon S ,Allen-Hoffmann BL, Lambert PF (1995) Integration of human papilloma- virus type 16 into the human genome correlates with a selective growth advantage of cells. J Virol 69:2989–2997
- Jeon S, Lambert PF (1995) Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. Proc Natl Acad Sci U S A 92:1654–1658
- Jiang M, Xi LF, Edelstein ZR, Galloway DA, Olsem GJ, Lin WC, Kiviat NB (2009) Identification of recombinant human papillomavirus type 16 variants. Virology 394:8-11

- Jiang M, Baseman JG, Koutsky LA, Feng Q, Mao C, Kiviat NB, Xi LF (2009) Sequence variation of human papillomavirus type 16 and measurement of viral integration by quantitative PCR. *J Clin Microbiol* 47:521–52643.
- Johannsen E, Lambert PF (2013) Epigenetics of human papillomaviruses. *Virology* 445:205-212
- Johnson KM, Kines RC, Roberts JN, Lowy DR, Schiller JT, Day PM (2009) Role of heparan sulfate in attachment to and infection of the murine female genital tract by human papillomavirus. *J Virol* 83:2067–2074
- Junes-Gill K, Sichero L, Maciag PC, Mello W, Noronha V, Villa LL (2008) Human papillomavirus type 16 variants in cervical cancer from an admixed population in Brazil. *J Med Virol* 80:1639-1645
- Kadaja M, Silla T, Ustav E, Ustav M (2009) Papillomavirus DNA replication – from initiation to genomic instability. *Virology* 384:360-368
- Kajitani N, Satsuka A, Kawate A, Sakai H (2012) Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation. *Front Microbiol* 3:152
- Kalantari M, Karlsen F, Kristensen G, Holm R, Hagmar B, Johansson B (1998) Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *Int J Gynecol Pathol* 17:146-153
- Kämper N, Day PM, Nowak T, Selinka HC, Florin L, Bolscher J, Hilbig L, Schiller J T, Sapp M (2006) A membrane-destabilizing peptide in capsid protein L2 is required for egress of papillomavirus genomes from endosomes. *J Virol* 80:759-768
- Kim SH, Oh JM, No JH, Bang YJ, Juhn YS, Song YS (2009) Involvement of NF-kappa B and AP-1 in COX-2 upregulation by human papillomavirus 16 E5 oncoprotein. *Carcinogenesis* 30:753–757
- Kivi N, Greco D, Auvinen P, Auvinen E (2008) Genes involved in cell adhesion, cell motility and mitogenic signaling are altered due to HPV16 E5 protein expression. *Oncogene* 27:2532–2541
- Korzeniewski N, Treat B, Duensing S (2011) The HPV-16 E7 oncoprotein induces centriole multiplication through deregulation of Polo-like kinase 4 expression. *Mol Cancer* 10:61
- Kosiol C, Bofkin L, Whelan S (2006) Phylogenetics by likelihood: evolutionary modeling as a tool for understanding the genome. *J Biomed Informatics* 39:51–61
- Krawczyk E, Suprynowicz FA, Sudarshan SR, Schlegel R (2010) Membrane orientation of the human papillomavirus type 16 E5 oncoprotein. *J Virol* 84:1696–1703
- Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS Genet* 4:e1000304

- Kulmala SM, Syrjänen SM, Gyllensten UB, Shabalova IP, Petrovichev N, Tosi P, Syrjänen KJ, Johansson BC (2006) Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 59:513-517
- Lee D, Lee B, Kim J, Kim DW, Choe J (2000) cAMP response element-binding protein binds to human papillomavirus E2 protein and activates E2-dependent transcription. *J Biol Chem* 275:7045-7051
- Lee K, Magalhaes I, Clavel C, Briolat J, Birembaut P, Tommasino M, Zehbe I (2008) Human papillomavirus 16 E6, L1, L2 and E2 gene variants in cervical lesion progression. *Virus Res* 131:106–110.
- Lehr E, Hohl D, Huber M, Brown D (2004) Infection with Human Papillomavirus alters expression of the small proline rich proteins 2 and 3. *J Med Virol.* 72:478-483
- Li H, Yang Y, Zhang R, Cai Y, Yang X, Wang Z, Li Y, Cheng X, Ye X, Xiang Y, Zhu B (2013) Preferential sites for the integration and disruption of human papillomavirus 16 in cervical lesions. *J Clin Virol* 56:342-347
- Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM (2011) Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer* 128:927-935
- Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D (2008) The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355-1361
- Liu J.S, Kuo SR, Broker TR, Chow LT (1995) The functions of human papillomavirus type 11 E1, E2, and E2C proteins in cell-free DNA replication. *J Biol Chem* 270:27283-272891
- Liu T, Lei Z, Pan Z, Chen Y, Li X, Mao T, He Q, Fan D (2013) Genetic association between p53 codon 72 polymorphism and risk of cutaneous squamous cell carcinoma. *Tumour Biol* [ahead of print]
- Longworth MS, Laimins LA (2004) Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol* 68:362–372
- Loo YM, Melendy T (2004) Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *J Virol* 78:1605-1615
- Lukaszuk K, Liss J, Wozniak I, Emerich J, Wójcikowski C (2003) Human papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex PCR. *J Clin Microbiol* 41:608-612
- Ma T, Zou N, Lin BY, Chow LT, Harper JW (1999) Interaction between cyclin-dependent kinases and human papillomavirus replication initiation protein E1 is required for efficient viral replication. *Proc Natl Acad Sci USA* 96: 382-387

- Maglennon GA, McIntosh P, Doorbar J (2011) Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology* 414:153–163
- Manawapat A, Stubenrauch F, Russ R, Munk C, Kjaer SK, Iftner T (2012) Physical state and viral load as predictive biomarkers for persistence and progression of HPV16-positive cervical lesions: results from a population based long-term prospective cohort study. *Am J Cancer Res* 2:192-203
- Masterson PJ, Stanley MA, Lewis AP, Romanos MA (1998) A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit. *J Virol* 72:7407-7419
- Matovina M, Sabol I, Grubisić G, Gasperov NM, Grce M (2009) Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecol Oncol* 113:120-127
- McBride AA, McPhillips MG, Oliveira JG (2004) Brd4: tethering, segregation and beyond. *Trends Microbiol* 12:527-529
- McBride AA (2008) Replication and partitioning of papillomavirus genomes. *Adv Vir Res* 72:155–205
- McIntosh PB, Laskey P, Sullivan K, Davy C, Wang Q, Jackson DJ, Griffin HM, Doorbar J (2010) E1-E4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium. *J Cell Sci* 123:2810-2822
- Meissner J (1997) Sequencing errors in reference HPV clones. In: Myers G, Munger K, Sverdrup F, McBride AA, Bernard HU, Meissner J (eds) *Human Papillomavirus 1997: a compilation and analysis of nucleic acid and amino acid sequences*. Los Alamos National Laboratory, NM, pp 110–123
- Meissner J (1999) Nucleotide sequences and further characterization of human papillomavirus DNA present in the CaSki, SiHa and HeLa cervical carcinoma cell lines. *J Gen Virol* 80:1725–1733
- Middleton K, Peh W, Southern S, Griffin H, Sotlar K, Nakahara T (2003) Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J Virol* 77:10186–10201
- Modis Y, Trus BL, Harrison SC (2002) Atomic model of the papillomavirus capsid. *EMBO J* 21:4754-4762
- Muller A, Ritzkowsky A, Steger G (2002) Cooperative activation of human papillomavirus type 8 gene expression by the E2 protein and the cellular coactivator p300. *J Virol* 76:11042-11053
- Munger K, Baldwin A, Edwards KM, Hayakawa H, Nguyen CL, Owens M. et al. (2004) Mechanisms of human papillomavirus-induced oncogenesis. *J Virol* 78:11451–11460

- Munoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, et al. (2003) Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 348:518–527
- Munoz N, Bosch FX, Castellsague X, Diaz M, de Sanjose S, Hammouda D et al. (2004) Against which human papillomavirus types shall we vaccinate and screen? The international perspective. *Int J Cancer* 111:278-285
- Munoz N, Castellsagué X, de González AB, Gissmann L (2006) Chapter 1: HPV in the etiology of human cancer. *Vaccine* 3:S3/1-10
- Narechania A, Chen Z, DeSalle R, Burk RD (2005) Phylogenetic incongruence among oncogenic genital alpha human papillomaviruses. *J Virol* 79:15503-15510
- Nath R, Mant CA, Kell B, Cason J, Bible JM (2006) Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. *Cancer Cell Int* 6:19
- Nguyen CL, Munger K (2008) Direct association of the HPV16 E7 oncoprotein with cyclin A/CDK2 and cyclin E/CDK2 complexes. *Virology*. 380:21–25
- Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197-218
- Nindl I, Gottschling M, Stockfleth E (2007) Human papillomaviruses and nonmelanoma skin cancer: basic virology and clinical manifestations. *Dis Markers* 23:247–259
- Nominé Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckere F, et al. (2006) Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell* 21:665–678
- Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, Sharma R, Batra S, Das BC (2008) Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J Clin Microbiol* 46:1060-1066
- Parish JL, Bean AM, Park RB, Androphy EJ (2006) ChlR1 is required for loading papillomavirus E2 onto mitotic chromosomes and viral genome maintenance. *Mol Cell* 24:867–76
- Parker LM, Harris S, Gossen M, Botchan MR (2000) The bovine papillomavirus E2 transactivator is stimulated by the E1 initiator through the E2 activation domain. *Virology* 270:430-443
- Peitsaro P, Johansson B, Syrjanen S (2002) Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *J Clin Microbiol* 40:886–891
- Peter M, Rosty C, Couturier J, Radvanyi F, Teshima H, et al. (2006) MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors. *Oncogene* 25:5985–5993

- Peter M, Stransky N, Couturier J, Hupe P, Barillot E, et al. (2010) Frequent genomic structural alterations at HPV insertion sites in cervical carcinoma. *J Pathol* 221: 320–330
- Peter W, Titolo S, Brault K et al. (2003) Inhibition of Human Papillomavirus DNA Replication by Small Molecule Antagonists of the E1-E2 Protein Interaction *Biol Chem Vol* 29:26765-26772
- Pett MR, Herdman MT, Palmer RD, Yeo GSH, Shivji MK, Stanley MA, Coleman N (2006) Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *Proc Natl Acad Sci USA* 103:3822–3827
- Pett M, Coleman N (2007) Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *J Pathol* 2007;212:356–367
- Pfister H (2003) Human papillomavirus and skin cancer. *J Natl Cancer Inst Monogr* 31:52–56
- Piccini A, Storey A, Massimi P, Banks L (1995) Mutations in the human papillomavirus type 16 E2 protein identify multiple regions of the protein involved in binding to E1. *J Gen Virol* 76:2909–2913
- Pim D, Banks L (2010) Interaction of viral oncoproteins with cellular target molecules: infection with high-risk vs low-risk human papillomaviruses. *APMIS* 118:471–493
- Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531-2533
- Puranen M, Saarikoski S, Syrjänen K, Syrjänen S (1996) Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391-395
- Pyeon D, Pearce SM, Lank SM, Ahlquist P, Lambert PF (2009) Establishment of human papillomavirus infection requires cell cycle progression. *PLoS Pathog* 5:e1000318
- Quint KD, de Koning MN, van Doorn LJ, Quint WG, Pirog EC (2010) HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix. *Gynecol Oncol* 117:297–301
- Raj K, Berguerand S, Southern S, Doorbar J, Beard P (2004) E1^{E4} protein of human papillomavirus type16 associates with mitochondria. *J Virol* 78:7199–7207
- Ray SC (1998) SimPlot for Windows (version 1.6). Baltimore, Md. Distributed by author (<http://www.welch.jhu.edu/~sray/download>)
- Reagan JW, Seidermann IL, Saracusa Y (1953) The cellular morphology of carcinoma in situ and dysplasia or atypical hyperplasia of the uterine cervix. *Cancer* 6:224–235
- Rehtanz M, Schmidt HM, Warthorst U, Steger G (2004) Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription. *Mol Cell Biol* 24:2153-2168

- Richart RM (1967) Natural history of cervical intraepithelial neoplasia. *Clinical Obstetrics and Gynecology* 10:748–784
- Richards RM, Lowy DR, Schiller JT, Day PM. (2006) Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. *Proc Natl Acad Sci USA* 103:1522-1527
- Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH (1994) Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J Virol* 68:6432-6445
- Roberts S, Ashmole I, Rookes SM, Gallimore PH (1997) Mutational analysis of the human papillomavirus type 16 E1-E4 protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments. *J Virol* 71:3554-3562
- Roberts S, Kingsbury SR, Stoeber K, Knight GL, Gallimore PH, Williams GH (2008) Identification of an arginine-rich motif in human papillomavirus type 1 E1/E4 protein necessary for E4-mediated inhibition of cellular DNA synthesis in vitro and in cells. *J Virol* 82:9056-9064
- Roman A (2005) The human papillomavirus E7 protein shines a spotlight on the pRB family member, p130. *Cell Cycle* 5:567–568
- Rous P, Kidd JG (1940) The activating, transforming, and carcinogenic effects of the rabbit papilloma virus (Shope) upon implanted tar tumors. *J Exp Med* 71:787–812
- Rous P, Beard JW (1935) The progression to carcinoma of virus-induced rabbit papilloma (Shope). *J Exp Med* 60:701-722
- Rous, P, Kidd JG, Smith WE (1952) Experiments on the cause of the rabbit carcinomas derived from virus-induced papillomas. II. Loss by the Vx2 carcinoma of the power to immunize hosts against the papilloma virus. *J Exp Med* 96:159-174
- Sabol I, Matovina M, Gasperov NM, Grce M (2008) Identification of a novel human papillomavirus type 16 E1 gene variant with potentially reduced oncogenicity. *J Med Virol* 80:2134-2140
- Sabol I, Matovina M, Si-Mohamed A, Grce M (2012) Characterization and whole genome analysis of human papillomavirus type 16 e1-1374^Δ63nt variants. *PLoS One* 7:e41045
- Sanchez GI, Kleter B, Gheit T, van Doorn LJ, de Koning MN, de Sanjosé S, Alemany L, Bosch XF, Tommasino M, Muñoz N, Quint WG (2011) Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *J Clin Virol* 51:165-169

- Sathish N, Abraham P, Peedicayil A, Sridharan G, Shaji RV, Chandy G (2004) E2 sequence variations of HPV 16 among patients with cervical neoplasia seen in the Indian subcontinent. *Gynecol Oncol* 95:363–369
- Schelhaas M, Shah B, Holzer M, Blattmann P, Kühling L, Day PM, et al. (2012) Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. *PLoS Pathog* 8:e1002657
- Schiffman M, Herrero R, Desalle R, Hildesheim A, Wacholder S, Rodriguez AC, et al. (2005) The carcinogenicity of human papillomavirus types reflects viral evolution. *Virology* 337:76–84
- Schiffman M, Clifford G, Buonaguro FM (2009) Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infect Agents Cancer* 4:8
- Schmitz M, Driesch C, Beer-Grondke K, Jansen L, Runnebaum IB, Dürst M (2012) Loss of gene function as a consequence of human papillomavirus DNA integration. *Int J Cancer* 131:E593-602
- Schmitz M, Driesch C, Jansen L, Runnebaum IB, Dürst M (2012) Non-random integration of the HPV genome in cervical cancer. *PLoS One* 7:e39632
- Schweiger MR, You J, Howley PM. Bromodomain protein 4 mediates the papillomavirus E2 transcriptional activation function. *J Virol* 80:4276-4285
- Sedman J, Stenlund A. (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J Virol* 72:6893-9897
- Shafti-Keramat S, Handisurya A, Kriehuber E, Meneguzzi G, Slupetzky K, Kirnbauer R (2003) Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. *J Virol* 77:13125-13135
- Shah SD, Doorbar J, Goldstein RA (2010) Analysis of host-parasite incongruence in papillomavirus evolution using importance sampling. *Mol Biol Evol* 27:1301–1314
- Shope RE, Hurst EW (1933) Infectious papillomatosis in rabbits. *J Exp Med* 58:607-624.
- Sichero L, Ferreira S, Trottier H, Duarte-Franco E, Ferenczy A, Franco EL, Villa LL (2007) High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *Int J Cancer*. 120:1763-1768
- Smith B, Chen Z, Reimers L, van Doorslaer K, Schiffman M, Desalle R, Herrero R, Yu K, Wacholder S, Wang T, Burk RD (2011) Sequence imputation of HPV16 genomes for genetic association studies. *PLoS One* 6:e21375
- Solomon D, Davey D, Kurman R, et al. (2002) The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 287:2114-2119
- Sotlar K, Diemer D, Dethleffs A, Hack Y et al. (2004) Detection and typing of human papillomavirus by E6 nested multiplex PCR. *J Clin Microbiol* 42:3176-3184

- Steger G, Corbach S (1997) Dose-dependent regulation of the early promoter of HPV 18 by the viral E2 protein. *J Virol* 71:50-58
- Stenlund A (2003) E1 initiator DNA binding specificity is unmasked by selective inhibition of non-specific DNA binding. *EMBO J* 22:954-963
- Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T (2007) Selecton: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Research* 35:W506-W511
- Stoppler MC, Straight SW, Tsao G, Schlegel R, McCance DJ (1996) The E5 gene of HPV-16 enhances keratinocyte immortalization by full-length DNA. *Virology* 223:251-254
- Straight SW, Hinkle PM, Jewers RJ, McCance DJ (1993) The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the down regulation of the epidermal growth factor receptor in keratinocytes. *J Virol* 67:4521-4532
- Surviladze Z, Dziduszko A, Ozbun MA (2012) Essential roles for soluble virion-associated heparan sulfonated proteoglycans and growth factors in human papillomavirus infections. *PLoS Pathog* 8:e1002519
- Suzuki Y, Gojobori T (1999) A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* 16:1315-1328
- Swan DC, Rajeevan M, Tortolero-Luna G, Follen M, Tucker RA, Unger ER (2005) Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* 96:695-700
- Swindle CS, Engler JA (1998) Association of the human papillomavirus type 11 E1 protein with histone H1. *J Virol* 72:1994-2001
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-1599
- Tan SH, Baker CC, Stünkel W, Bernard HU (2002) A transcriptional initiator overlaps with a conserved YY1 binding site in the long control region of human papillomavirus type 16. *Virology* 305:486-501
- Theelen W, Litjens RJ, Vinokurova S, Haesevoets A, Reijans M, Simons G, Smedts F, Herrington CS, Ramaekers FC, von Knebel Doeberitz M, Speel EJ, Hopman AH (2013) Human papillomavirus multiplex ligation-dependent probe amplification assay for the assessment of viral load, integration, and gain of telomerase-related genes in cervical malignancies. *Hum Pathol* 2013 44:2410-2418.
- Thiery F (2009) Transcription regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology* 384:375-379
- Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, et al. (2000) Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. *Cancer Res* 60:5916-5921

- Thorland EC, Myers SL, Gostout BS, Smith DI (2003) Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* 22:1225-1237
- Titolo S, Pelletier A, Pulichino AM, Brault K, Wardrop E, White PW, Cordingley MG, Archambault J (2000) Identification of domains of the human papillomavirus type 11 E1 helicase involved in oligomerization and binding to the viral origin. *J Virol* 74:7349-7361
- Titolo S, Brault K, Majewski J, White PW, Archambault J (2003) Characterization of the minimal DNA binding domain of the human papillomavirus e1 helicase: fluorescence anisotropy studies and characterization of a dimerization-defective mutant protein. *J Virol* 77:5178-5191
- Tonon SA, Picconi MA, Bos PD, Zinovich JB, Galuppo J, Alonio LV, Teyssie A (2001) Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J Clin Virol* 21:129-134
- Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Greggi S, Buonaguro L, Buonaguro FM. (2011) Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 1:32-42
- Van Regenmortel MH, Fauquet CM, Bishop DH, Calisher CH, Carsten EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (2002) Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses. Academic Press, New-York, San Diego
- Vandepoele K, Van Roy N, Staes K, Speleman F, van Roy F (2005) A novel gene family NBPF: intricate structure generated by gene duplications during primate evolution. *Mol Biol Evol* 22:2265-2274
- Varsani A, van der Walt E, Heath L, Rybicki EP, Williamson AL, Martin DP (2006) Evidence of ancient papillomavirus recombination. *J Gen Virol* 87:2527-31
- Veress G, Szarka K, Dong XP, Gergely L, Pfister H (1999) Functional significance of sequence variation in the E2 gene and the long control region. *J Gen Virol* 80:1035-1043
- Vinokurova S, Wentzensen N, Kraus I, Klaes R, Driesch C, Melsheimer P, et al. (2008) Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. *Cancer Res* 68:307-313
- Vinokurova S, von Knebel Doeberitz M (2011) Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation. *PloS one* 6: e24451
- Wang J, Sampath A, Raychaudhuri P, Bagchi S (2001) Both Rb and E7 are regulated by the ubiquitin proteasome pathway in HPV-containing cervical tumor cells. *Oncogene* 20:4740-4749

- Wang L, Dai SZ, Chu HJ, Cui HF, Xu XY (2013) Integration sites and genotype distributions of human papillomavirus in cervical intraepithelial neoplasia. *Asian Pac J Cancer Prev* 14:3837-3841
- Wang Q, Griffin H, Southern S, Jackson D, Martin A et al. (2004) Functional analysis of the human papillomavirus type 16 E1^{E4} protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J Virol* 78:821-833
- Watts KJ, Thompson CH, Cossart YE, Rose BR (2002) Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. *Int J Cancer* 97:868–874
- Wentzensen N, Vinokurova S, von Knebel Doeberitz M (2004) Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res* 64:3878–3884
- Wetherill LF, Holmes KK, Verow M, Muller M, Howell G, Harris M, Fishwick C, Stonehouse N, Foster R, Blair GE, Griffin S, Macdonald A (2012) High- risk human papillomavirus E5 oncoprotein displays channel-forming activity sensitivetosmall-moleculainhibitors *J Virol* 86:5341-5351
- Wheeler CM, Yamada T, Hildesheim A, Jenison SA (1997) Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* 35:11-19
- White PW, Pelletier A, Brault K, Titolo S, Welchner E, Thauvette L, Fazekas M, Cordingley MG, Archambault J (2001) Characterization of recombinant HPV6 and 11 E1 helicases: effect of ATP on the interaction of E1 with E2 and mapping of a minimal helicase domain. *J Biol Chem* 276:22426-22438
- Wickner, RB *Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses.* Academic Press, New-York, San Diego. 2002.
- Wilson VG, West M, Woytek K, Rangasamy D (2002) Papillomavirus E1 proteins: form, function, and features. *Virus Genes* 24:275-290
- Winder DM, Pett MR, Foster N, et al. (2007) An increase in DNA double-strand breaks, induced by Ku70 depletion, is associated with human papillomavirus 16 episome loss and de novo viral integration events. *J Pathol* 213: 27-34
- Woodman CB, Collins S, Rollason TP, Winter H, Bailey A, Yates M, et al. (2003) Human papillomavirus type 18 and rapidly progressing cervical intraepithelial neoplasia. *Lancet* 361:40-43
- Woodman CB, Collins SI, Young LS (2007) The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer* 7:11-22
- Woolhouse M, Gaunt E (2007) Ecological origins of novel human pathogens. *Crit Rev Microbiol* 33:231-42

- Wu Y, Chen Y, Li L, Yu G, He Y, Zhang Y (2006) Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China. *J Gen Virol* 87:1181–1188
- Xu B, Chotewutmontri S, Wolf S, Klos U, Schmitz M, Durst M, Schwarz E (2013) Multiplex identification of human papillomavirus 16 DNA integration sites in cervical carcinomas. *PLoS ONE* 8:e66693
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743-7753
- Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX (1997) Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol*. 71:2463-2472
- Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431-449
- Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19:908-917
- Yu T, Ferber MJ, Cheung TH, Chung TK, Wong YF, Smith DI (2005) The role of viral integration in the development of cervical cancer. *Cancer Genet Cytogenet* 158:27-34
- Zanier K,ould M'hamed ould Sidi A, Boulade-Ladame C, Rybin V, Chappelle A, Atkinson A, et al. (2012) Solution structure analysis of the HPV16 E6 oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53. *Structure* 20:604–617
- Zehbe I, Mytilineos J, Wikstrom I, Henriksen R, Edler L, Tommasino M (2003) Association between human papillomavirus 16 E6 variants and human leukocyte antigen class I polymorphism in cervical cancer of Swedish women. *Hum Immunol* 64:538–542
- Zhang Y, Lehman JM, Petti LM (2002) Apoptosis of mortal human fibroblasts transformed by the bovine papillomavirus E5 oncoprotein. *Mol Cancer Res* 1:122-136
- Zheng ZM, Baker CC (2006) Papillomavirus genome structure expression and post transcriptional regulation. *Front Biosci* 11:2286-2302
- Ziegert C, Wentzensen N, Vinokurova S, Kisseljov F, Eienkel J, Hoeckel M, von Knebel Doeberitz M (2003) A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques. *Oncogene* 22:3977-3984
- Zuna RE, Moore WE, Shanesmith RP, Dunn ST, Wang SS, Schiffman M, Blakey GL, Teel T (2009) Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. *Int J Cancer* 125:2609-2613
- zur Hausen H (1996) Papillomavirus infections – a major cause of human cancers. *Biochim Biophys Acta* 1288:F55-78

zur Hausen H (2009) Papillomaviruses in the causation of human cancers - a brief historical account. *Virology* 384:260–265

Abstract

Human Papillomaviruses (HPVs) are non enveloped, epitheliotropic, double stranded, circular DNA viruses that infect cutaneous and/or mucosal epithelia. Over 150 different HPV types have been characterized, and approximately 40 of them are detected in the cervical epithelia and are classified as members of the genus Alphapapillomavirus. Epidemiological studies have revealed that persistent infection with high-risk human papillomavirus is the main risk factor for the development of high grade precancerous lesions and cervical carcinoma, with HPV16 and HPV18 types being the most frequently identified HPV types in invasive cervical cancer cases worldwide. Cancer progression is associated with persistent high-risk HPV infection and with deregulated expression of viral oncogenes E6 and E7. The E6 and E7 oncoproteins mediate mitogenic and antiapoptotic stimuli by interacting with numerous regulatory proteins of the host cell that control the cell cycle. As a consequence, HPV infection causes excessive cell proliferation, deficient DNA repair, and the accumulation of genetic damages in the infected cell.

Persistent infection with high-risk HPV types is associated with an increasing risk of genome integration into the host cell chromosome and progression to cancer. Viral integration deregulates the expression of the E6 and E7 proteins in the high-risk HPVs types. As a consequence, the virus stimulates cell cycle entry, providing selective growth advantage to the infected cells. In addition, the extensive proliferation of the infected epithelial cells is directly responsible for the accumulation of genetic errors, and genome destabilization that result in cancer development.

The present thesis, focused on the molecular analysis of HPV16 DNA. Moreover, the HPV16 DNA physical state and the HPV16 integration sites into the cellular chromosome were determined. The study was carried out in HPV16 positive cervical samples derived from Greek population, that were diagnosed as high grade, low grade squamous intraepithelial lesions and cervical cancer cases. Nucleotide and phylogenetic analysis of HPV16 DNA, showed that the prototype strain and the European variant are both prevalent in Greek population, while non European variants are circulating among Greek women. To the best of our knowledge, novel sequence variations were reported for the first time within E1, E2, E4, E6 and E7 genes, as well as duplications and intratypic recombination events among distinct HPV16 variants were also reported. However, molecular evolutionary analysis proposed that different genes at the early region of HPV16

genome are dominated by different selective pressure and that might be associated with the essential role of viral proteins in the productive viral life cycle.

In addition, the present thesis studied the physical state of HPV16 DNA, considering that the key importance for the progression of cervical intraepithelial neoplasias (CIN) to invasive cervical cancer is the integration of HPV into the host genome. The physical state of viral genome was determined, taking into consideration the most frequently disrupted sites of HPV16 DNA, revealing that the viral integration is an early event in the viral life cycle. Finally, according to integration sites analysis, rearrangements within HPV16 genome were determined. In particular, a model of looping was proposed by which viral - host DNA concatamers - mediate replication and recombination resulting to rearrangements in inverted orientation. According to this hypothesis, after HPV16 integration into the host chromosome, the virus promotes genome destabilization leading to cancer development. In conclusion, the extensive mapping analysis of disruption sites of E1 and E2 genes, in combination with the nucleotide analysis of E1, E2 and E6 genes, and the detection of HPV16 genome rearrangements, is a prerequisite in order to further understand and analyze the physical state of HPV16 DNA form in cancer progression.

ΔΗΜΟΣΙΕΥΣΕΙΣ

Sequence variation analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from women in Greece

D. Tsakogiannis · I. G. A. Ruether · Z. Kyriakopoulou ·
V. Pliaka · A. Theoharopoulou · V. Skordas ·
E. Panotopoulou · C. Nepka · P. Markoulatos

Received: 7 June 2011 / Accepted: 15 December 2011 / Published online: 1 February 2012
© Springer-Verlag 2012

Abstract The E2 gene of human papilloma virus is expressed at the early stage of the viral life cycle, encoding the E2 transcription factor, and regulates the expression of E6 and E7 oncogenes. Disruption of E2 gene due to viral integration inhibits the transcriptional suppression of the HPV oncogenes, inducing cell proliferation. In the present study, a total of 22 HPV16-positive cytological specimens derived from high- and low-grade cervical intraepithelial lesions were investigated in order to identify sequence variations in the HPV16 E2 ORF. The E2 gene was amplified by PCR using external and internal overlapping sets of primers. Amplicons were cloned and sequenced. Disruption sites were detected in cervical samples diagnosed as high-grade cervical intraepithelial lesions. Moreover, sequence variations were identified in the E2 ORF and specific variations were associated with non-European variants such as African type I, African type II and Asian American. A total of three new sequence variations were identified at positions 2791, 2823 (transactivation domain) and 3361 (hinge region). Distinct phylogenetic branches

were formed according to E2 analysis that characterized the different HPV16 variants. It was ascertained that non-European variants are circulating in the Greek population.

Introduction

Human papillomaviruses (HPVs) are non-enveloped, double-stranded DNA viruses that infect the basal layer of epithelia and are associated with benign and malignant lesions of the skin and the mucous membranes in humans. To date, more than 120 different HPV types have been identified, and about 40 of them infect the genital tract [1]. The mucosatropic HPVs are typically classified as high-risk (HR) genotypes, including HPV-16, and -18, and low-risk (LR) genotypes, such as HPV-6, -11, and -42, based on their association with clinical lesions [2, 3]. A further classification of HPV16 has revealed five distinct phylogenetic branches according to intratypic sequence polymorphisms in specific portions of the genome related to the geographic origin of the strains. Consequently, five HPV16 variants have been characterized: East Asian, European, African type I, African type II and Asian American [3, 4]. Moreover, HPV16, HPV18 and HPV45 are the most frequently identified HR HPV types found in invasive cervical cancer [5–9]. Nevertheless, it has been proposed that epigenetic alterations may also be necessary for the establishment of cervical neoplastic progression [10, 11].

The HPV genome encodes six early (E1, E2, E4, E5, E6, E7) and two late (L1, L2) genes. The early functional domain of the viral genome is regulated by the early promoter. This promoter is referred to as P97 in HPV-16 and -31 and as P105 in HPV-18 [12]. The early genes encode proteins responsible for replication, transcription and translation of infected cells. In particular, the HPV E6 and

D. Tsakogiannis · I. G. A. Ruether · Z. Kyriakopoulou ·
V. Pliaka · A. Theoharopoulou · V. Skordas ·
P. Markoulatos (✉)

Microbiology-Virology Laboratory, Department of Biochemistry and Biotechnology, School of Health Sciences, University of Thessaly, Ploutonos 26 and Aioulou, 41221 Larissa, Greece
e-mail: markoulatos@bio.uth.gr

E. Panotopoulou
Papanicolaou Research Centre of Oncology and Experimental Surgery, Virology Laboratory, Anticancer Oncology Hospital of Athens “St Savvas”, Athens, Greece

C. Nepka
Cytopathology Laboratory, University Hospital of Thessaly, Larissa, Greece

E7 early genes encode oncoproteins that are able to transform the infected cells. In addition, they are involved in maintenance of the viral genome extrachromosomally. The HPV E2 protein is also involved in both the replication of viral DNA and the transcriptional regulation of early genes and contributes to the control of the viral copy number in undifferentiated epithelial cells [13–15].

In previous studies, it has been reported that in low-grade CIN lesions, the high-risk HPV genomes persist as episomes, while during progression to high-grade lesions the viral genome is also found integrated into chromosomes of the host cells [16–27]. The integration of HPV usually disrupts the E2 gene open reading frame, resulting in the loss of the E2 protein function. These data led to the assumption that disruption of the HPV-16 E2 gene results in the deregulation of P97 promoter activity and an increase in the expression of the E6 and E7 viral oncogenes [28, 29]. As a result, E6 and E7 oncoproteins are persistently expressed, and they deregulate p53 and pRb proteins, respectively, providing the infected human keratinocytes with a selective growth advantage [29–31]. Consequently, it has been proposed that mutations in the E2 regulatory gene increase the immortalizing activity of HPV16 [32]. In the present study, a total of 22 HPV16-positive cytological specimens derived from high- and low-grade cervical intraepithelial lesions were investigated in order to determine the physical state and sequence variations of the E2 gene in high- and low-grade precancerous lesions.

Materials and methods

Clinical cervical specimens

A total of 22 HPV16-positive cervical samples were investigated. Five clinical Thin Prep cases were received from the Virology Laboratory of the Anticancer Hospital of Athens “Agios Savvas”, eight paraffin-embedded samples from cervical biopsies were from patients with cervical intraepithelial lesions (CIN), and nine clinical samples were from stained cervical smears submitted for cytologic analysis to the Cytopathology Laboratory of the University Hospital of Larissa. All these cases were diagnosed as cervical intraepithelial lesion stage I, II and III. Of the 22 specimens, 11 were diagnosed as high-grade (CIN II, III) and eleven as low-grade (CIN I) cervical intraepithelial neoplasia. All cervical samples that were assayed in the present study were from Greek patients.

DNA preparation

Three different DNA extraction methods were used for the three different groups of specimens. DNA from Thin

Prep was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) according to Casas et al. [33]. DNA from stained cervical smears was isolated as described by Puramen et al. [34]. The extraction of DNA from paraffin-embedded tissue sections is briefly described below. Sections of 5–10 μm in diameter were cut and collected in 0.5-ml microtubes using a clean blade. Each individual section was deparaffinized in xylene, washed with 96% ethanol, pelleted and air-dried. The sections were treated with 200 μl of 1X extraction buffer, which contained 20 $\mu\text{g}/\mu\text{l}$ of proteinase K, 100 μl of 10X PCR buffer (Stratagene, La Jolla, USA) and ultrapure water (Sigma, St. Louis, USA) up to a final volume of 1 ml. The samples were incubated either at 37°C overnight or at 55°C for 3 hours, and proteinase K was inactivated at 95°C for 15 min. The lysate was removed by a brief spin, and the resulting supernatant was transferred into a new tube and stored at -20°C . Three μl of the supernatant was used directly for DNA amplification by PCR assays.

The quality of the extracted DNA from cervical specimens and from formalin-fixed, paraffin-embedded human cervical tissues was tested by amplification of a 498-bp fragment of the human β -actin gene [20]. HPV typing of the specimens was performed using a nested multiplex PCR assay as described previously by Sotlar et al. [35].

E2 amplification by PCR

The HPV16 samples were subjected to PCR amplification of the E2 ORF using the set of primers (HPV-16 2734/HPV-16 3912) (Table 1). The primers were designed using the online Primer3 program, and the nucleotide positions are numbered according to the E2 sequence of the reference HPV16 strain available in the NCBI GenBank sequence database (<http://www.ncbi.nlm.nih.gov/>) under accession number NC_001526. The forward primer was designed to hybridize 21 bp upstream of the E2 gene, and the reverse primer was designed to hybridize 90 bp downstream of the E2 ORF.

PCR was performed in a final volume of 50 μl , and the PCR mixture contained 50 pmol of each primer, 10X PCR buffer (Stratagene, La Jolla, USA) containing 2 mM MgCl_2 , 0.25 mM each dNTP, and 2.5 U of thermostable DNA polymerase (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA). The cycling conditions were as follows: 40 cycles of 40 s at 95°C, 30 s at 56°C and 40 s at 72°C. The first cycle was preceded by a 2-min denaturation step at 95°C, and the last cycle was followed by a 5-min elongation step at 72°C. Subsequently, in order to increase the amplification signal, 2 μl of the first-round PCR product was used as template for auto-nested PCR using the primer set HPV-16 2734/HPV-16 3912. The

Table 1 Sequences of the outer and two overlapping inner primers and their respective positions on the E2 gene

	Primers	Location	Sequence (5'-3')	PCR product
E2 complete ORF	HPV-16 2734	2734	GACGAGGACAAGGAAAACGA	1176 bp
	HPV-16 3912	3912	GCACACAAAGCAAAGCAAAA	
5' region	HPV-16 2732	2732	GAGGACGAGGACAAGGAAAA	632 bp
	HPV-16 3364	3364	CGTTGCTGCTAAACACAGATG	
3' region	HPV-16 3312	3312	ATGCGGGTGGTCAGGTAAT	598 bp
	HPV-16 3912	3912	GCACACAAAGCAAAGCAAAA	

The primers were designed using Primer3

auto-nested PCR was carried out for 25 cycles under the same conditions.

The samples that were not amplified with the primer set HPV-16 2734/HPV-16 3912, were tested further with two inner overlapping primer sets (Table 1) in separate reactions. These primers were designed to amplify overlapping fragments of the E2 ORF in order to locate the sites of disruption in the E2 gene. All of the PCR reactions were performed in a final volume of 50 μ l, and each PCR mixture contained 50 pmol of each primer (HPV-16 2732/HPV-16 3364, HPV-16 3312/HPV-16 3912). Amplification conditions were the same as described above. In addition, 2 μ l of the first-round PCR product was used as template for auto-nested PCR, which was performed for 25 cycles under the same conditions as above. Ten microliters of the amplified samples was electrophoresed in a 2% agarose gel containing 1 μ g/ml ethidium bromide in Tris-borate-EDTA buffer using a 100-bp DNA ladder (Invitrogen Life Technologies, Paisley, UK) as a molecular weight marker.

Sequence analysis of E2

PCR amplicons were excised from the agarose gel using a QIAquick Gel Extraction Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer's instructions and were cloned into pGEM-T Easy vector (Promega, Madison, USA). Recombinant plasmids containing the 1176-bp, 632-bp or 598-bp fragment were prepared by miniprep using a NucleoSpin Plasmid kit (Macherey-Nagel GmbH, Duren, Germany). In order to confirm the nucleotide variations, two independent PCR assays were carried out for each individual cervical sample, and amplicons from the different PCR assays were subjected to cloning. Three clones from each individual amplicon were selected. The samples were sequenced at Macrogen Inc., Korea. The HPV16 E2 cloned sequences were analyzed using the ORF Finder tool on the NCBI server of the National Institutes of Health (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Multiple sequence alignment of the cloned E2 sequences and the

reference HPV16 sequence (HPV16R) available in the HPV16 Sequence Database (Los Alamos National Laboratory) was done using the multiple sequence alignment program Mega5 in order to detect nucleic acid variations.

Phylogenetic analysis

A phylogenetic tree was constructed on the basis of the cloned sequences of the complete E2 ORFs derived from eleven low-grade cervical samples, the HPV16 E2 reference sequences and those of representative sequences of HPV16 variants, such as African type I, African type II, Asian-American, European and Asian (data not shown). In addition, considering that some cervical samples were amplified only by the two internal overlapping primers sets (Table 1), two phylogenetic trees in the portion of the 5' region and of the 3' region of the E2 gene were also constructed using the corresponding partial E2 sequences of the reference HPV16 strain and of the HPV16 variants (data not shown).

The sequences of the HPV16 variants that were used to construct the distinct HPV16 phylogenetic branches are available in the GenBank sequence database under the following accession numbers: for the Asian American variant, AF402678 (Asian American) and AF407220, AF115832, AF115831, DQ191465, DQ191464, AY686582, AY686579; for the Asian variant, AF534061 (Asian) and AF40719, AF407218, FJ610150; for the African type I variant, AF536180 (African type 1) and AF472508; for the African type II variant, AF472509 (African type 2) and AF193429; and finally, for the European variant, AF536179 (European) and AF067034, AF193425, AF19347, AY686580, EU430675, EU430667, EU430683, FJ610148, FJ610149, FJ610152, FJ610146, FJ610147, DQ191463. The phylogenetic trees were calculated by the neighbor-joining method using the program Mega5 (1000 bootstrap replicates). ClustalW2 multiple sequence alignment analysis was also carried out in order to calculate the percentage of sequence similarity between the E2 amplicons and the representative sequences of the HPV16 variants.

Results

E2 ORF disruption

Intact E2 ORF was detected in eleven low-grade cytological specimens, while the eleven cases of high-grade intraepithelial neoplasia revealed E2 disruption. The E2 genomic region from nucleotides 3312 to 3912 was not amplified in the specimen D30 (Table 2), while the 5' genomic region of the E2 gene from nucleotides 2732 to 3364 was not amplified in three more specimens, D2, D4 and D14 (Table 2). The amplification of the human β-actin gene in all paraffin-embedded cervical samples revealed that the disruption in E2 ORF was caused by viral integration and not by the formalin fixation process. These findings indicate that the viral genome must be integrated into chromosomes of the host cells and that E2 disruption took place in the corresponding genomic region of the E2 ORF.

E2 Sequence variations

Three colonies were selected from each cloned amplicon, and the isolated plasmids were used for direct sequencing.

The selection of clones was performed according to the sequence heterogeneity among the different cloned sequences derived from the same cervical sample. The numbering of the nucleotide variations was based on the reference sequence of HPV16 (HPV-16R), which is available in the HPV16 Sequence Database (Los Alamos National Laboratory).

A total of nine sequence variations were identified in the portion of the E2 gene that encodes the transactivation domain. Two new variations in this domain were detected at positions 2823 and 2791 (Tables 2, 3). The sequence variation at position 2823 results in the amino acid substitution S23I (Table 2) and was detected only in cervical samples derived from high-grade lesions (samples A13 and D30). The sequence variation at position 2791 was detected only in the low-grade clinical samples D7 and D34 (Table 3) with no amino acid change. Moreover, ten variations were detected in the hinge domain, including one new sequence change at position 3361. The C3361T variation was detected in three cloned E2 sequences derived from one low-grade (A17) and two high-grade (A3 and A4) samples (Tables 2, 3) without resulting in an amino acid change. In addition, seven nucleic acid variations (A3622T,

Table 2 Sequence variations in the E2 gene and amino acid substitutions detected in eleven cervical samples with high-grade intraepithelial neoplasia

Table with 14 columns: CIN: HG, Clones, Nucleotide positions (2-33), Variants, Accession number in GenBank, and Amino acid. The table lists various cervical samples and their corresponding sequence variations in the E2 gene and amino acid substitutions. It also includes domain labels like 'Hinge domain', 'DNA binding domain', and 'Transactivation domain'.

The HG cervical cases that are listed in this table were amplified only by the internal overlapping primer sets. The corresponding HPV16 variants were identified using phylogenetic trees and ClustalW2 analysis. The codes A3, A4, A6, A13, A15, A31, D2, D4, D14, D16, and D30 correspond to the cervical samples. The numbers 632 and 598 correspond to the size of the amplified PCR product, and the last letter (C1 or C2) corresponds to the number of the clone.

Table 3 Sequence variations and amino acid substitutions of eleven low-grade cervical samples with intact E2 genes

CIN: LG	Clones	Sequence variations																								Variants	Accession number in GenBank			
		2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3			3		
		7	8	8	9	1	1	1	2	2	3	3	3	4	4	5	5	5	5	6	6	6	7	7	7	8				
		9	2	6	7	5	6	8	2	4	6	6	7	1	4	1	1	3	6	2	8	9	0	7	8	0				
		1	3	0	7	9	1	2	4	9	1	2	7	0	3	9	6	7	8	6	2	4	4	6	8	7	5			
		G	G	C	A	C	C	G	T	G	C	A	C	C	G	G	C	T	A	T	A	C	T	T	G	C	T			
CIN I	D7 632 C1	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422083	
CIN I	D7 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422085	
CIN I	D7 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422084	
CIN I	D7 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422086	
CIN I	D34 632 C1	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422087	
CIN I	D34 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422089	
CIN I	D34 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422088	
CIN I	D34 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422090	
CIN I	D42 632 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422091	
CIN I	D42 598 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422093	
CIN I	D42 598 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JF422094	
CIN I	747 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	A	-	-	-	-	-	T	-	-	-	-	-	European	HM162464	
CIN I	747 1176 C2	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	T	-	-	-	-	-	European	HM162465	
CIN I	747 1176 C3	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	T	-	-	-	-	-	European	HM162466	
CIN I	747 1176 C4	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	T	-	-	-	-	-	European	HM162467	
CIN I	747 1176 C5	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	T	-	-	-	-	-	European	HM162468	
CIN I	796 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	A	-	-	-	-	-	T	-	-	-	-	-	European	HM162469	
CIN I	858 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM162475	
CIN I	868 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM162477	
CIN I	793 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM162478	
CIN I	A17 1176 C1	-	-	A	-	A	T	A	-	A	-	A	G	G	T	-	A	A	C	-	G	-	A	G	C	T	A	G	African type II	HQ398228
CIN I	A17 1176 C2	-	-	A	-	A	T	A	-	A	T	G	G	T	-	A	A	C	C	G	-	A	A	C	T	A	G	African type II	HQ398229	
CIN I	A24 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM162463	
CIN I	A38 1176 C1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	European	HM162462	
Amino acid		I	S	H	E	T	H	A	G	R	S	N	P	P	A	E	T	T	S	F	V	T	T	S	W	D	V			
		0	0	0	0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3		
		1	2	3	7	3	3	4	5	6	0	0	0	1	2	3	5	5	6	7	8	1	1	1	4	4	5			
		5	3	5	4	5	6	3	6	5	2	3	8	9	0	2	4	4	1	1	9	0	3	7	1	4	0			
		L	I	Q	D	K	Y	T	-	Q	-	D	A	S	T	K	N	N	-	V	-	K	-	-	C	E	-			
		Transactivation domain										Hinge domain						DNA binding domain												

The corresponding HPV16 variants were identified using phylogenetic tree and ClustalW2 analysis. The codes 747, 793, 796, 858, 868, A24, A38, A17, D7, D34, D42 correspond to the cervical samples. The numbers 1176, 632 and 598 correspond to the size of the amplified PCR product, and the last letter (C1 or C2) corresponds to the number of the clone

C3684A, T3694A, T3706G, G3778T, C3787A, T3805G) were found in the DNA binding domain. These variations were found in low- and high-grade cervical cases that have been correlated with HPV16 African type I, African type II, Asian American and European variants (Tables 1, 2).

Specific variations in the E2 gene, such as T3224A and G3413A, were identified in a group of clones (A3, A15) (Tables 2, 3) that were closely related to the Asian American variant as calculated using the online ClustalW2 multiple sequence alignment program. A high degree (>97%) of nucleotide sequence similarity was noted between the Asian American variant and the cloned E2 sequences (Table 2).

As a result, this strong association between specific variations and the HPV16 variants could be a significant step towards the identification of the HPV16 Asian American variant in premalignant intraepithelial neoplasia.

Phylogenetic clustering of HPV16 isolates

Phylogenetic and ClustalW2 analysis were performed based on the complete HPV16 E2 coding sequence, the 3' region and the 5' region of the E2 gene as described above. Several HPV16 variants were detected in

pre-malignant cervical samples. In particular, samples that had been diagnosed as CIN I lesions (Table 3) contained an intact E2 coding region, and phylogenetic and ClustalW2 analysis revealed a strong relationship between the cloned sequences and the European variant in the corresponding genomic portion (Table 3). ClustalW2 analysis revealed that the similarities between the characterized HPV16 strains and the European variant were 98% in E2 ORF and less than 97% with the non-European HPV16 variants in the same portion of the genome. In contrast, the LG sample A17 was clustered with the African type II variant, with 99% sequence similarity in the E2 ORF (Table 3).

Phylogenetic tree and multiple DNA sequence alignment analysis of the 3' region and the 5' region of the E2 gene revealed a strong association between cloned sequences derived from five high-grade cervical cases (A6, A13, A31, D16, D30) and the European variant (Table 2). In addition, cloned sequences derived from the high-grade samples A3 and A15 were classified with the HPV16 Asian American variant, with 97%-98% sequence similarity, while the high-grade cervical case A4 was clustered in the same branch with the African type I variant, with 98% sequence similarity (Table 2).

Discussion

The purpose of the present study was to identify HPV16 E2 sequence variations in 22 cervical samples that were diagnosed as premalignant intraepithelial neoplasia. One outer primer set was designed to amplify the complete E2 ORF, and two inner primer sets were used to amplify internal overlapping fragments in the 5' and the 3' region of the E2 gene in order to detect disruption sites. Disruption sites were dominant in high-grade clinical cases. Nucleotide variations were detected in the E2 sequences, mainly in the region encoding the hinge and the transactivation domain of the E2 protein. Multiple sequence alignment analysis revealed that specific nucleotide variations were associated with non-European variants such as African type I, African type II and Asian American variants. The nucleotide variations were studied by analysis of the cloned sequences derived from the high- and low-grade cervical samples. Cloned sequences were selected according to the nucleotide homology between the different clones derived from the same sample. In particular, only one cloned sequence was used in nucleotide and phylogenetic analysis of cervical samples when the cloned sequences showed 100% nucleotide homology (Tables 1, 2). In particular, a total of six cloned sequences derived from the cervical sample D30 were characterized. From these cloned sequences, four clones were identical to the HPV16 reference sequence, while in two cloned sequences, the nucleotide variation A2823G was identified. One representative cloned sequence from each HPV16 strain (the A2823G variant and the HPV16 reference strain) is listed in Tables 1 and 2. Taking these data into account, it was assumed that distinct HPV16 strains are present in the same cervical sample. In addition, different HPV16 strains were identified in the samples A3, A4, A13, D7, D30, D34, 747 and A17 according to nucleotide polymorphisms that were detected in the E2 region (Tables 1, 2).

In previous studies, it was reported that the disruption of the E2 gene is a common and early event following HPV16 or HPV18 infection [16–18, 22, 23, 36–39]. In the present report, intact E2 gene was detected in HPV16 samples derived from low-grade intraepithelial neoplasia, while E2 gene disruptions were detected only in high-grade premalignant lesions (Table 2). These data suggest that the E2 gene was disrupted and that the HPV16 genome was integrated most frequently into the host genome in high-grade lesions, as has been reported in earlier studies [19–21, 24–27].

In the present report, a total of 29 sequence variations were identified in the cloned E2 sequences. Of the 29 sequence variations, 19 resulted in amino acid changes in the three HPV16 E2 protein domains (Tables 2, 3). The

majority of the nucleotide variations identified have also been reported in previous studies [40–50], but three new variations were reported for the first time in the present report. In the 5' region of the E2 gene, which encodes the transactivation domain of the protein, two new variations were detected at position 2791 and 2823 in cervical cases that were associated with HPV16 European variants (Tables 2, 3). The sequence variation at position 2791 does not result in an amino acid change (Table 3), while the variation at position 2823 results in the amino acid substitution S23I (Table 2). The S23I amino acid substitution is located within the E1 interaction site at residues 18–41 [51]. This genomic region is implicated in the formation of the E1–E2 complex responsible for the viral DNA replication.

In addition, the analysis of the HPV16 E2 ORF sequence revealed a new nucleic acid variation at position 3361. The C3361T variation failed to induce an amino acid substitution in the E2 protein. However, the specific nucleotide variation was reported in cervical samples strongly related to HPV16 Asian American, African type I and African type II variants (Tables 2, 3). The strong association of C3361T with non-European variants requires further investigations in order to use the specific nucleotide change as a molecular marker for the identification of non-European variants. Consequently, it is reported that non-European variants are circulating in the Greek population.

The hinge region harbors the amino acid change P219S (C3410T). This variation was detected in different HPV16 variants derived from both high- and low-grade cervical premalignant intraepithelial lesions (Tables 2, 3). In an earlier study by Graham et al. [43], it was reported that this variation was found in high-grade clinical samples, and it had been suggested that HPV16 variants that carry this variation could be involved in the development of high-grade and invasive malignancy. In the present report, the P219S nucleotide variation was detected in both high- and low-grade cervical samples (Table 3). Amino acid variations in the hinge region could influence the three-dimensional structure of the protein, inhibiting the functional activity of the protein.

Sequence analysis of the HPV16 E2 gene revealed that the nucleotide variations C2860A, A2977T, C3159A, C3161T, G3182A, G3249A (transactivation domain) and A3362G, C3377G, G3449A, C3516A (hinge domain) were present in cervical samples with high sequence homology to HPV16 Asian American, African type I and African type II variants (Tables 2, 3) [48]. In addition, the nucleotide change A3538C was reported in cervical cases with strong association with HPV16 Asian American and African type II variants (Tables 2, 3). Moreover, the specific sequence variations T3224A and G3413A were only identified in HPV16 Asian American variants (Tables 2, 3).

The underlying mechanism of cervical carcinogenesis induced by oncogenic HPVs is mediated through protein-protein interactions of the host and virus. From this point of view, amino acids conserved across viral HPV16 variants are more likely to be functionally significant. Therefore, variations that change these amino acids might be more likely to be associated with altered protein functions involved in virulence or enhanced capacity to transform the cervical epithelium. The E2 sequence variations could also be involved in an alternative mechanism for deregulating the expression of the HPV E6 and E7 oncogenes.

In conclusion, this is the first study reporting specific E2 sequence variations and E2 disruption sites in HPV16 strains from women in Greece. Such findings might have implications for future studies of HPV infection and the design of new viral biomarkers for the elucidation of mechanisms involved in cervical carcinogenesis.

Acknowledgments This work was supported by research grants of the Postgraduate Programme ‘Applications of Molecular Biology-Genetics. Diagnostic Biomarkers’, code 3817, of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology.

Conflict of interest All authors declare that they have no conflicting or dual interests.

References

- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79
- Zur Hausen H (1996) Papillomavirus infections—a major cause of human cancers. *Biochim Biophys Acta* 1288:F55–F78
- Bernard HU, Calleja-Macias IE, Dunn ST (2006) Genome variation of human papillomavirus types: phylogenetic and medical implications. *Int J Cancer* 118:1071–1076
- Yamada T, Wheeler CM, Halpern AL, Stewart ACM, Hildesheim A, Jenison SA (1995) Human Papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2 and L1 coding segments. *J Virol* 69:7743–7753
- Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, Tortolero-Luna G, Kjaer SK, Muñoz N (2008) Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. *Vaccine* 26:K1–K17
- Schiffman M, Clifford G, Buonaguro FM (2009) Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infect Agent Cancer* 4:8
- Sundström K, Eloranta S, Sparén P, Arnheim Dahlström L et al (2010) Prospective study of human papillomavirus (HPV) types, HPV persistence, and risk of squamous cell carcinoma of the cervix. *Cancer Epidemiol Biomarkers Prev* 19:2469–2478
- Castle PE, Schiffman M, Wheeler CM, Wentzensen N, Gravitt PE (2010) Human papillomavirus genotypes in cervical intraepithelial neoplasia grade 3. *Cancer Epidemiol Biomarkers Prev* 19:1675–1681
- Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX, de Sanjose S (2010) Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *J Infect Dis* 202:1789–1799
- Dong SM, Kim HS, Rha SH, Sidransky D (2001) Promoter hypermethylation of multiple genes in carcinoma of the uterine cervix. *Clin Cancer Res* 7:1982–1986
- Feng Q, Balasubramanian A, Hawes SE, Toure P, Sow PS, Dem A, Dembele B, Critchlow CW, Xi L, Lu H, McIntosh MW, Young AM, Kiviat NB (2005) Detection of hypermethylated genes in women with and without cervical neoplasia. *J Natl Cancer Inst* 97:273–282
- Smotkin D, Wettstein FO (1986) Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. *Proc Natl Acad Sci USA* 83:4680–4684
- Soeda E, Ferran MC, Baker CC, McBride AA (2006) Repression of HPV16 early region transcription by the E2 protein. *Virology* 351:29–41
- Longworth MS, Laimins LA (2004) Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev* 68:362–372
- Hamid NA, Brown C, Gaston K (2009) The regulation of cell proliferation by the papillomavirus early proteins. *Cell Mol Life Sci* 66:1700–1717
- Huang LW, Choa SL, Lee BH, Lee BH (2008) Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. *J Clin Pathol* 61:627–631
- Cricca M, Morselli-Labate AM, Venturoli S et al (2007) Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions. *Gynecol Oncol* 106(549):557
- Briolat J, Dalstein V, Saunier M et al (2007) HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. *Int J Cancer* 121:2198–2204
- Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM (2006) Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 44:1755–1762
- Li W, Wang W, Si M et al (2008) The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355–1361
- Tonon SA, Picconi MA, Bos PD, Zinovich JB, Galuppo J, Alonio LV, Teyssie AR (2001) Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J Clin Virol* 21:129–134
- De Marco L, Gillio-Tos A, Bonello L, Ghisetti V, Ronco G, Merletti F (2007) Detection of human papillomavirus type 16 integration in preneoplastic cervical lesions and confirmation by DIPS-PCR and sequencing. *J Clin Virol* 38:7–13
- Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B (2005) Type distribution, viral load and integration state of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer* 92:2195–2200
- Hudelist G, Manavi M, Pischinger KID, Watkins-Riedel T, Singer CF, Kubista E, Czerwenka KF (2004) Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol Oncol* 92:873–880
- Pett MR, Herdman MT, Palmer RD, Yeo GSH, Shivji MK, Stanley MA, Coleman N (2006) Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *Proc Natl Acad Sci USA* 103:3822–3827

26. Matovina M, Sabol I, Grubisić G, Gasperov NM, Grce M (2009) Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecol Oncol* 113:120–127
27. Cheung JL, Lo KW, Cheung TH, Tang JW, Chan PK (2006) Viral load, E2 gene disruption status, and lineage of human papillomavirus type 16 infection in cervical neoplasia. *J Infect Dis* 194:1706–1712
28. Doeberitz V, Bauknecht M, Bartsch T, Zur Hausen H (1991) Influence of chromosomal integration on glucocorticoid-regulated transcription of growth-stimulating papillomavirus genes E6 and E7 in cervical carcinoma cells. *Proc Natl Acad Sci USA* 18:1411–1415
29. Jeon S, Lambert PF (1995) Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proc Natl Acad Sci USA* 92:1654–1658
30. Jeon S, Allen-Hoffmann B, Lambert P (1995) Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol* 69:2989–2997
31. Boulet G, Horvath C, Vanden Broeck D, Sahebali S, Bogers J (2007) Human papillomavirus: E6 and E7 oncogenes. *Int J Biochem Cell Biol* 39:2006–2011
32. Storey A, Greenfield I, Banks L, Pim D, Crook T, Crawford L, Stanley M (1992) Lack of immortalizing activity of a human papillomavirus type 16 variant DNA with a mutation in the E2 gene isolated from normal human cervical keratinocytes. *Oncogene* 7:459–465
33. Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25–36
34. Puranen M, Saarikoski S, Syrjänen K, Syrjänen S (1996) Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391–395
35. Sotlar K, Diemer D, Dethleffs A, Hack Y et al (2004) Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol* 42:3176–3184
36. Collins SI, Constandinou-Williams C, Wen K, Young LS, Roberts S, Murray PG, Woodman CB (2009) Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer Res* 69:3828–3832
37. Kulmala SM, Syrjönen SM, Gyllenstein UB et al (2006) Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 59:513–517
38. Peitsaro P, Johansson B, Syrjanen S (2002) Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *J Clin Microbiol* 40:886–891
39. Gallo G, Bibbo M, Bagella L et al (2003) Study of viral integration of HPV-16 in young patients with LSIL. *J Clin Pathol* 56:532–536
40. Azizi N, Brazete J, Hankins C, Money D et al (2008) Canadian Women's HIV Study Group. Influence of human papillomavirus type 16 (HPV-16) E2 polymorphism on quantification of HPV-16 episomal and integrated DNA in cervicovaginal lavages from women with cervical intraepithelial neoplasia. *J Gen Virol* 89:1716–1728
41. Giannoudis A, Duin M, Snijders PJ, Herrington CS (2001) Variation in the E2-binding domain of HPV 16 is associated with high-grade squamous intraepithelial lesions of the cervix. *Br J Cancer* 84:1058–1063
42. Jiang M, Baseman JG, Koutsky LA, Feng Q, Mao C, Kiviat NB, Xi LF (2009) Sequence variation of human papillomavirus type 16 and measurement of viral integration by quantitative PCR. *J Clin Microbiol* 47:521–526
43. Graham DA, Herrington CS (2000) HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and invasive squamous cell carcinomas of the cervix: relation to numerical chromosome abnormalities. *Mol Pathol* 53:201–206
44. Eriksson A, Herron JR, Yamada T et al (1999) Human papillomavirus type 16 variant lineages characterised by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* 80:595–600
45. Casas L, Galvan SC, Ordoñez RM, Lopez N, Guido M, Berumen J (1999) Asian-american variants of human papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. *Int J Cancer* 83:449–455
46. Sathish N, Abraham P, Peedicayil A, Sridharan G, Shaji RV, Chandy G (2004) E2 sequence variations of HPV 16 among patients with cervical neoplasia seen in the Indian subcontinent. *Gynecol Oncol* 95:363–369
47. Watts KJ, Thompson CH, Cossart YE, Rose BR (2002) Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. *Int J Cancer* 97:868–874
48. Swan DC, Rajeevan M, Tortolero-Luna G, Follen M, Tucker RA, Unger ER (2005) Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* 96:695–700
49. Meissner J (1997) Sequencing errors in reference HPV clones. In: Myers G, Munger K, Sverdrup F, McBride AA, Bernard HU, Meissner J (eds) *Human Papillomavirus 1997: a compilation and analysis of nucleic acid and amino acid sequences*. Los Alamos National Laboratory, NM, pp 110–123
50. Veress G, Szarka K, Dong XP, Gergely L, Pfister H (1999) Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J Gen Virol* 80:1035–1043
51. Piccini A, Storey A, Massimi P, Banks L (1995) Mutations in the human papillomavirus type 16 E2 protein identify multiple regions of the protein involved in binding to E1. *J Gen Virol* 76:2909–2913

Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece

D. Tsakogiannis · I. G. A. Ruether · Z. Kyriakopoulou ·
V. Pliaka · V. Skordas · C. Gartzonika ·
S. Levidiotou-Stefanou · P. Markoulatos

Received: 17 January 2012 / Accepted: 23 April 2012 / Published online: 9 June 2012
© Springer-Verlag 2012

Abstract The HPV16 E1[^]E4 protein is thought to contribute to the release of newly formed viral particles from infected epithelia. In order to investigate amino acid mutations in the HPV16 E1[^]E4 protein, the complete E4 ORF was amplified by PCR in 27 HPV16-positive cervical samples, and the amplicons were cloned. Fifteen nucleic acid variations were identified in the E4 ORF, including seven silent nucleic acid mutations. In addition, nine amino acid mutations (A7V, A7P, L16I, D45E, L59I, L59T, Q66P, S72F, H75Q) were detected in the E1[^]E4 protein, and these were associated with the severity of cervical malignancy. A maximum-likelihood phylogenetic tree was constructed based on the E4 ORF, and nucleotide sequence analysis of the E4, E6 and E7 genes from the same samples was conducted in order to determine the phylogenetic origin of the cloned sequences from the amplified HPV16 E4. Based on the nucleotide sequence and phylogenetic analysis it was revealed that even though E4 ORF constitutes a small polymorphic portion of the viral genome (288 bp), it could provide valuable information about the origins of the HPV16 genome. In addition, molecular evolutionary analysis of the E4 coding region revealed that neutral selection is dominant in the overlapping region of the E4 and E2 ORFs.

Introduction

Human papilloma viruses (HPVs) constitute a group of small double-stranded DNA viruses that infect the basal layer of the epithelium of skin and mucous membranes. Over 100 different HPV types infect humans, and 40 HPV types are associated with anogenital tract malignancy [1]. HPV16 and HPV18 are the most common types identified in invasive cervical cancer cases [2–6]. Genomic intratypic heterogeneity has been mainly characterized in HPV16, and five distinct HPV16 variants have been identified (African type I, African type II, Asian-American, European and Asian) according to their geographical distribution [7–12]. Previous studies have suggested that non-European HPV16 variants constitute a high risk factor for more-severe cervical malignancies [13–15].

Infection with mucosatropic HPV types is a major factor for cell proliferation induced by overexpression of the viral oncogenes E6 and E7. HPV genome integration into the host chromosome usually results in the functional elimination of the viral E2 transcriptional repressor. As a result, viral transcriptional control is depressed, and the viral oncoproteins are persistently expressed, inactivating the tumor suppressor proteins p53 and pRB [16–19]. The loss of function of the cell-cycle regulatory proteins induces extended proliferation of the infected epithelial cells, leading to premalignant and malignant lesions [20].

The HPV genome is divided into two functional domains, which are expressed via two major promoters during the early and late phases of the viral life cycle. The majority of HPV mRNAs are polycistronic, including more than one open reading frame (ORF) and gene expression is regulated by alternative splicing mechanisms [21, 22], as is the case for the E4 protein, also known as E1[^]E4 late protein. The first five amino acids are derived from the E1

D. Tsakogiannis · I. G. A. Ruether · Z. Kyriakopoulou ·
V. Pliaka · V. Skordas · P. Markoulatos (✉)
Department of Biochemistry and Biotechnology, Microbiology-
Virology Laboratory, School of Health Sciences, University
of Thessaly, Ploutonos 26 and Aiolou, 41221 Larissa, Greece
e-mail: markoulatos@bio.uth.gr

C. Gartzonika · S. Levidiotou-Stefanou
Department of Microbiology, Medical School, University
of Ioannina, Ioannina, Greece

ORF, while the remaining amino acids are derived from the E4 ORF [23, 24]. The E1[^]E4 protein is the first late protein that is expressed in the upper layers of the epithelium and is the most highly expressed of the HPV proteins [24–26]. The HPV16 E1[^]E4 protein is thought to be responsible for the total collapse of the epithelial cell intermediate filament during the late stage of viral infection, releasing infective novel viruses [27]. In previous studies, it has been reported that overexpression of the E1[^]E4 protein induces G₂ arrest in several HPV-infected cell types [28–30]. However, Davy et al. [31] have suggested that the E2 protein is stabilized by the presence of the E1[^]E4 protein and that it influences E2 protein activity.

In the present report, considering the essential role of the E4 protein during the productive phase of HPV infection, we have investigated the HPV16 E4 ORF in order to detect sequence variations in premalignant cervical lesions. The main objective of the present report was to reveal E4 gene mutations, to conduct phylogenetic analysis based on the E4 ORF, and to correlate these sequences with HPV16 variants, considering that HPV16 variants have been mainly characterized in E6, E7, L1, E2 and LCR, while data on E4 variants remain scarce.

Materials and methods

Cervical specimens

Eleven Thin Prep slides, nine clinical samples from stained cervical smears and seven paraffin-embedded cervical biopsy samples were tested from women diagnosed as having HPV16 associated with cervical intraepithelial neoplasia (CIN), stage I, II or III. Fifteen of these cases were diagnosed as high-grade cervical intraepithelial neoplasia (CIN II, III), and 12 were diagnosed as low-grade cervical intraepithelial neoplasia (CINI).

DNA preparation

One hundred μ l of Thin Prep solution was used for DNA extraction using the chaotropic agent guanidine thiocyanate (GuSCN) according to Casas et al. [32]. DNA from stained cervical smears was extracted as described previously by Puramen et al. [33]. DNA from formalin-fixed, paraffin-embedded cervical tissues was extracted by the proteinase K extraction method as described previously [10]. Three μ l of the eluted DNA was used for DNA amplification by PCR assay.

The quality of extracted DNA was tested by amplification of a 248-bp fragment of the beta-globin gene, while HPV typing of the cervical samples was performed by

nested multiplex PCR assay as described previously by Sotlar et al. [34].

Amplification of the E4 ORF

PCR amplification of the E4 ORF was performed using one primer set (HPV16 3312, 5'-ATGCGGGTGGTCAGGTAAT-3'/HPV16 3912, 5'-AGCACACAAAGCAAAGCAA A-3'). The Primer3 program was used to design the primer pair (<http://frodo.wi.mit.edu/primer3/>). The nucleotide numbering of the primer positions is according to the prototype E4 sequence of the reference HPV-16 strain available in the NCBI GenBank sequence database (<http://www.ncbi.nlm.nih.gov/>) under accession number NC_001526.

PCR assay was performed in a final volume of 50 μ l, including 50 pmol of each primer, (HPV-16 3312/HPV-16 3912), 10X PCR buffer (Stratagene, La Jolla USA) containing 2 mM MgCl₂, 0.25 M each dNTP and 2.5 U of thermostable DNA polymerase (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA). The cycling conditions were as follows: A DNA denaturation step for 2 min at 95 °C, followed by 40 cycles of denaturation for 40 s at 95 °C, annealing for 30 s at 56 °C and extension for 40 s at 72 °C. The last cycle was followed by a 5-min elongation step at 72 °C. Two μ l of the obtained product was used as template in auto-nested PCR using the same primer set in order to improve the amplified signal. The auto-nested PCR was performed for 25 cycles under the same conditions. Ten μ l of the amplified samples was electrophoresed in a 2 % agarose gel containing 1 μ g ml⁻¹ ethidium bromide in Tris-borate-EDTA buffer using a 100-bp DNA ladder (Invitrogen Life Technologies, Paisley, UK) as a molecular weight marker.

Amplification of the E6 and E7 genes

HPV16 variant determination was performed through nucleotide sequence analysis of the HPV16 E6 and E7 genes. The E6 and E7 region was amplified using the primer set HPV16 41, 5'AGCGACCCAGAAAGTTACCA-3'/HPV16 757, 5'-ATGGGGCACACAAT TCCTAG-3'. The primers were designed using the Primer3 program (<http://frodo.wi.mit.edu/primer3/>), and the numbering of nucleotide positions was based on the reference sequence of the HPV16 genome (NC_001526). The PCR reaction was performed in a final volume of 50 μ l. The PCR mixture contained 50 pmol of each primer, 5X Green GoTaq Flexi Buffer (Promega), 2 mM MgCl₂, 1 mM dNTPs and 1.25 U of thermostable DNA polymerase (GoTaq[®] DNA Polymerase, Promega). The cycling conditions were as follows: 40 cycles of 40 s at 95 °C, 50 s at 55 °C and 1 min at 72 °C. The first cycle was preceded by a 2-min

denaturation step at 95 °C and the last cycle was followed by a 5-min elongation step at 72 °C.

DNA sequencing analysis

The E6-E7 and E4 amplicons were excised from an agarose gel and purified using a QIAquick Gel Extraction Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer's instructions. The PCR products were ligated into pGEM-T Easy Vector (Promega, Madison, USA), and the plasmid constructs were used to transform *Escherichia coli* DH5a cells. The plasmid DNA was purified using a NucleoSpin Plasmid kit (Macherey-Nagel GmbH, Duren, Germany), and samples were sequenced at Macrogen Inc., Korea. In order to confirm the nucleotide sequence variations in the E4, E6 and E7 genes, two independent PCR assays were carried out for each viral domain, and amplicons from the distinct PCR assays were subjected to cloning. For E4 gene analysis, two clones from each individual amplicon were selected, while one clone was selected for E6 and E7 gene analysis. However, both strands of all tested clones were used in the sequencing reactions. The DNA strands were sequenced three times to ensure minimum reliability of results.

Multiple sequence alignment was carried out using the program MEGA v.5 [35] and the Muscle algorithm [36], comparing the cloned E4 sequences and the reference HPV16 E4 sequence (HPV16R) available in the HPV16 Sequence Database (Los Alamos National Laboratory) to detect sequence variations and amino acid substitutions. However, nucleotide polymorphisms in the E6 and E7 genes were detected using multiple sequence alignment of the amplified E6-E7 sequences, the reference sequence of the HPV16 genome (HPV16R), and those of representative sequences of HPV16 variants available in the GenBank sequence database under accession numbers AF536179 for the European variant (European-German type), AF534061 for the East Asian variant, AF472508 for the African type I variant, AF472509 for the African type II variant, and AF402678 for the Asian American variant.

For E4 nucleotide sequence and phylogenetic analysis, two clones were used from the cervical samples A3, A4, A13, A17 PNL2, PNL4, PNL5, PNL7, D7, D34, D42, one clone from the cervical samples A6, A15, A24, A31, A38, D2, D4, D14, D16, PNL11, 796, 858, 868, 793, and three clones from the sample PNL6 (Tables 1, 2). From the cervical sample 747, five clones were sequenced and used for phylogenetic and nucleotide sequence analysis. Clones were selected based on the sequence heterogeneity of the different cloned sequences derived from the same cervical sample. Only one cloned sequence was used in nucleotide and phylogenetic analysis for cervical samples in which the nucleotide sequences of all of the cloned sequences were identical (Tables 1, 2).

E4 Phylogenetic tree

A maximum-likelihood phylogenetic tree (1000 bootstrap replicates) was constructed from the alignment of the reference HPV16 E4 sequence (HPV16R), the complete E4 ORF cloned sequences derived from the 27 cervical samples (Tables 1, 2), and those of representative sequences of HPV16 variants, African type I, African type II, Asian-American, European and Asian (Fig. 1). Prior to phylogenetic estimation, the online program FindModel (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>) was used to identify the best evolutionary model. The GTR nucleotide substitution model was used for the E4 phylogenetic tree. The nucleotide sequences that were used to construct the E4 phylogenetic tree are available in the GenBank sequence database under accession numbers FJ610152, FJ610150, FJ610151, FJ610149, FJ610148, FJ610147, FJ610146, FJ006723, AY686584, AY686583, AY686581, AY686580, AF125673 and AF536179 for the European variant (European-German type), AF534061 for the East Asian variant, AF472508 and AF536180 for the African type I variant, AF472509 for the African type II variant, and AF402678 for the Asian American variant. The sequence of the HPV16 reference strain was also included in the present phylogenetic analysis.

Selection detection

The ratio of dN (nonsynonymous changes per nonsynonymous site) to dS (synonymous changes per synonymous site) was calculated in the two different reading frames in the overlapping portion of E2 and E4 genes using the SLAC (single likelihood ancestor counting) method in Datamonkey [37, 38] in order to examine the selective pressure acting on the two protein-coding genes. The median mutation rate in the first, second and third positions of the codon in the two different reading frames was determined by Bayesian analysis using Markov chain Monte Carlo (MCMC) and BEAST v.1.6.2 [39]. Bayesian analysis consisted of four chains with random starting trees. Results were examined using the program Tracer v1.5 from the BEAST package. Convergence was assessed with ESS (effective sample size) values after a burning of 2 million steps. However, the GTR nucleotide substitution model was used, as it was calculated using the online program FindModel (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.htm>).

Molecular evolutionary analysis of the overlapping region of the E2 and E4 genes was performed using the cloned E4 sequences from the present study, the reference sequence of the HPV16 genome, and 49 E2 gene sequences that included the complete E4 ORF, available in the GenBank sequence database under accession numbers

Table 1 Sequence variations and amino acid substitutions of 15 high-grade (HG) HPV16-positive cervical samples

CIN	E4 ORF	Variants																				Accession number in Genbank
		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
	CLONES	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
		6	6	6	7	8	1	1	4	7	7	1	1	3	5	5	5	5	5	5	5	
		0	1	2	7	7	0	3	9	6	6	6	7	8	6	6	6	6	6	6	6	
		G	C	A	C	C	T	C	G	C	C	C	C	A	C	C	A	C	C	T	T	
HG	A3 288 C1	-	-	G	G	G	C	T	A	A	-	-	A	C	-	-	-	-	-	-	-	
	A3 288 C2	-	T	G	G	G	C	T	A	A	-	-	A	C	-	-	-	-	-	-	-	
	A4 288 C1	-	-	G	G	G	-	T	-	A	-	-	A	-	-	-	-	-	-	-	-	
	A4 288 C2	-	-	G	G	G	-	T	-	A	-	-	A	-	-	-	-	-	-	-	-	
	A6 288 C1	-	-	-	-	-	-	T	-	-	-	-	A	-	-	-	-	-	-	-	-	
	A13 288 C1	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	A13 288 C2	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	A15 288 C1	-	-	-	G	C	-	-	A	A	T	-	A	C	-	-	-	-	-	-	-	
	A31 288 C1	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	
	D2 288 C1	-	-	-	-	-	-	T	-	-	-	-	A	-	-	-	-	-	-	-	-	
	D4 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	D14 288 C1	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	D16 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL4 288 C1	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL4 288 C3	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL5 288 C1	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL5 288 C2	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL6 288 C1	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL6 288 C2	C	-	G	G	C	C	T	-	A	-	-	A	C	-	-	-	-	-	-	-	
	PNL6 288 C3	-	-	-	-	-	-	T	A	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL7 288 C1	C	-	G	G	C	C	T	-	A	-	-	A	C	-	-	-	-	-	-	-	
	PNL7 288 C3	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PNL11 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	E4 protein	A7P	A7V	A7A	L12L	L16L	L16L	T23T	P24P	P33P	D45E	D45E	L59I	L59T	Q66P	S72F	H75Q					
	E2 protein	S202S	S202S	N2023D	P208A	I211T	I211T	P219S	A220T	E232K	P241S	P241T	T254N	T254N	S261S	I267I	F271V					

Nucleotide analysis of HPV16 E4 ORF in high-grade cervical samples

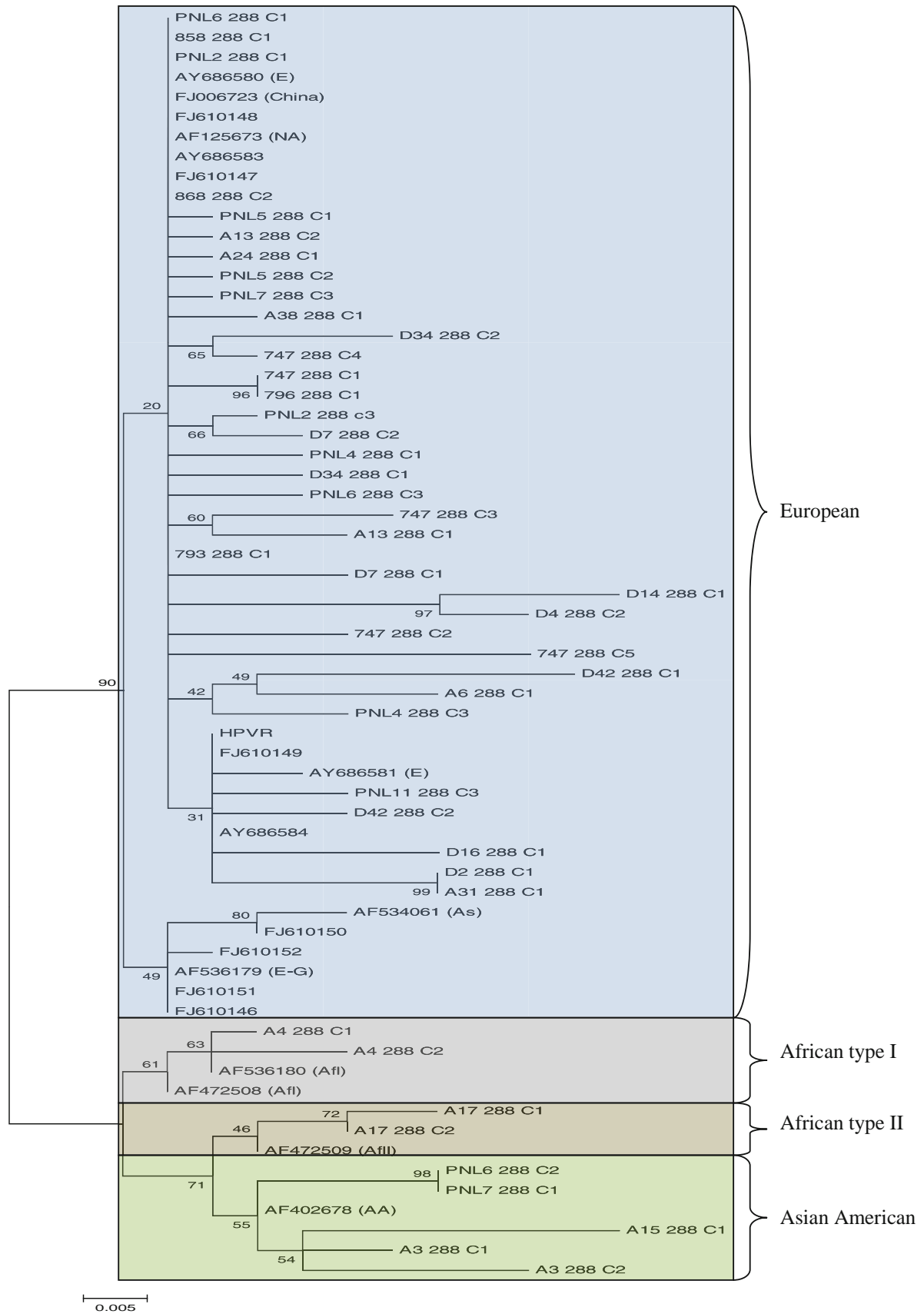
The corresponding HPV16 variants were calculated by phylogenetic tree analysis. Codes A3, A4, A6, A13, A15, D14, D34, D42, 747, 793, PNL2, PNL6, PNL7, etc. correspond to the cervical samples. The number 228 corresponds to the size of the E4 ORF, and the last letter (C1, C2, C3) corresponds to the number of the E4 cloned sequences. Fifteen nucleic acid variations were identified in the E4 ORF. Eight of the nucleotide variations resulted in amino acid substitutions in the E1⁺E4 protein. Six nucleotide variations that caused amino acid substitutions in the protein were located in the 3' region of the E4 gene. The amino acid substitutions that occurred in into the E2 protein are also presented

Table 2 Sequence variations and amino acid substitutions of 12 low-grade (LG) HPV16-positive cervical samples

CIN	E4 ORF	Variants												Accession number in Genbank			
		3	3	3	3	3	3	3	3	3	3	3	3				
LG	CLONES	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
	PNL2 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	JN032663
	PNL2 288 C3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	JN032664
	D7 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693895
	D7 288 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693896
	D34 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693897
	D34 288 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ683998
	D42 288 C1	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693899
	D42 288 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693900
	747 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693905
	747 288 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693906
	747 288 C3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702327
747 288 C4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702328	
747 288 C5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702329	
796 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702330	
858 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702331	
868 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702332	
793 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ702333	
A17 288 C1	-	T	G	A	A	T	-	A	-	A	-	A	-	-	-	HQ693901	
A17 288 C2	-	T	G	G	G	T	-	A	-	A	-	A	-	-	-	HQ693902	
A24 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693904	
A38 288 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	HQ693903	
E4 protein	A7P	A7V	A7A	L12L	L16I	T23T	P24P	P33P	R40R	D45E	L59I	L59T	Q66P	S72F	H75Q		
E2 protein	S20S	S20S	N2023D	P208A	I211V	P219S	A220T	E232K	T236M	P241T	T254N	T254N	S261S	I267I	F271V		

Nucleotide analysis of HPV16 E4 ORF in low-grade cervical samples

The corresponding HPV16 variants were calculated by phylogenetic tree analysis. Codes PNL2, D7, D42, D34, 793, 796, 858, 868, A17, A24, and A38 correspond to the cervical samples. The number 228 corresponds to the size of the E4 ORF, and the last letter (C1, C2, C3) corresponds to the number of the cloned E4 sequences. The nucleotide variation T3387A in low-grade cervical samples resulted in the amino acid substitution L16I. The European HPV16 variants had a restricted number of nucleotide variations. Amino acid substitutions that occurred in the E2 protein are presented



◀ **Fig. 1** The phylogenetic tree was calculated by the maximum-likelihood method (1000 bootstrap replicates) using MEGA v.5. The GTR substitution model was selected. The phylogenetic tree was inferred from Muscle alignment of cloned E4 sequences, the prototype sequence of HPV16, and those of representative sequences of HPV16 variants. The phylogenetic tree construction based on the E4 gene revealed distinct phylogenetic branches. Each individual branch represents a specific HPV16 variant. The results of phylogenetic analysis revealed that the low-grade cervical sample A17 and cloned sequences derived from high-grade cervical samples (A4 288 C1, A4 288 C2, A3 288 C1, A3 288 C2, A15 288 C1, PNL6 288 C2 and PNL7 288 C1) were associated with non-European HPV16 variants

AF472509, AF402678, AF472508, AF536180, AF534061, AF536179, FJ610150, FJ610151, FJ610147, AY686583, FJ610149, FJ610146, AY686581, AF125673, FJ610152, FJ610148, FJ006723, AY686584, AY686580, AF193429, DQ191463, AF067034, AF193425, AF193427, AF115832, DQ191465, AY686579, NC001526, AY686582, AF067035, AF115831, EU430675, EU430667, EU430683, EU118173, AF515620, AF067029, DQ191462, AF193430, AF193426, AF067032, AF067028, AF067030, K02718, DQ191461, AF193428, AF067031, EU918764, and HM057182.

Results

E4 sequence analysis

In the present report, 27 HPV16 samples were investigated in order to detect sequence variations in the E4 ORF. Multiple sequence analysis of the cloned E4 sequences and the reference sequence of the HPV16 strain was carried out using MEGA v.5 [35] with the Muscle [36] algorithm. The numbering of the nucleotide variations was based on the HPV16 genome sequence, which was obtained from the HPV16 Sequence Database (Los Alamos National Laboratory).

The nucleotide variations at positions 3360 and 3361 caused the amino acid substitutions A7P and A7V, respectively, in the N-terminal domain of the E1[^]E4 protein (Tables 1, 2). The G3360C variation was identified in two high-grade cervical cases (PNL6 and PNL7) and in one low-grade cervical sample (D42). The C3361T variation was detected in the low-grade cervical case A17 and in the high-grade cervical samples A3 and A4 (Tables 1, 2). Additionally, six silent nucleotide variations were identified at positions 3362, 3377, 3387, 3410, 3413 and 3449 of the E4 ORF (Tables 1, 2). Nevertheless, two different nucleotide changes at position 3387 were observed. Nucleotide variation T3387C (A3, A15, PNL6, PNL7) does not result in an amino acid substitution, while T3387A (A4, A17) causes the amino acid change L16I (Tables 1, 2).

However, two nucleotide variations were detected at positions 3476 and 3516, resulting in the amino acid

changes D45E and L59I, respectively (Table 1). The variation C3476A was identified in two high-grade clinical cases (A31 and D2) (Tables 1, 2). In the same portion of the high-grade sample A15, the nucleotide change C3476T failed to induce an amino acid change in the E1[^]E4 protein (Table 1). The C3516A variation leads to an Ile substitution in amino acid 59 in the cervical samples A4 and A6. In the remaining four high-grade specimens A3 and A15, and in one low-grade case, A17, the C3516A was found in combination with the nucleotide variation T3517C, causing the amino acid change L59T (Tables 1, 2).

In the C-terminal domain of the E1[^]E4 protein, which is responsible for keratin collapse and the oligomerisation of the protein, three nucleotide variations at positions 3538, 3556 and 3566 were detected. The A3538C variation induces the amino acid substitution Q66P, which was identified in cervical samples (A3, A4, A15, PNL6, PNL7, and A17) that were related to non-European variants (Tables 1, 2; Fig. 1). The C3556T variation causes the amino acid substitution S72F in the high-grade samples A4, PNL6 and PNL7 (Table 1). This variation was strongly associated with the African type I and the Asian American variants (Table 1; Fig. 1). However, the T3566G variation causes the amino acid change H75Q in one low-grade (A17) and three high-grade (A3, A4, and A15) cervical intraepithelial cases (Table 1). Considering the phylogenetic analysis (Table 1; Fig. 1), amino acid substitutions in the C-terminal domain were strongly associated with non-European variants such as the HPV16 Asian American, African type I and African type II variants (Tables 1, 2; Fig. 1).

Nucleic acid mutations that were detected in the present report were identified in cervical samples derived from Thin Prep samples and stained cervical smears. Three variations at positions 3360, 3410 and 3476 were also detected in paraffin-embedded cervical tissue. Taking these data into account, the probability that these nucleic acid mutations were induced by formalin fixation must be excluded.

Evolutionary analysis of the overlapping region of E4 and E2

E4 ORF is entirely contained within the E2 ORF. Results derived from molecular evolution analysis showed that in the E4 reading frame the dN/dS = 0.97 and the relative substitution rates for all three codon positions was CP1 = 0.929, CP2 = 1.023, CP3 = 1.048. In the E2 reading frame of the overlapping portion of the E2 and E4 genes, the dN/dS ratio was 1, and the relative substitution rates for all three codon positions were CP1 = 1.056, CP2 = 0.927 and CP3 = 1.018. The dN/dS ratio was 1 for both the E2 and E4 reading frames. Taking these data into

account, it was assumed that neutral selection dominated in the overlapping fragment of the E2 hinge region and the E4 gene.

Phylogenetic clustering of E4 amplicons

In previous studies, the amplified HPV16 sequences were classified with the HPV16 variants according to sequence variations in the E6, E7, L1, E2 ORFs and the long control region (LCR) [7–12]. In order to investigate the possible association between nucleotide variations in the E4 ORF and HPV16 variants, E6 and E7 nucleotide analysis was performed on the same cervical samples (Table 3). Nevertheless, seven cervical cases (D2, D4, D7, D14, D16, D34, D42) have been associated with the European HPV16 variant by nucleotide and phylogenetic analysis of the E2 gene, as we have reported previously (Table 3) [11].

Nucleotide and phylogenetic analysis of cloned HPV16 E4 sequences showed that the E4 ORF may be used to classify the HPV16 genome into distinct phylogenetic branches like those of the E6 and E7 genes. In particular, nucleotide sequence analysis of the E4, E6 and E7 genes revealed that isolates from six high-grade and seven low-grade cervical samples were related to the European variant (T350G), while isolates from the cervical samples PNL11 and 796 were related to the reference strain. In addition, the high-grade cervical samples A3, A15 and PNL7 (Tables 1, 3) showed sequence homology to the HPV16 Asian American variant, while the high-grade cervical sample A4 and the low-grade sample A17 showed sequence similarity with the African type I and African type II variants, respectively (Tables 1, 2, 3).

Nucleotide and phylogenetic analysis of HPV16 E4 also revealed that different cloned sequences derived from each individual high-grade cervical case, PNL6 and PNL7, clustered in different phylogenetic branches. In particular, the cloned sequences PNL6 288 C2 and PNL7 288 C1 clustered in the same phylogenetic branch with the Asian American variant, while the other cloned sequences (PNL6 288 C1 and PNL7 288 C3) derived from the same cervical samples clustered with the European variant (Fig. 1) revealing that these cervical samples might carry multiple viral infections by Asian American and European variants.

Discussion

The E4 ORF is located entirely within the E2 gene and is expressed during the late phase of infection [40]. In the present report, a total of 15 nucleic acid variations were detected in the HPV16 E4 ORF in high- and low-grade cervical cases (Tables 1, 2). The initial study by Eriksson et al. [41], which described HPV16 variants in this region,

did not report the amino acid mutations that were reported in the present study. According to sequence analysis of the E4 ORF, two amino acid changes were detected in the N-terminal domain of the E1^{E4} protein at amino acid position 7 (A7V and A7P). Previous studies revealed that the N-terminal domain was necessary for E1^{E4} protein attachment in a keratin network, through a conserved leucine motif (LLXLL) [23, 27, 42]. In addition, McIntosh et al. [27] suggested that the N-terminal domain also contributes to the regulation of protein function in the cell. The significance of these changes was also reinforced by Wang et al. [43], who proposed that the effective LLXLL keratin interaction may be strongly related to the integrity of the N-terminal 16 amino acids of the HPV16 E1^{E4} protein.

The protein region from amino acids 17 to 45 has been termed the “arrest domain” as is implicated in G2 arrest of the cell cycle [28, 29]. However, the “arrest domain” includes a single cyclin-dependent kinase phosphorylation site at T22 and T23, and a point mutation at amino acid T23 interferes with G2 arrest [28, 29]. In the present study, the nucleotide variation C3410T was detected in 21 cervical samples (low and high grade) (Tables 1, 2) but did not cause an amino acid substitution at the Cdk phosphorylation site T23. Based on a multiple sequence alignment, it was concluded that the phosphorylation site (T22, T23) is strongly conserved between the different characterized HPV16 isolates (Tables 1, 2). Two nucleotide variations at positions 3516 and 3517 were detected in the E4 ORF. The covariation of C3516A and T3517C leads to the amino acid change L59T (Tables 1, 2). These mutations were located between the N-terminal and C-terminal domains of the protein. Phylogenetic analysis revealed that the L59T mutation was encountered in HPV16 African type II and Asian American variants (Tables 1, 2).

The C-terminal domain of the HPV16 E1^{E4} protein is implicated in the disruption of keratin filaments and in the formation of multimeric complexes of the protein [30]. In this region, three amino acid changes (Q66P, S72F and H75Q) were identified among high- and low-grade cervical samples (Tables 1, 2). However, the Q66P and S72F mutations were strongly associated with non-European variants. The role of these mutations in the C-terminal domain of the protein must be further investigated in order to clarify the influence of point mutations in the oncogenic potential of the virus. Taking these data into consideration, it is proposed that the nucleotide variations A3362G, C3377G, G3413A, C3476T/A, C3516A, T3517C, A3538C, C3556T and T3566G are related to high-grade cervical malignancies. Nevertheless, further studies are required in order to estimate the role of these particular changes in progression of malignancy.

Table 3 Nucleotide sequence variations in the E6 and E7 genes that were identified in the examined cervical samples

Samples	E6 region										E7 region							Variants	Accession number in Genbank	E2 region		Accession number in Genbank
																				Samples	Variants	
	1	1	1	1	1	1	2	2	3	3	4	5	6	7	7	7	7					
A3	-	-	-	-	T	-	A	G	T	-	-	G	-	C	C	C	G	Asian American	JQ796772	D2	European	HQ398221
A4	-	C	G	T	-	A	G	T	-	-	-	-	-	-	C	C	G	African type I	JQ422059	D4	European	HQ398219
A6	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422097	D14	European	HQ398220
A13	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422072	D16	European	HQ398222
A15	-	-	-	-	T	-	A	G	T	G	-	-	-	C	C	G	Asian American	JN205116	D7	European	JF422083	
A31	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422098	D34	European	JF422087
PNL4	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422080	D42	European	JF422091
PNL5	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422082			
PNL6	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422083			
PNL7	-	-	-	-	T	-	A	G	T	G	-	G	-	C	C	G	Asian American	JQ422085				
PNL11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Prototype	JQ422094			
PNL2	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	European	JQ422078			
747	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422116			
796	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Prototype	JQ422099			
858	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422101			
868	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422103			
793	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422106			
A17	-	T	G	T	-	A	G	T	-	G	-	G	-	C	C	G	African type II	JQ422056				
A24	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JQ422075			
A38	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	European	JN205111			
European	G	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-					
African type I	-	C	G	T	-	A	G	T	-	-	-	-	-	-	C	C	G					
African type II	-	T	G	T	-	A	G	T	-	G	-	G	-	C	C	G						
Asian American	-	-	-	-	T	-	A	G	T	G	-	G	-	C	C	G						
Asian	-	-	-	-	T	G	-	-	-	-	-	-	-	G	-	-	-					

Nucleotide analysis of E6, E7 and E2 genes

These changes were used to describe the distinct HPV16 variants, and they were associated with nucleotide sequence variations in the E4 ORF. However, the cervical samples D2, D4, D14, D16, D7, D34 and D42 have been associated with the European variant in a previous study [10]

Phylogenetic and nucleotide analysis of HPV16 E6, E7 and E4 sequences from the same samples revealed that specific nucleotide variations in the E4 region were associated with intratypic nucleotide polymorphisms into E6 and E7 genes and clustered HPV16 E4 sequences into distinct phylogenetic branches [7–9, 11] (Tables 1, 2, 3). In particular, in the E4 ORF, the nucleotide variations A3362G, C3377G, G3449A, C3516A, T3566G were related to sequence changes in the E6 and E7 genes (G145T, T286A, A289G, C335T, T789C, T795G) that are encountered in the HPV16 African type I, African type II and Asian American variants (Tables 1, 2, 3). However, the nucleotide changes T3517C, A3538C were identified in the African type II and Asian American variants, while the nucleotide variation C3556T was detected in Asian American and African type I variants. In addition, the nucleotide variations T3387C and G3413A were associated with the E6, E7 nucleotide changes A532G and T732C, which converge only in the HPV16 Asian American genome (Tables 1, 2, 3). Taking these data into account, it was concluded that although the E4 ORF is a small region of the viral genome (288 bp) it can be used to classify the HPV16 variants into distinct phylogenetic branches. Nucleotide and phylogenetic analysis of the E4 ORF revealed that non-European variants such as the African type I and Asian American variants were preferentially detected in high-grade cervical malignancies [13–15]. Non-European variants (African type I, African type II and Asian American) are circulating in the Greek population, and results derived from the phylogenetic study revealed that the cervical samples PNL6 and PNL7 contained both HPV16 Asian American and European variants (Fig. 1).

However, the nucleotide variations that were recorded for the E4 gene resulted in nucleotide variations and amino acid substitutions in the E2 hinge region (Tables 1, 2). In order to investigate the selective pressure acting on the two distinct overlapping reading frames, molecular evolution analysis was performed using the cloned E4 sequences and 49 E2 sequences available in GenBank sequence database that included the overlapping fragment of the E2 hinge region and the E4 ORF. Our results indicate that neutral selection is acting on the two overlapping reading frames. In contrast, previous reports have proposed that positive diversifying selection on the E2 hinge region coexists with purifying selection on the overlapping fragment of the E4 gene [44, 45].

In conclusion, little information about the HPV16 E4 ORF is known. HPV16 variant determination is carried out using mainly the LCR and L1, and E6 and E7 coding regions. Our data presented in Tables 1, 2 prove that the E4 ORF is a polymorphic region when different HPV16 variants are compared, and although the E4 ORF constitutes a small portion of the viral genome (288 bp),

sequence variability in the E4 ORF may provide valuable information about the origins of HPV16 variants.

Acknowledgments This work was supported by research grants of the Postgraduate Programmes ‘Applications of Molecular Biology-Genetics, Diagnostic Biomarkers’, code 3817, and “Biotechnology”, code 3439, of the University of Thessaly, School of Health Sciences, Department of Biochemistry and Biotechnology.

Conflict of interest All authors declare that they have no conflicting or dual interests.

References

- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79
- Sundström K, Eloranta S, Sparén P, Arnheim Dahlström L, Gunnell A, Lindgren A, Palmgren J et al (2010) Prospective study of human papillomavirus (HPV) types, HPV persistence, and risk of squamous cell carcinoma of the cervix. *Cancer Epidemiol Biomarkers Prev* 19:2469–2478
- Castle PE, Schiffman M, Wheeler CM, Wentzensen N, Gravitt PE (2010) Human papillomavirus genotypes in cervical intraepithelial neoplasia grade 3. *Cancer Epidemiol Biomarkers Prev* 19:1675–1681
- Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX, de Sanjosé S (2010) Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *J Infect Dis* 202:1789–1799
- Kedzia W, Józefiak A, Pruski D, Rokita W, Marek S (2010) Human papilloma virus genotyping in women with CIN 1. *Ginekol Pol* 81:664–667
- Tomesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Reggi S, Buonaguro L, Buonaguro FM (2011) Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 121:32–42
- Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM (1997) Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* 71:2463–2472
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743–7753
- Wheeler CM, Yamada T, Hildesheim A, Jenison SA (1997) Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* 35:11–19
- Tsakogiannis D, Ruether IG, Kyriakopoulou Z, Pliaka V, Theoharopoulou A, Skordas V, Panotopoulou E, Nepka C, Markoulatos P (2012) Sequence variation analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from women in Greece. *Arch Virol*. doi:10.1007/s00705-012-1236-8
- Swan DC, Rajeevan M, Tortolero-Luna G, Follen M, Tucker RA, Unger ER (2005) Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* 96:695–700
- Eriksson A, Herron JR, Yamada T, Wheeler CM (1999) Human papillomavirus type 16 variant lineages characterised by

- nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* 80:595–600
13. Slichero L, Ferreira S, Trotter H, Duarte-Franco E, Ferenczy A, Franco EL, Villa LL (2007) High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *Int J Cancer* 120:1763–1768
 14. Sanchez GI, Kleter B, Gheit T, van Doorn LJ, de Koning MN, de Sanjosé S et al (2011) Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *J Clin Virol* 51:165–169
 15. Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Greggi S et al (2011) Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 121:32–42
 16. Lehoux M, D'Abramo CM, Archambault J (2009) Molecular mechanisms of human papillomavirus-induced carcinogenesis. *Public Health Genomics*. 12:268–280
 17. Doorbar J (2007) Papillomavirus life cycle organization and biomarker selection. *Dis Markers* 23:297–313
 18. Snijders PJ, Steenbergen RD, Heideman DA, Meijer CJ (2006) HPV-mediated cervical carcinogenesis: concepts and clinical implications. *J Pathol* 208:152–164
 19. Münger K, Baldwin A, Edwards KM, Hayakawa H, Nguyen CL, Owens M, Grace M, Huh K (2004) Mechanisms of human papillomavirus-induced oncogenesis. *J Virol* 78:11451–11460
 20. Fehrmann F, Laimins LA (2003) Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene* 22:5201–5207
 21. Schwartz S (2008) HPV-16 RNA processing. *Front Biosci* 13:5880–5891
 22. Schwartz S, Zhao S, Öberg D, Rush M (2004) Regulation of papillomavirus late gene expression. *Recent Res Dev Virol* 6:29–45
 23. Doorbar J, Foo C, Coleman N, Medcalf L, Hartley O, Prospero T et al (1997) Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. *Virology* 238:40–52
 24. Middleton K, Peh W, Southern S, Griffin H, Sotlar K, Nakahara T et al (2003) Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J Virol* 77:10186–10201
 25. Chow LT, Broker TR (2007) Human papillomavirus RNA transcription. In: Garcea RL, DiMaio D (eds) *The Papillomaviruses*. Springer, New York, pp 109–144
 26. Howley PM (1996) Papillomavirinae: the viruses and their replication. In: Fields BN, Knipe DM, Howley PM (eds) *Fields virology*, 3rd edn. Lippincott-Raven, Philadelphia, pp 2045–2076
 27. McIntosh PB, Laskey P, Sullivan K, Davy C, Wang Q, Jackson DJ, Griffin HM, Doorbar J (2010) E1–E4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium. *J Cell Sci* 123:2810–2822
 28. Davy CE, Jackson DJ, Wang Q, Raj K, Masterson PJ, Fenner NF et al (2002) Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16. *J Virol* 76:9806–9818
 29. Davy CE, Jackson DJ, Raj K, Peh WL, Southern SA, Das P, Sorathia R et al (2005) Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes. *J Virol* 79:3998–4011
 30. Roberts S, Kingsbury SR, Stoerber K, Knight GL, Gallimore PH, Williams GH (2008) Identification of an arginine-rich motif in human papillomavirus type 1 E1;E4 protein necessary for E4-mediated inhibition of cellular DNA synthesis in vitro and in cells. *J Virol* 82:9056–9064
 31. Davy C, McIntosh P, Jackson DJ, Sorathia R, Miell M, Wang Q, Khan J, Soneji Y, Doorbar J (2009) A novel interaction between the human papillomavirus type 16 E2 and E1–E4 proteins leads to stabilization of E2. *Virology* 394:266–275
 32. Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25–36
 33. Puranen M, Saarikoski S, Syrjänen K, Syrjänen S (1996) Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391–395
 34. Sotlar K, Diemer D, Dethleffs A, Hack Y et al (2004) Detection and typing of human papillomavirus by E6 nested multiplex PCR. *J Clin Microbiol* 42:3176–3184
 35. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
 36. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
 37. Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533
 38. Suzuki Y, Gojobori T (1999) A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* 16:1315–1328
 39. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
 40. Doorbar, J and Myers G (1996) The E4 protein. In: Myers HDG, Icenogele J (eds) *Human papillomaviruses*, vol. 3. Los Alamos National Laboratory, Los Alamos, NM, pp 58–80
 41. Eriksson A, Herron JR, Yamada T, Wheeler CM (1999) Human papillomavirus type 16 variant lineages characterized by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* 80:595–600
 42. Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH (1994) Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J Virol* 68:6432–6445
 43. Wang Q, Griffin H, Southern S, Jackson D, Martin A et al (2004) Functional analysis of the human papillomavirus type 16 E1–E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J Virol* 78:821–833
 44. Hughes AL, Hughes MA (2005) Patterns of nucleotide difference in overlapping and non-overlapping reading frames of papillomavirus genomes. *Virus Res* 113:81–88
 45. Burk RD, Chen Z, Van Doorslaer K (2009) Human papillomaviruses: genetic basis of carcinogenicity. *Public Health Genomics* 12:281–290

Identification of novel E6-E7 sequence variants of human papillomavirus 16

D. Tsakogiannis · Z. Kyriakopoulou ·
G. Amoutzias · I. G. A. Ruether · T. G. Dimitriou ·
E. Panotopoulou · P. Markoulatos

Received: 10 May 2012 / Accepted: 16 October 2012 / Published online: 4 December 2012
© Springer-Verlag Wien 2012

Abstract The rate of evolution of the human papillomavirus 16 (HPV16) genome is low. However, the ability of the E6 oncoprotein to interact with distinct p53 variants causes selective pressure on the E6 gene. In addition, intratypic recombination events in the HPV16 E6 and E7 genes have been characterized as extraordinary phenomena during the evolutionary history of virus. In the present study, we identified two new sequence variants through nucleotide analysis of the E6-E7 region of the HPV16 genome. Maximum-likelihood and empirical Bayesian methods were used in order to identify positive selection at particular residues of the E6 and E7 genes. Using the single recombination breakpoint (SBP) method, we found evidence of recombination events in the E6 ORF. Nucleotide sequence analysis showed that the new sequence variants are phylogenetically distant from the other members of the population. Our results indicate that new evolutionary intermediates of HPV16 might be formed either through positive selective pressure or through recombination events by multiple infections with distinct HPV16 variants.

Introduction

Human papillomaviruses (HPVs) are double-stranded, circular DNA viruses that infect cutaneous or mucosal epithelial cells [1]. HPVs are classified into different types and variants according to sequence homology of the L1 capsid gene [2, 3]. Intratypic genomic heterogeneity has been mainly described in HPV16. In particular, the study of specific viral domains (E6, E7, E2, L1, L2 and LCR) has revealed a number of intratypic nucleotide polymorphisms in the HPV16 genome. These sequence polymorphisms have been identified in different geographical locations and have been associated with the ethnicity of the studied populations. The geographical relatedness of the virus resulted in a further classification of HPV16 into five distinct molecular variants known as the HPV16 European, African type II, African type I, Asian American and Asian variants [4–10]. The best-defined region of the viral genome, which has been used to describe distinct HPV16 variants, is the E6 oncogene [4–8, 11, 12].

The results of molecular evolution analysis have suggested that the HPV16 genome has a low evolutionary rate [13]. Nevertheless, the ability of the E6 oncoprotein to interact with distinct p53 variants causes selective pressure on the E6 gene [14]. Carvajal-Rodríguez [15] proposed that coevolution with cell-cycle regulators drives divergence in the E6 genes within the HPV16 group. However, a molecular mechanism that also contributes to HPV16 genome evolution is intratypic recombination. Previous studies that used human alpha PVs sequences from the public database at Los Alamos National Laboratory reported evidence that recombination signals are located in the E6, E7, L2 and L1 genes, mainly within the high-risk HPV16 [15, 16]. Nevertheless, the first intratypic recombination in natural infection was described by Jiang et al. in

D. Tsakogiannis · Z. Kyriakopoulou · G. Amoutzias ·
I. G. A. Ruether · T. G. Dimitriou · P. Markoulatos (✉)
Microbiology-Virology Laboratory, Department of Biochemistry
and Biotechnology, School of Health Sciences,
University of Thessaly, Ploutonos 26 and Aiolou,
41221 Larissa, Greece
e-mail: markoulatos@bio.uth.gr

E. Panotopoulou
Virology Laboratory, Anticancer Oncology Hospital of Athens
“St Savvas”, Papanicolaou Research Centre of Oncology
and Experimental Surgery, Athens, Greece

2009 [17]. This was the first report that recombinant sequences in the HPV16 genome were detected in a concurrent infection, showing the genomic variability of HPV16 E6 and LCR genes.

In the present report, HPV16-positive cervical samples, diagnosed as normal, high- and low-grade cervical intraepithelial neoplasia, were assayed in order to investigate the E6-E7 genomic region, looking into the nucleotide variability between different HPV16 variants circulating in the Greek population. According to the nucleotide sequence analysis, novel HPV16 E6-E7 sequence variants were present in three distinct cervical samples.

Materials and methods

Sample collection and DNA isolation

In the present report, HPV16-positive cervical samples from Thin Prep and stained cervical smears were received from the Virology Laboratory of Anticancer Oncology Hospital of Athens “St Savvas”. DNA from Thin Prep specimens was isolated following the method described by Casas et al. [18]. DNA from cervical stained smears was extracted according to Puranen et al. [19]. The quality of the extracted DNA was assayed by amplification of a 498-bp fragment of the human β -actin gene [20]. HPV typing was performed using nested multiplex PCR as described previously by Sotlar et al. [21].

E6-E7 amplification by PCR

HPV16-positive samples were subjected to PCR amplification of the HPV16 E6-E7 genes using the primer set HPV16 41, 5'-AGCGACCCAGAAAGTTACCA-3'/ HPV16 757, 5'-ATGGGGCACACAATTCCTAG-3'. The primers were designed using the Primer3 program, and the nucleotide positions were numbered according to the nucleotide sequence of the prototype HPV16 strain available in the NCBI GenBank sequence database (<http://www.ncbi.nlm.nih.gov/>) under accession number NC_001526. The set of primers was designed to amplify both the HPV16 E6 and E7 ORFs. PCR was performed in a final volume of 50 μ l. The PCR mixture contained 50 pmol of each primer, 5X Green GoTaq Flexi buffer (Promega), 2 mM MgCl₂, 1 mM dNTPs and 1.25 U of thermostable DNA polymerase (GoTaq[®] DNA Polymerase, Promega). The cycling conditions were as follows: 40 cycles of 40 s at 95 °C, 50 s at 55 °C and 1 min at 72 °C. The first cycle was preceded by a 2-min denaturation step at 95 °C, and the last cycle was followed by a 5-min elongation step at 72 °C.

Cloning and sequencing of E6-E7 genes

The 715-bp amplicon of the HPV16 E6-E7 genes was subjected to cloning using pGEM-T Easy Vector (Promega, Madison, USA). Recombinant plasmid DNA was extracted using a Nucleospin Plasmid Kit (Macherey-Nagel GmbH, Duren, Germany), and the plasmids were subjected to sequencing at MacroGen Inc., Korea. In order to confirm the sequence variations that were detected in the E6 and E7 ORFs, three independent PCR assays were performed, and amplicons were subjected to cloning. Two to three colonies were collected from each individual specimen, and both strands of all tested clones were sequenced. The DNA strands were sequenced three times to insure minimum reliability of results. In order to identify HPV16 variants that were circulating in the Greek population, a multiple sequence alignment was made using the program MEGA v.5 [22] and the Muscle algorithm [23]. The alignment included the cloned HPV16 E6-E7 sequences, the reference sequence of HPV16 (HPV16R) available in the HPV16 Sequence Database (Los Alamos National Laboratory), and the corresponding sequences of the following HPV16 variants: African type I, African type II, Asian-American, European and Asian, available on line in the GenBank sequence database under accession numbers AF536180, AF472509, AF402678, AF536179 and AF534061, respectively (Table 1).

E6-E7 selective pressure

To investigate the selective pressure acting on specific codons in the E6 and E7 genes, maximum-likelihood (ML) and empirical Bayesian (EB) methods were used to examine the new E6-E7 sequences and those of representative sequences of HPV16 variants that were described in previous work by Carvajal-Rodríguez [15] (accession numbers AF125673, AY686583, AY686580, AY686584, AF536179, AF534061, AF472508, AF536180, AF472509, AF402678, AY686579, and AY686582). The maximum-likelihood method was used using the fixed effects likelihood model (FEL) on the Datamonkey server [24, 25]. Empirical Bayesian analysis was performed using the program Selection 2.4 [26] at the highest precision level. In order to identify evidence of recombination, the single recombination breakpoint (SBP) method [27] was used in the current study. In addition, the GTR + G model was used to describe the process of nucleotide substitution according to AIC criterion using the online program FindModel (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>). Bootscan analysis was incorporated in the present study using the program SimPlot v.3.5.1. In addition, ClustalW2 analysis was carried out in order to determine the percentage of sequence homology

Table 1 Nucleotide sequence analysis of the cervical samples A15, A38 and 1143. The first letter or number corresponds to the code of the cervical sample (A15, A38 and 1143). The character E6-E7 corresponds to the genomic portion that was amplified. The last letter, C1, C2, or C3, corresponds to the number of the cloned sequence. Specific nucleotide polymorphisms that were associated with particular HPV16 variants are presented. Cervical samples A38 and 1143

carried the nucleotide changes G132T and C143G, which were detected in the African type II variant, and the nucleotide change T350G, which was identified exclusively in European variants. In addition, cervical sample A15 showed sequence homology between nucleotides 123 and 350 with the Asian American variant, and between nucleotides 350 and 839, with the African type I variant

	E6 region								E7 region							Accession number in GenBank
	1	1	1	1	2	2	3	3	4	5	6	7	7	7		
CLONES	3	3	4	4	8	8	3	5	0	3	4	3	8	9		
	1	2	3	5	6	9	5	0	3	2	7	2	9	5		
	A	G	C	G	T	A	C	T	A	A	A	T	T	T		
A15 E6-E7 C1	-	-	-	T	A	G	T	G	-	-	-	-	C	G	JN205116	
A15 E6-E7 C2	-	-	-	T	A	G	T	G	-	-	-	-	C	G	JN205117	
A38 E6-E7 C1	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205111	
A38 E6-E7 C2	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205112	
A38 E6-E7 C3	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205113	
1143 E6-E7 C1	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205114	
1143 E6-E7 C2	-	T	G	-	-	-	-	G	-	-	-	-	-	-	JN205115	
European	G	-	-	-	-	-	-	G	-	-	-	-	-	-		
African type I	-	C	G	T	A	G	T	-	-	-	-	-	C	G		
African type II	-	T	G	T	A	G	T	-	G	-	G	-	C	G		
Asian American	-	-	-	T	A	G	T	G	-	G	-	C	C	G		
Asian	-	-	-	-	-	-	-	-	-	-	G	-	-	-		

between the E6-E7 amplicons and the representative sequences of the HPV16 variants (<http://www.ebi.ac.uk/Tools/msa/clustalw2/analysis>).

Phylogenetic trees

Maximum-likelihood, maximum-parsimony and Bayesian phylogenetic trees were constructed using one representative sequence from the newly identified variant, the reference sequence of the HPV16 genome (HPV16R), and representative sequences of the HPV16 variants described above in order to determine the position of the new E6-E7 genes in the major HPV16 phylogenic branches [15] (Fig. 1). The maximum-likelihood phylogenetic tree (1000 bootstrap replicates) was constructed using the program MEGA v.5 [21] (Fig. 1). Maximum parsimony (MP) analysis was carried out using the close-neighbor-interchange algorithm to search for the optimal tree, and confidence in the nodes was evaluated by 1000 bootstrap replicates using the program MEGA v.5 [21] (Fig. 1). In addition, a Bayesian analysis using the Markov chain Monte Carlo (MCMC) method was also carried out on the E6-E7 dataset using BEAST v.1.6.2 [28]. Bayesian analysis consisted of four chains with random starting trees. The results were examined using the program Tracer v.1.5 from the BEAST package. Convergence was assessed with ESS (effective sample size) values, after a burning of

500,000 steps. The Bayesian phylogenetic tree was constructed using the programs TreeAnnotator V.1.6.2 and FigTree v.1.3.1 from the BEAST package. The first 1000 trees were discarded as burn-in, and the remaining trees were used to construct a 50 % majority-rule consensus tree on which posterior probability scores are indicated (Fig. 1). Bayesian analysis was performed using the GTR + G substitution model and was calculated using the online program FindModel (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>).

Results

Characterization of new HPV16 E6-E7 gene variants

Phylogenetic classification of HPV16 genomes in distinct branches has usually been done based on specific nucleotide changes in the E6 and E7 genes [4–8] as follows: prototype strain (no changes), European variant E-G350 (T350G), E-C109G (T109C, T350G), E-G131G (A131G, T350G), East Asian (T178G, A647G), African type I (G132C, C143G, G145T, T286A, A289G, C335T, T789C, T795G), African type II (T109C, G132T, C143G, G145T, T286A, A289G, C335T, G403G, A647G, T789C, T795G), Asian American (G145T, T286A, A289G, C335T, T350G, A532G, T732C, T789C, T795G). In the present study,

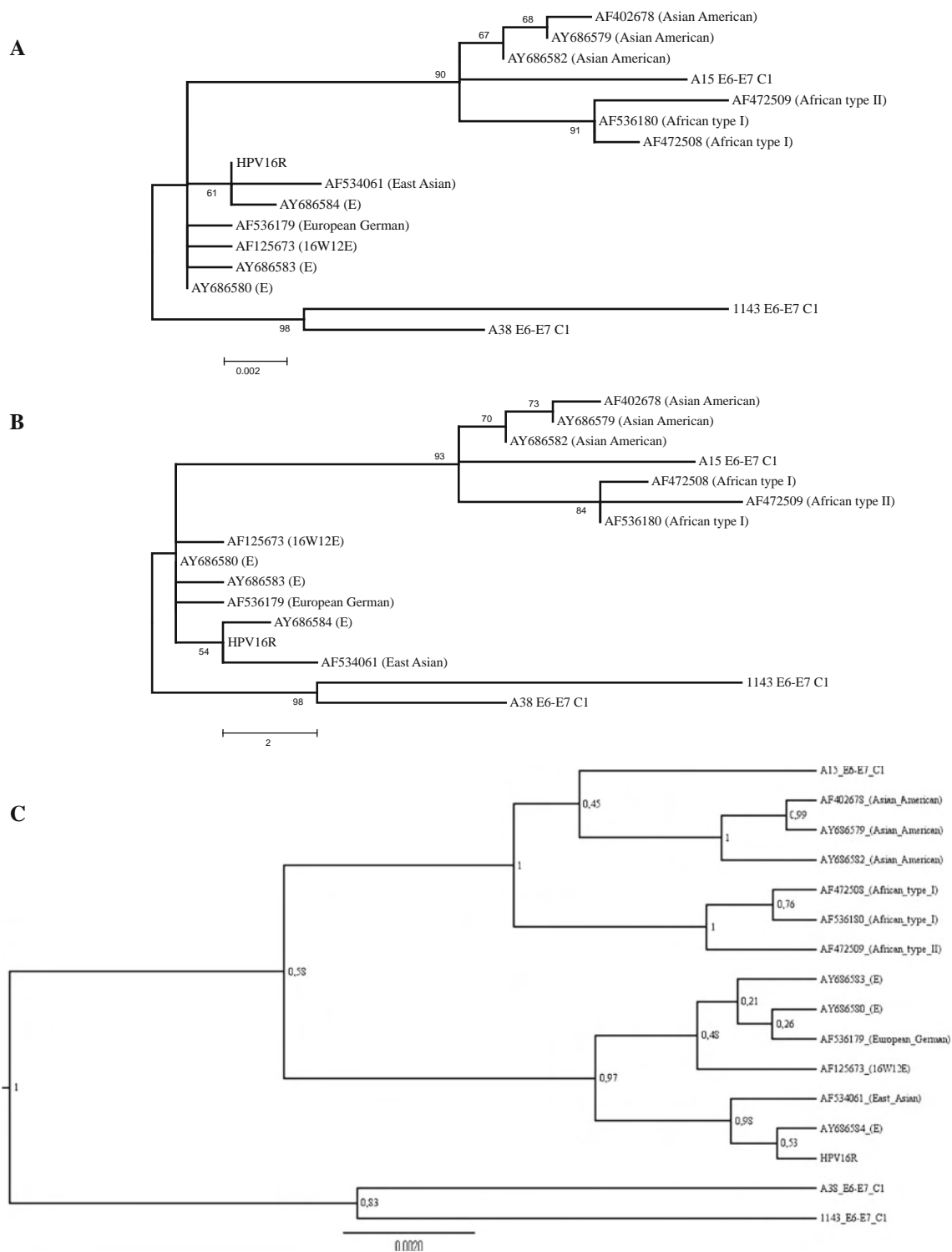


Fig. 1 Phylogenetic analysis of the new HPV16 E6-E7 sequence variants, the reference sequence of the HPV16 genome, and representative sequences of HPV16 variants that were described in a previous study by Carvajal-Rodríguez [15]. **A)** A maximum-likelihood phylogenetic tree (1000 bootstrap value) was constructed using the program MEGA v.5. **B)** Maximum-parsimony analysis was carried out using the close-neighbor-interchange algorithm to search for the optimal tree, and confidence in the nodes was evaluated using

1000 bootstrap replicates in the program MEGA v.5. **C)** Bayesian phylogenetic tree constructed using the programs TreeAnnotator V1.6.2 and FigTree v1.3.1 from the BEAST package, on which posterior probability scores are indicated. Maximum-likelihood, maximum-parsimony and embryrical Bayesian phylogenetic trees showed that the viruses in cervical samples A38 and 1143 are phylogenetically distant from the established phylogenetic branches of HPV16 variants

nucleotide sequence analysis of E6-E7 genes from three cervical samples (A15, A38 and 1143) revealed two novel sequence variants that were circulating in the Greek population (Table 1).

Nucleotide sequence analysis revealed that the virus in the high-grade cervical sample A15 had sequence homology to two distinct HPV16 variants in different positions of the E6-E7 sequence. In particular, between nucleotides 123-350, the cervical sample A15 showed nucleotide homology to Asian American variant (G145T, T286A, A289G, C335T, T350G), with maximum sequence similarity of 98 % based on ClustalW2 analysis, while nucleotides 350-839 had sequence homology with African type I variant (T789C, T795G), with a maximum sequence similarity of 99 % (Table 1) [4–8]. The SBP program revealed evidence of a recombination event at position 406 (codon 101) of the E6 gene according to the AIC criterion (average support for AIC = 92.3 % and cAIC = 90 %) in the multiple sequence alignment between the cloned sequences derived from the cervical sample A15, the reference sequence of the HPV16 genome (HPVR), and representative sequences of HPV16 Asian American, African type I and European variants. In addition, Bootscan analysis supported the clustering of the cervical sample A15 with the Asian American and African type I variants in the E6-E7 genomic region (Fig. 2).

Moreover, sequence analysis in low-grade cervical samples A38 and 1143 between positions 123-143 revealed the nucleotide changes G132T, C143G, which were encountered in the HPV16 African type II variant (Table 1). In contrast, between nucleotides 144-839, only

the nucleotide variation T350G was detected, suggesting sequence homology with the HPV16 European variant (T350G) [4–8] (Table 1). Accordingly, the SBP program showed evidence of a recombination event at position 143 (codon 14) in the E6 gene in the multiple sequence alignment of the A38 and 1143 cloned sequences, the reference sequence of the HPV16 genome, and the representative sequences of HPV16 Asian American, African type I, African type II and European variants (average support for AIC = 100 % and cAIC = 100 %).

Phylogenetic clustering and selective pressure

In the E6 gene, positive selection ($dN/dS > 1$) was indicated at codons 10, 14 and 83 by empirical Bayesian analysis [15, 29–31]. No positively selected sites were found in the E7 gene. Considering the results derived from nucleotide and positive selection analyses, it was assumed that codons 10 and 14 are sufficient for the classification of HPV16 variants. Codon 10 distinguishes African type I and African type II variants, while codon 14 determines only the African type II branch [15]. Considering our results, it was reported that the new sequence variant isolated from cervical samples A38 and 1143 showed sequence homology with the HPV16 European variant (T350G), with amino acid substitutions in codons 10 and 14 that were only identified in the African population.

In order to investigate the classification of novel sequence variants in the distinct phylogenetic branches of the HPV16 genome, maximum-likelihood, maximum-parsimony and Bayesian phylogenetic trees were constructed.

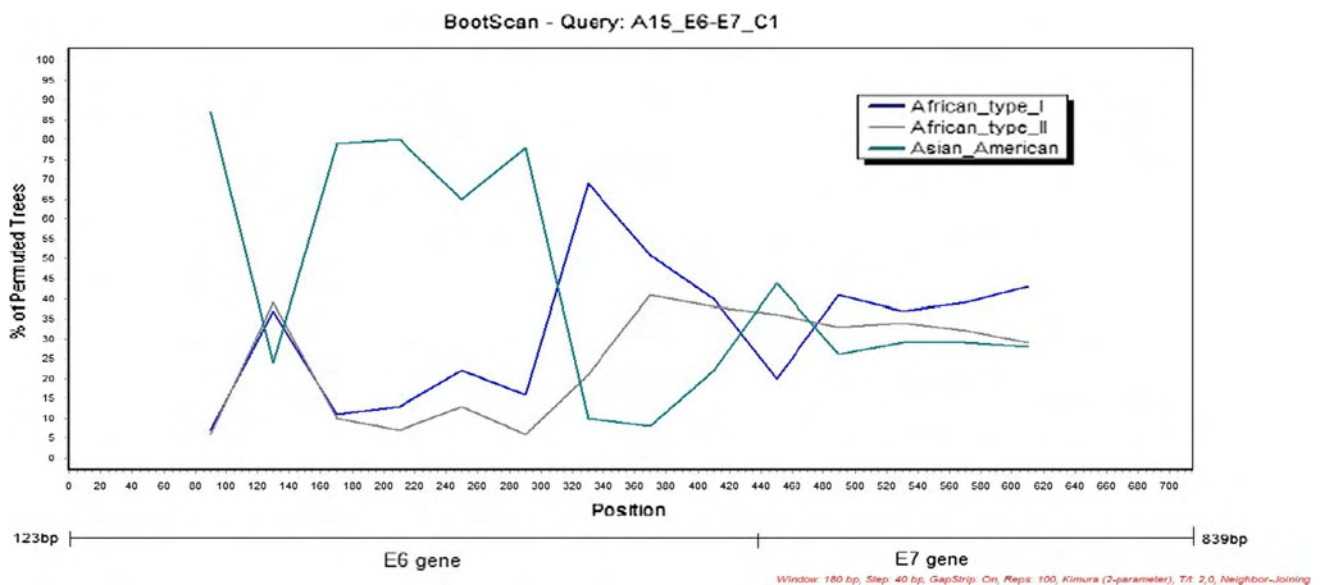


Fig. 2 Bootscan plot of A15 E6-E7 C1 supports the clustering of cloned sequences of cervical sample A15 with the representative HPV16 variants. The y-axis indicates the percentage of bootstrap

values that support the clustering of A15 E6-E7 C1 with different HPV16 variants in different positions of the viral genome between nucleotides 123 and 839

The results of the phylogenetic analysis suggested that the E6-E7 sequences derived from cervical samples A38 and 1143 (10I, 14E, 83 V) are phylogenetically distinct from the established phylogenetic branches (Fig. 1). In addition, the virus in cervical sample A15 was phylogenetically distant from the Asian American branch (Fig. 1).

Discussion

The main objective of this study was to examine the nucleotide polymorphism of HPV16 E6 and E7 oncogenes from positive cervical samples in order to identify HPV16 variants that were circulating in the Greek population [4–8]. According to nucleotide analysis two novel E6-E7 sequence variants were detected in three cervical samples. In order to avoid artificial variations due to PCR errors, different PCR assays were carried out for each individual cervical sample. Amplicons from different PCR assays were subjected to cloning. PCR errors must be excluded, as all of the nucleotide alterations matched those of the African type II and European variants (samples A38, 1443) and Asian American and African type I variant (sample A15) at the corresponding positions and were present in overlapping clones.

Nucleotide sequence analysis of the E6 and E7 genes in the low-grade cervical cases A38 and 1143 revealed sequence homology with the European variant (T350G). However, in the E6 gene, we identified the sequence variations G132T and C143G, which were encountered only in the African type II variant (Table 1). Using the EB method, it was revealed that codons 10 (G132T), 14 (C143G) and 83 (T350G) were characterized as statistically significant positively selected sites [15, 29–31] ($dN/dS > 1$). The combination of positively selected residues 10I, 14E, 83 V that was found in the E6-E7 sequences of A38 and 1143 were not present in the reference sequence of HPV16 variants. Moreover, Carvajal-Rodríguez [15] has proposed that positive selection at codons 10 and 14 plays a significant role in the separation of African type I and African type II variants from the other populations. In the present analysis, codons 10 and 14 were detected as positively selected sites that were encountered in the HPV16 European genome (T350G). In addition, SBP analysis revealed evidence of a recombination event with a potential break point at position 143. Maximum-likelihood, maximum-parsimony and Bayesian phylogenetic clustering revealed that the viruses in cervical samples A38 and 1143 are phylogenetically distant from the other populations (Fig. 1).

Accordingly, in cervical sample A15, nucleotide sequence analysis of the E6-E7 sequence between nucleotides 123 and 350 revealed sequence homology with

Asian American variant, while nucleotides 350 and 839 showed nucleotide homology with the African type I variant (Table 1). SBP analysis revealed evidence of a recombination event in the E6 ORF at position 406 (average support for AIC = 92.3 % and cAIC = 90 %). In addition, Bootscan analysis supported the clustering of cervical sample A15 with the Asian American and African type I variants at different positions in the E6-E7 genomic region. A recent study by Cornet et al. [32] suggested that particular nucleotide polymorphisms in E6 and the LCR region may allow classification of the major HPV16 variant lineages at a deeper level, known as variant sublineages. Cornet et al. [32] demonstrated that the Asian American variant lineage is divided in the Asian American 1, Asian American 2, and North American variant sublineages. The intratypic variation A532G in the E6 gene is absent in the North American variant sublineage and in the Asian American 2 variant sublineage. Taking these data into account, it was concluded that the new E6-E7 sequence variant might have originated due to selective pressure acting on particular residues in the Asian American variant genome and might have led to the emergence of a HPV16 Asian American variant sublineage that was circulating in Greece over evolutionary time. Moreover, in previous studies, we have demonstrated that the Asian American, African type II and African type I variants are circulating in the Greek population, through the study of intratypic nucleotide polymorphisms in the E2 and E4 genes [10, 33]. Taking these data into account, it is possible that the novel sequence variant might have emerged from Asian American or/and African type I variants that circulated for a long time in the Greek population. However, further investigations are required in order to determine the prevalence of this new sequence variant in the Greek population.

Considering that the E6 oncoprotein interacts with distinct p53 variants [30], it can be concluded that new evolutionary E6 intermediates may be formed, sharing characteristics with their relative HPV16 variants. Two possible explanations for the identification of new evolutionary intermediates can be presented. First, as positively selected sites indicate that past evolutionary pressures have led to increased genetic variability [34–36] and the E6 protein interacts with distinct p53 variants, evolutionary pressure might act on the E6 gene, promoting specific nucleotide variations. These sequence changes may lead to a selective advantage of the E6 oncoprotein in effective interaction [15, 29, 30]. Second, intratypic recombination events might be induced by multiple infections with distinct HPV16 variants. Previous studies have suggested that the evolutionary rate of the HPV genome is low, characterized by the absence of recent recombination [13, 37, 38]. However, statistical and phylogenetic evidence suggest that recombination can occur within the HPV16 genome [39].

In particular, Carvajal-Rodríguez [15] identified a potential intratypic recombination breakpoint at nucleotide 171 of the HPV16 E7 gene. In addition, in a previous study by Jiang et al. [17], an intratypic recombination event was reported between the HPV16 European and HPV16 African type II variants, with potential breakpoints between positions 7868 and 31, and between positions 145 and 286.

In summary, in the present analysis, we identified two novel E6-E7 sequence variants that were circulating in the Greek population. The evidence of intratypic recombination events in the HPV16 E6-E7 region that have been described in previous reports imply a rare molecular mechanism that may contribute to the evolution of the HPV16 genome. In addition, the role of positively selected residues in E6 oncoprotein function and phylogenetic clustering of HPV16 allow the E6 gene as a significant evolutionary marker of the viral genome. According to our analyses, it was concluded that new evolutionary HPV16 intermediates might be formed either through positive selective pressure acting in particular on residues in the E6 gene or through recombination events by multiple infections with distinct HPV16 variants. However, further studies are required in order to understand the molecular and evolutionary mechanisms that led to novel E6-E7 sequence variants, their prevalence in distinct populations, and their role in viral pathogenicity.

Acknowledgments This work was supported by research grants from the Postgraduate Programme ‘Applications of Molecular Biology-Genetics. Diagnostic Biomarkers’, code 3817, and “Biotechnology”, code 3439, of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology.

Conflict of interest All authors declare that they have no conflicting or dual interests.

References

- Muñoz N (2000) Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol* 19:1–5
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H (2004) Classification of papillomaviruses. *Virology* 324:17–27
- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79
- Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM (1997) Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* 71:2463–2472
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743–7753
- Wheeler CM, Yamada T, Hildesheim A, Jenison SA (1997) Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* 35:11–19
- Wu Y, Chen Y, Li L, Yu G, He Y, Zhang Y (2006) Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China. *J Gen Virol* 87:1181–1188
- Swan DC, Rajeevan M, Tortolero-Luna G, Follen M, Tucker RA, Unger ER (2005) Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* 96:695–700
- Eriksson A, Herron JR, Yamada T, Wheeler CM (1999) Human papillomavirus type 16 variant lineages characterised by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* 80:595–600
- Tsakogiannis D, Ruether IG, Kyriakopoulou Z, Pliaka V, Theoharopoulou A, Skordas V, Panotopoulou E, Nepka C, Markoulatos P (2012) Sequence variation analysis of the E2 gene of human papillomavirus type 16 in cervical lesions from women in Greece. *Arch Virol* 157:825–832
- Nindl I, Rindfleisch K, Teller K, Schneider A, Dürst M (1999) Cervical cancer, HPV16 E6, variant genotypes, and serology. *Lancet* 353:152
- Sathish N, Abraham P, Peedicayil A, Sridharan G, Chandy G (2005) HPV16 E6 sequence variations in Indian patients with cervical neoplasia. *Cancer Lett* 229:93–99
- Halpern AL (2000) Comparison of papillomavirus and immunodeficiency virus evolutionary patterns in the context of a papillomavirus vaccine. *J Clin Virol* 19:43–56
- DeFilippis VR, Ayala FJ, Villarreal LP (2002) Evidence of diversifying selection in human papillomavirus type 16 E6 but not E7 oncogenes. *J Mol Evol* 55:491–499
- Carvajal-Rodríguez A (2008) Detecting recombination and diversifying selection in human alpha-papillomavirus. *Infect Genet Evol* 8:689–692
- Angulo M, Carvajal-Rodríguez A (2007) Evidence of recombination within human alpha-papillomavirus. *Virol J* 4:33
- Jiang M, Xi LF, Edelstein ZR, Galloway DA, Olsem GJ, Lin WC, Kiviat NB (2009) Identification of recombinant human papillomavirus type 16 variants. *Virology* 394:8–11
- Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25–36
- Puranen M, Saarikoski S, Syrjänen K, Syrjänen S (1996) Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391–395
- Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D (2008) The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355–1361
- Sotlar K, Diemer D, Dethleffs A, Hack Y et al (2004) Detection and typing of human papillomavirus by E6 nested multiplex PCR. *J Clin Microbiol* 42:3176–3184
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533
- Suzuki Y, Gojobori T (1999) A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* 16:1315–1328

26. Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T (2007) Selecton: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Research* 35:W506–W511
27. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD (2006) Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 23:1891–1901
28. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
29. Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD (2005) Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 79:7014–7023
30. DeFilippis VR, Ayala FJ, Villarreal LP (2002) Evidence of diversifying selection in human papillomavirus type 16 E6 but not E7 oncogenes. *J Mol Evol* 55:491–499
31. Lee K, Magalhaes I, Clavel C, Briolat J, Birembaut P, Tommasino M, Zehbe I (2008) Human papillomavirus 16 E6, L1, L2 and E2 gene variants in cervical lesion progression. *Virus Res* 131:106–110
32. Cornet I, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, Tommasino M, Clifford GM, IARC HPV Variant Study Group (2012) Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol* 86:6855–6861
33. Tsakogiannis D, Ruether IG, Kyriakopoulou Z, Pliaka V, Skordas V, Gartzonika C, Levidiotou-Stefanou S, Markoulatos P (2012) Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece. *Arch Virol* 157:1729–1739
34. Bush RM (2001) Predicting adaptive evolution. *Nat Rev Genet* 2:387–392
35. Kosiol C, Bofkin L, Whelan S (2006) Phylogenetics by likelihood: evolutionary modeling as a tool for understanding the genome. *J Biomed Informatics* 39:51–61
36. Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431–449
37. Chan SY, Ho L, Ong CK, Chow V, Drescher B, Dürst M, ter Meulen J, Villa L, Luande J, Mgaya HN et al (1992) Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J Virol* 66:2057–2066
38. Chan SY, Bernard HU, Ratterree M, Birkebak TA, Faras AJ, Ostrow RS (1997) Genomic diversity and evolution of papillomaviruses in rhesus monkeys. *J Virol* 71:4938–4943
39. Varsani A, van der Walt E, Heath L, Rybicki EP, Williamson AL, Martin DP (2006) Evidence of ancient papillomavirus recombination. *J Gen Virol* 87:2527–2531

Molecular and evolutionary analysis of HPV16 E6 and E7 genes in Greek women

D. Tsakogiannis,¹ A. Papadopoulou,¹ G. Kontostathi,¹ I. G. A. Ruether,¹ Z. Kyriakopoulou,¹ T. G. Dimitriou,¹ G. Orfanoudakis² and P. Markoulatos¹

Correspondence

P. Markoulatos

markoulatos@bio.uth.gr

¹University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology, Microbiology-Virology Laboratory, Larissa, Greece

²University of Strasbourg, Ecole Supérieure de Biotechnologie de Strasbourg, CNRS FRE 3211-Oncoprotein Group, 67412 Illkirch, France

Human papillomavirus type 16 (HPV16) non-European variants have been associated with persistent infection and cervical cancer development, while the L83V variant of the E6 gene has been correlated with the progression of cervical malignancy. The present study investigated the presence of the HPV16 L83V variant in Greek women. Molecular evolutionary analysis of the HPV16 E6 and E7 oncogenes was conducted in order to estimate the evolution of the HPV16 genome in the Greek population. The E6 L83V variant was found in 78.2% of high- and 64.28% of low-grade specimens. Moreover, the prototype and E6 L83V variants were both prevalent in high- and low-grade malignancies in Greek women. Selective pressure analysis of the individual amino acid residues of HPV16 sequences from the Greek population indicates that codon 83 of the E6 protein, as well as codon 85 of the E7 protein, are undergoing positive selection. Novel sequence variations were recorded within the E6 and E7 genes in cervical samples, characterized as (T350G) European variants. However, no signal of intratypic recombination event was identified within the E6–E7 region. Molecular and evolutionary analyses of HPV16 genomes from distinct geographical locations might provide valuable information about viral evolution and oncogenicity.

Received 27 November 2012

Accepted 10 August 2013

INTRODUCTION

Human papillomavirus (HPV) is composed of a 7.9 kb double-stranded circular DNA encapsidated in an icosahedral capsid. HPV infects cutaneous and mucosal epithelial cells, leading to malignant transformation (Muñoz, 2000). A total of 40 HPV types have been related to anogenital tract malignancy and have been grouped as high-risk (HR) and low-risk (LR) genotypes (Bernard *et al.*, 2010, 2006; zur Hausen, 1996). Persistent infection with HR HPV is a major cause of cervical cancer worldwide. HPV16 and HPV18 are the most common types identified in invasive cervical cancer cases (Li *et al.*, 2011; de Sanjose *et al.*, 2010).

The HPV genome is divided into three viral domains, consisting of the early region that encodes five early genes (E1, E2, E4, E6, E7), the late region that encodes two late genes (L1, L2) and the long control region (LCR). HPV

Abbreviations: CIN, cervical intraepithelial neoplasia; EB, empirical Bayesian; GTR, general time reversible; HPV, human papillomavirus; HR, high risk; LCR, long control region; LR, low risk; ML, maximum-likelihood; SBP, single recombination breakpoint.

The GenBank/EMBL/DDBJ accession numbers for the E6–E7 region sequences are given in the supplementary data available with the online version of this paper.

types are defined according to the percentage of sequence identity at the L1 gene. Sequences from any two distinct HPV types differ by at least 10% at the nucleotide level. Isolates derived from the same HPV type are referred as variants when less than 10% sequence diversity in L1 gene is encountered (Bernard *et al.*, 2006, 2010). The HPV genome has a low evolutionary rate, and is also characterized by the absence of recent recombination events (Chan *et al.*, 1997; Halpern, 2000). Nevertheless, previous studies that have used alpha HPV sequences from the public database at Los Alamos National Laboratory identified significant recombination signals, located in E6, E7, L1 and L2 genes at different HPV types and namely within HPV16 (Carvajal-Rodriguez, 2008; Angulo *et al.*, 2007).

The HPV16 genome has been extensively investigated for nucleotide polymorphisms (that occur via random mutations) and a number of HPV16 variants have been detected in distinct geographical locations and ethnic groups (Schiffman *et al.*, 2010; Chen *et al.*, 2005, 2009). In particular, HPV16 is divided into four distinct phylogenetic branches, the distribution of which varies geographically. These four distinct groups are designated as (i) European–Asian (including the sublineages European and Asian), (ii) African I, (iii) African II and (iv) North

American/Asian American (Cornet *et al.*, 2012; Yamada *et al.*, 1995, 1997). Several analyses revealed a number of intratypic nucleotide polymorphisms within the E2, E6, E7, L1 and LCR regions (Wu *et al.*, 2006; Swan *et al.*, 2005; Eriksson *et al.*, 1999; Wheeler *et al.*, 1997; Yamada *et al.*, 1995, 1997). Interestingly, intratypic nucleotide polymorphisms within the HPV16 E4 ORF are capable of correctly classifying an HPV16 genome in any of the four phylogenetic branches mentioned above (Tsakogiannis *et al.*, 2012). Different HPV16 variants might present distinct biological functions, thus resulting in different pathogenicity and immunogenicity. Epidemiological studies have suggested that the presence of HPV16 non-European variants constitute HR factors for development of cervical malignancy and invasive cancer (Sanchez *et al.*, 2011; Tornesello *et al.*, 2011; Quint *et al.*, 2010). It has also been proposed that nucleotide polymorphisms within the HPV16 E6 gene (rather than in other HPV16 genes) influence the progression of cervical malignancy, and that additional factors are likely to play a significant role as well (Lee *et al.*, 2008).

This study considers the significant role of the E6 and E7 oncogenes, and focuses on these two regions in order to investigate the nucleotide variability between different HPV16 variants that circulate in the Greek population. The most dominant HPV16 variants were detected in high- and low-grade cervical malignancies in Greek patients. Finally, molecular evolution analyses of the E6 and E7 genes were conducted in order to better understand viral evolution and pathogenicity.

METHODS

Sample collection and DNA isolation. A total of 53 HPV16 positive cervical specimens were derived from ThinPrep (46 specimens) and cervical-stained smears (seven specimens, A3, A4, A6, A13, A31, PNL7 and PNL16). In particular, 23 cervical samples were diagnosed as high-grade cervical intraepithelial neoplasia (CIN II, III), 28 as low-grade cervical intraepithelial neoplasia (CIN I) and two cervical specimens (PNL7, PNL16) were characterized as cervical cancers. The cervical samples were collected from patients attending for the annual Pap test, while smears were collected from patients with a history of abnormal smears. Patient ages ranged from 25 to 60 years.

DNA from ThinPrep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) (Casas *et al.*, 1995). DNA from cervical-stained smears was extracted following the method described by Puranen *et al.* (1996). The quality of extracted DNA was evaluated by PCR amplification of a 498 bp fragment of human β -actin gene (Li *et al.*, 2008). HPV typing was performed using the Nested Multiplex PCR described by Sotlar *et al.* (2004).

E6–E7 amplification by PCR. The cervical samples were subjected to PCR amplification for HPV16 E6–E7 genes using the primer set: HPV-16 41 5'-AGCGACCCAGAAAGTTACCA-3'/HPV-16 757 5'-ATGGGGCACACAATTCCTAG-3'. The primers were designed with the Primer3 program (<http://bioinfo.ut.ee/primer3-0.4.0/primer3/>). The numbering of nucleotide positions was based upon the sequence of the prototype HPV16 strain (GenBank ID: NC001526). The set of primers was designed to amplify both HPV16 E6 and E7 genes. The forward primer hybridized 20 bp downstream from the transcriptional start site of the HPV16 E6

gene, while the reverse primer hybridized 10 bp upstream from the 3' end of the HPV16 E7 gene.

PCR was performed in a final volume of 50 μ l. PCR mixture contained 50 pmol of each primer, 5 \times Green GoTaq Flexi buffer (Promega), 2 mM MgCl₂, 1 mM dNTPs and 1.25 U of thermostable DNA Polymerase (GoTaq, Promega). The cycling conditions were as follows: 40 cycles of 40 s at 95 °C, 50 s at 55 °C and 1 min at 72 °C. The first cycle was preceded by a 2 min denaturation step at 95 °C and the last cycle was followed by a 5 min elongation step at 72 °C.

Cloning and sequencing of E6–E7 genes. The 715 bp amplicons of the HPV16 E6–E7 genes were ligated into pGEM T-easy vector (Promega) and transformed into *Escherichia coli* DH5a competent cells. Recombinant plasmid DNA was purified using the Nucleospin plasmid kit (Macherey-Nagel GmbH) and the plasmids were subjected to sequencing at Macrogen. In order to confirm nucleotide variations, two independent PCR assays were carried out for each individual cervical sample and amplicons from the two distinct PCR assays were subjected to cloning. Three clones from each individual amplicon were isolated. In addition, the two strands of all tested E6–E7 clones were used in sequencing reaction analysis.

In order to identify nucleotide polymorphism and amino acid substitutions into the E6 and E7 ORFs, multiple sequence alignments were performed with the MUSCLE algorithm, in MEGA v.5 software (Edgar, 2004; Tamura *et al.*, 2007). The multiple sequence alignment was conducted between the cloned E6–E7 sequences and the reference sequence of HPV16 genome (HPV16R), available at the HPV16 Sequence Database (Los Alamos National Laboratory).

Phylogenetic analysis. Specific intratypic sequence polymorphisms within the E6 and E7 genes were used in order to classify the E6–E7 cloned sequences to the corresponding HPV16 variants. In order to examine the phylogenetic clustering of HPV16 E6–E7 cloned sequences, a multiple sequence alignment was carried out among the E6–E7 cloned sequences, the reference sequence of HPV16 (HPV16R) available at the HPV16 Sequence Database (Los Alamos National Laboratory), and those of representative sequences of HPV16 variants available in the GenBank sequence database under accession numbers for European variant AF536179 (European-German type), AY686580, AY686581, AF125673 and FJ006723 (North American), for East Asian variant AF534061, for African type I variant AF472508 and AF536180, for African type II variant AF472509 and for the Asian American variant, AF402678. The sequence alignment was performed using the MUSCLE algorithm (Edgar, 2004) of MEGA v.5 (Tamura *et al.*, 2007).

Selection and recombination analysis. Maximum-likelihood (ML) and empirical Bayesian (EB) methods were implemented in order to identify the selective pressure acting upon the E6 and E7 genes. For the ML analysis, the fixed effects likelihood model (FEL) was implemented at the Datamonkey server (Kosakovsky Pond & Frost, 2005; Suzuki & Gojobori, 1999) in order to identify selection pressure. FEL estimates the rate between dN (non-synonymous changes per non-synonymous site) and dS (synonymous changes per synonymous site) at every site. The FEL method assigns a codon as positively selected at a 0.05 *P*-value cut-off (Carvajal-Rodriguez, 2008). Implementation of the EB analysis was through the Selecton 2.4 program (Stern *et al.*, 2007), using the highest precision level. In addition, a Bayesian analysis using Markov chain Monte Carlo methods using the program BEAST v.1.6.2 (Drummond & Rambaut, 2007) was also implemented. Bayesian analysis consisted of four chains with random starting trees. Convergence was assessed with effective sample size values, after a burning of 1×10^6 steps. The median mutation rate in the first, second and third position of the codon in the E6 and E7 reading frames was estimated using Tracer v.1.5 within the BEAST software package. The general time reversible

(GTR) plus Gamma substitution model was selected according to Akaike information criterion by the FindModel online software (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html). In addition, the single recombination breakpoint (SBP) method in Datamonkey server (Kosakovsky Pond *et al.*, 2006) was used in order to identify any evidence of recombination signal.

Phylogenetic tree. An ML phylogenetic tree was computed (1000 bootstrap replicates) from the alignment of the reference HPV16 E6 sequence (HPV16R), the E6 cloned sequences derived from the present study (Tables 1 and 2) and the HPV16 E6 sequences that were presented in the study of Carvajal-Rodriguez (2008) (Fig. 1). Thus, it was possible to evaluate how the sequences in the present study classify with respect to codons 10, 14 and 83. Prior to phylogenetic estimation, the online software FindModel (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html) was used to identify the best evolutionary model as mentioned above. The GTR plus Gamma nucleotide substitution model was used for the E6 phylogenetic tree. The nucleotide sequences used are available in GenBank under the following accession numbers: for European variant AF125673, AY686583, AY686580, AY686584 and AF536179 (European-German type), for East Asian variant AF534061, for African type I variant AF472508 and AF536180, for African type II variant AF472509 and for the Asian American variant AF402678, AY686579 and AY686582.

RESULTS

E6–E7 sequence variations

Sequence analysis of the HPV16 E6 gene shows that this viral region is polymorphic. The most frequently observed nucleotide changes the T350G that induces the L83V amino acid substitution. In high-grade cases, the nucleotide variation T350G was detected in 69.6% (16/23) of cervical specimens, while in low-grade dysplasia, the T350G polymorphism was identified in 64.3% (18/28) of cervical samples (Tables 1 and 2). Nevertheless, the nucleotide variations A184G, G201A, G219A, A280T, T302A, A336G, G514A, C523T and C539T were identified for the first time in the present study (Table 1). The nucleotide variations at positions 184, 201, 219, 302, 336, 514 and 539 induced the amino acid substitutions I27M, C33Y, R39Q, L67I, H78R, M137I and R146C, respectively. However, two silent mutations (A280T, C523T) were identified within the E6 ORF.

The nucleotide variations A184G, G201A, G219A, C523T and C539T were detected in high- and low-grade cervical samples (Table 1). In addition, the silent nucleotide variation A280T was detected in cloned sequences derived from the high-grade cervical samples ATT5 and ATT20 and the low-grade sample PNL2, while the nucleotide changes T302A, A336G and G514A were identified only in low-grade cervical intraepithelia neoplasias (Table 1). The newly identified nucleotide variations were detected in HPV16 strains that harbour the T350G nucleotide variation. Nevertheless, G219A was identified in both European and African type II variants. In particular, G219A was found in low-grade sample 2279 and high-grade sample ATT9 that were characterized as European

variants. In addition, the G219A was identified in two cloned sequences derived from the low-grade cervical specimen A17 (A17 C1, A17 C2), characterized as African type II (Tables 1 and 2).

Sequence analysis of the HPV16 E7 gene revealed that this ORF appears to be more conserved among distinct HPV16 variants. However, seven nucleotide changes were identified for the first time in the present report. The nucleotide variations C565T, T596A, G685A, G813A, G814A, G815A and C818T were detected in cervical samples that harboured the nucleotide variation T350G. These variations induce the amino acid changes H2Y, M12K, A42T, M84I, G85S, G85D and T86I, respectively. The nucleotide variations T596A, G685A, G814A, G815A and C818T were detected in both high- and low-grade cervical intraepithelial neoplasias (CIN) (Table 1). In contrast, the nucleotide changes C565T, G813A were recorded only in low-grade cervical cases (Table 1). In particular, the C565T was detected in the low-grade cervical cases A24 and 1620, while the G813A was identified in two cloned sequences derived from the cervical sample A24 (A24 C2, A24 C3) and in the low-grade cervical sample 858 (Table 1).

HPV16 variant determination

Specific nucleotide changes within the E6 and E7 genes have been associated with the five major variant lineages: prototype strain (no changes), European variant E-G350 (T350G), E-C109G (T109C, T350G), E-G131G (A131G, T350G), East Asian (T178G, A647G), African type I (G132C, C143G, G145T, T286A, A289G, C335T, T789C, T795G), African type II (T109C, G132T, C143G, G145T, T286A, A289G, C335T, G403G, A647G, T789C, T795G) and Asian American (G145T, T286A, A289G, C335T, T350G, A532G, T732C, T789C, T795G) (Wu *et al.*, 2006; Swan *et al.*, 2005; Wheeler *et al.*, 1997; Yamada *et al.*, 1995, 1997).

According to nucleotide analysis of E6 and E7 ORFs, the HPV16 prototype sequence was identified in ten low-grade (2121, 2127, 2148, 2216, 796, ATT8, ATT10, ATT12, ATT18, ATT45) and in six high-grade cervical samples (PNL8, PNL10, PNL11, PNL12, ATT20, ATT24) (Tables 1 and 2). In addition, European sequences (T350G) were recorded in 18 low- and 15 high-grade CIN cases (Tables 1 and 2). Asian American sequences were reported in the high-grade cervical sample A3 and in the cancer case PNL7, while African type I sequence was identified in the high-grade cervical sample A4. The cancer case PNL16 was characterized as (T350G) European variant. In addition, two different HPV16 variants were identified in the low-grade cervical sample A17. In particular, one cloned sequence of the cervical specimen A17 (A17 C2) was characterized as a European variant, while in two other cloned sequences, A17 C1 and A17 C3, E6–E7 African type II sequences were detected (Tables 1 and 2). However, the possibility of an artefactual presence of both African type II

Table 2. Sequence analysis of HPV16 E6 and E7 ORFs in low- and high-grade cervical samples

-, absence of nucleotide variation or amino acid substitution; A, adenine; T, thymine; G, guanine; C, cytosine. The African type II variant was identified in two cloned sequences derived from the low-grade cervical sample A17 (A17 C1, A17 C3) (Table 1). However, the cloned sequence A17 C2 was characterized as European variant. This is the only multiple infection that was described in the present report.

Clones*	E6 region																		E7 region								Variants	Accession number in Genbank											
	131A	132G	143C	145G	145G	178T	184A	201G	219G	280A	286T	289A	302T	335C	336A	350T	403A	514G	523C	532A	539C	565C	596T	647A	685G	732T			789T	795T	813G	814G	815G	818C					
CIN-HG																																							
PNL1 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193802	
PNL9 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422089	
PNL9 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422090	
PNL11 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422094	
PNL11 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422093	
PNL12 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110693	
PNL15 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193804		
PNL15 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193811		
PNL16 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193810		
ATT21 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	KF110687		
ATT24 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110694		
ATT80 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	KF110690		
CIN-LG																																							
PNL3 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193803	
2209 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ796770	
2290 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ796774	
2290 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193812	
2186 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ864252	
2121 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422108		
2121 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422109		
2127 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110692		
2216 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422117		
A17 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422057		
747 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422116		
747 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422107		
793 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422105		
793 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	IQ422106		
796 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422099		
796 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	IQ422100		
ATT6 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	KF110689		
ATT6 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	KF110688		
ATT10 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110686			
ATT10 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110685			
ATT12 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	JX193805			
ATT18 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110695			
ATT22 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193806			
ATT22 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193807			
ATT45 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Pr	KF110684				
ATT51 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193808			
ATT51 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193808			
ATT78 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	JX193809			
European	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	KF110691			
African type I	-	-	C	G	T	-	-	-	-	-	A	G	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
African type II	-	-	T	G	T	-	-	-	-	-	A	G	-	T	-	-	G	-	-	-	-	-	-	G	-	-	-	C	G	-	-	-	-	-	-	-	-		
Asian American	-	-	-	-	-	T	-	-	-	-	A	G	-	T	-	G	-	-	-	-	G	-	-	-	-	-	-	C	C	G	-	-	-	-	-	-			
Asian	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-		
Amino acid	R10G	R10I	Q14E	Q14D	Q14H	D25E	I27M	C33Y	R39Q	P59P	A61A	V62V	L67I	H78Y	H78R	L83V	L100L	M137I	C140C	S143S	R146C	H2Y	M12K	N29S	A42T	F57F	I76I	T78T	M84I	G85S	G85D	T86I							

*The first letter or number corresponds to the code of the cervical sample. The last letter C1, C2, C3 corresponds to the number of the cloned sequence.

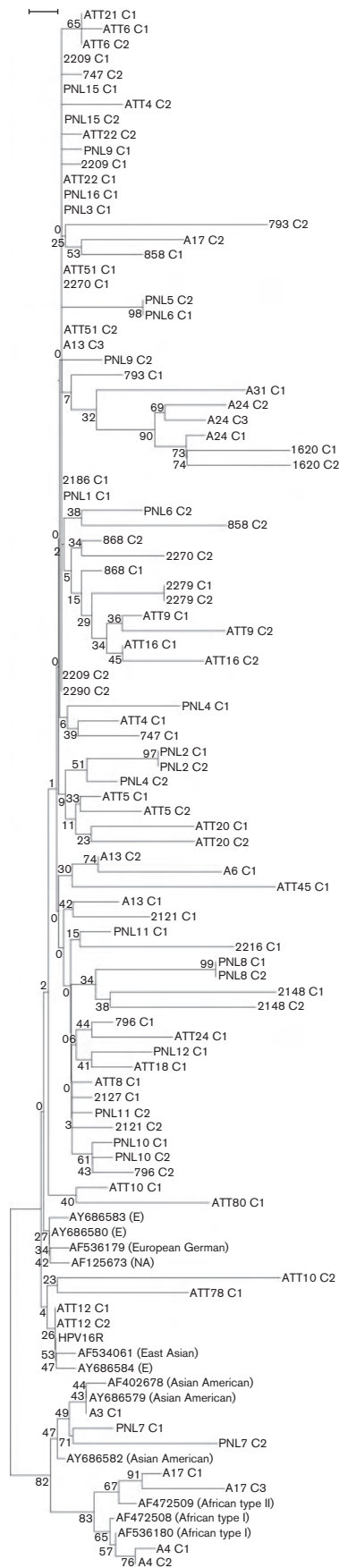


Fig. 1. A maximum-likelihood phylogenetic tree (1000 bootstrap replicates) from the alignment of the reference HPV16 E6 sequence (HPV16R), the E6 cloned sequences derived from the present study and codons 10, 14 and 83 that were presented in the study of Carvajal-Rodriguez (2008). Phylogenetic and positive evolution analysis reported that residue 83 (L83V) has undergone positive selection within the European population. Bar, 0.002.

and European variants in cervical sample A17 cannot be excluded, due to sequencing only two PCR products.

Selection and recombination analysis

Molecular evolution analyses for E6 gene revealed that the relative substitution rates for the three codon position was $CP1=0.826$, $CP2=1.015$, $CP3=1.159$, whereas a corresponding analysis of the E7 ORF estimates $CP1=0.728$, $CP2=1.2$ and $CP3=1.07$. Only under the EB method (Stern *et al.*, 2007) was the codon 83 detected as a positively selected site. Moreover, in the E7 gene, both the ML and EB methods detected codon 85 as a positively selected site, but only under the EB method was the finding statistically significant. The ML phylogenetic tree (Fig. 1) for the sequences in the present study and for codons 10, 14 and 83 (based on previous work by Carvajal-Rodriguez, 2008) revealed a sufficient classification of HPV16 genome in the corresponding phylogenetic branch. However, the SBP program (Kosakovsky Pond *et al.*, 2006) revealed no evidence of recombination in the multiple sequence alignment between the E6–E7 cloned sequences, the reference sequence of HPV16 and those of representative sequences of HPV16 variants.

DISCUSSION

The main objective of this study was to understand the genomic evolution of the HPV16 genome, via the study of genomic diversity and evolution of the E6 and E7 oncogenes. According to sequence analysis, specific nucleotide polymorphisms were detected and used to explore the intratypic heterogeneity of the HPV16 genome in the Greek population. The most frequently identified HPV16 variant was the European variant (T350G), following the HPV16 prototype strain (Tables 1 and 2) (Zuna *et al.*, 2009; Andersson *et al.*, 2000). In particular, the European variant (T350G) was identified in 65.2% of high-grade cervical cases and in 64.2% of low-grade specimens, while the prototype strain (350T) was detected in 26.1% of high-grade and in 35.7% of low-grade CIN samples (Tables 1 and 2). The nucleotide change T350G corresponds to amino acid substitution L83V, which is one of the changes detected as undergoing positive selection (Chen *et al.*, 2005; DeFilippis *et al.*, 2002; Carvajal-Rodríguez, 2008). Previous studies have revealed that HPV16 E6 L83V variant is prevalent in high-grade lesions and is associated with progression of cervical malignancy (Grodzki *et al.*, 2006;

Andersson *et al.*, 2000; Yamada *et al.*, 1997). In the present study, the E6 L83V variant was detected in 69.5 % of high-grade and 64.28 % of low-grade cases. Moreover, the L83V variant was identified in the two cervical cancer cases (PNL7 and PNL16). The difference among high- and low-grade malignancies was not statistically significant. The prototype and the E6 L83V variant were both prevalent in high- and low-grade malignancies in Greek women.

Previous studies have proposed that specific intratypic variants might influence the progression of preinvasive cervical lesions to cancer. In particular, it was proposed that there is a strong relation between Asian American variant and cervical cancer development (Sanchez *et al.*, 2011; Tornesello *et al.*, 2011; Quint *et al.*, 2010; Junes-Gill *et al.*, 2008). According to sequence analyses of the E6–E7 cloned sequences and the representative sequences of HPV16 variants, non-European variants were identified in four cervical samples (PNL7, A3, A4, A17). The HPV16 Asian American variant was identified in cervical sample PNL7, diagnosed as invasive cervical cancer and in the high-grade cervical sample A3 (Table 1). However, two E6–E7 cloned sequences derived from the low-grade cervical sample A17 were characterized as African type II variants, while one E6–E7 cloned sequence derived from the cervical sample A17 was characterized as European variant (Tables 1 and 2). This was the only mixed infection that was recorded in the present study although an artefactual event could not be excluded due to sequencing only two PCR products. In addition, the cervical sample A4 was characterized as an African type I variant. Taking these data into account, we conclude that non-European variants are circulating in Greek population and are detected in high- and low-grade CINs (Tables 1 and 2).

Sequence analysis of the E6 gene revealed significant sequence heterogeneity, clustering the HPV16 variants in distinct phylogenetic branches (Pande *et al.*, 2008; Swan *et al.*, 2005; Wheeler *et al.*, 1997; Yamada *et al.*, 1995, 1997). Nevertheless, nine new nucleotide variations (A184G, G201A, G219A, A280T, T302A, A336G, G514A, C523T and C539T) were detected in the present study (Table 1) in cervical samples that were characterized as (T350G) European variants. In particular, the nucleotide variations A184G, G201A, G219A, T302A and A336G induced the amino acid changes I27M, C33Y, R39Q, L67I and H78R, respectively. These amino acid changes were located at the N-terminal domain of the E6 oncoprotein (Ghittoni *et al.*, 2010; Boulet *et al.*, 2007). Additionally, the nucleotide changes G514A and C539T caused the amino acid substitutions M137I and R146C, respectively. These changes were located at the C-terminal domain of the E6 protein into the Zn²⁺ Finger portion (Ghittoni *et al.*, 2010; Boulet *et al.*, 2007). Moreover, the nucleotide variations A280T and C523T did not induce amino acid substitutions within the E6 protein. The nucleotide variation A280T was detected in both high- and low-grade cervical samples (Table 1).

Previous studies have proposed that the E7 gene is more conserved than E6 (de Boer *et al.*, 2004; Wu *et al.*, 2006;

Pande *et al.*, 2008). Nevertheless, seven new nucleotide variations were identified within the E7 gene in cervical samples that clustered to the same phylogenetic branch with the (T350G) European variant. The nucleotide changes C565T, T596A, G685A, G813A, G814A, G815A and C818T caused the amino acid substitutions H2Y, M12K, A42T, M84I, G85S, G85D and T86I, respectively. In particular, the amino acid substitutions H2Y and M12K were located at the CR1 domain of the E7 oncoprotein (Münger *et al.*, 2004). In addition, the amino acid changes M84I, G85S, G85D and T86I were detected in the C-terminal domain of the E7 protein into the Zn²⁺ Finger portion (Ghittoni *et al.*, 2010) (Table 1). The new nucleotide and amino acid substitutions of E6 and E7 ORFs that were detected in European variants require further investigations, in order to discern their functional implications.

However, the ability of the E6 oncoprotein to interact with distinct p53 variants might induce strong selective pressure upon the E6 gene (DeFilippis *et al.*, 2002). In particular, codons 10, 14 and 83 of the E6 gene are under positive selection and the L83V amino acid substitution is associated with cervical cancer development (Chen *et al.*, 2005; DeFilippis *et al.*, 2002; Carvajal-Rodríguez, 2008; Lee *et al.*, 2008). In the present study, positive selection was found in codon 83 of the E6 gene. However, no signal of positive selection was found in codons 10 and 14, by the ML and EB methods. Considering the positive selection analysis and the ML phylogenetic tree (Fig. 1), we assume that codon 83, which is implicated in cervical cancer progression, has undergone positive selection within the European population. However, nucleotide and phylogenetic analyses revealed that codon 10 distinguishes African I and II, while codon 14 separates African type II from the non-African variants (Carvajal-Rodríguez, 2008). Absence of positive selection in codons 10 and 14 might be caused by dominance of HPV16 European variant in the Greek population (Fig. 1).

It was proposed that positive selection is acting upon codon 29 of the E7 protein (Sun *et al.*, 2012). In the present analysis, only codon 85 was found to be under positive selection by EB analysis. Codon 85 is located within the C-terminal domain of E7 protein. This protein domain is implicated in protein interactions with the pRb and other cellular proteins, and may act as a dimerization domain (Münger *et al.*, 2004). Amino acid substitution in codon 85 was detected only in cervical samples that were characterized as European variants. The role of positive selection in codon 85 in the Greek population should be further investigated in order to discern the interaction between the E7 oncoprotein and host cellular targets.

In conclusion, the HPV16 E6 and E7 genes are polymorphic regions of the viral genome and provide significant information about HPV16 intratypic heterogeneity and evolution. To the best of our knowledge, this is the first study describing nucleotide variations and

evolutionary pressure acting on E6 and E7 regions of HPV16 genomes from the Greek population. Molecular and evolutionary analyses of the HPV16 genome in distinct geographical locations might provide valuable information about the viral pathogenicity and genome evolution.

ACKNOWLEDGEMENTS

We are grateful to E. Panotopoulou, Papanicolaou Research Centre of Oncology and Experimental Surgery, Anticancer Oncology Hospital of Athens 'St Savvas', Athens, Greece, for providing the clinical samples and for valuable discussions. The work was supported by research grants of the Postgraduate Programme 'Applications of Molecular Biology-Genetics. Diagnostic Biomarkers' (3817) and 'Biotechnology' (3439), of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology. All authors declare that they have no conflicting or dual interests.

REFERENCES

- Andersson, S., Alemi, M., Rylander, E., Strand, A., Larsson, B., Sällström, J. & Wilander, E. (2000). Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. *Br J Cancer* **83**, 307–310.
- Angulo, M. & Carvajal-Rodríguez, A. (2007). Evidence of recombination within human alpha-papillomavirus. *Virology* **4**, 33.
- Bernard, H. U., Calleja-Macias, I. E. & Dunn, S. T. (2006). Genome variation of human papillomavirus types: phylogenetic and medical implications. *Int J Cancer* **118**, 1071–1076.
- Bernard, H. U., Burk, R. D., Chen, Z., van Doorslaer, K., zur Hausen, H. & de Villiers, E. M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* **401**, 70–79.
- Boulet, G., Horvath, C., Vanden Broeck, D., Sahebali, S. & Bogers, J. (2007). Human papillomavirus: E6 and E7 oncogenes. *Int J Biochem Cell Biol* **39**, 2006–2011.
- Carvajal-Rodríguez, A. (2008). Detecting recombination and diversifying selection in human alpha-papillomavirus. *Infect Genet Evol* **8**, 689–692.
- Casas, I., Powell, L., Klapper, P. E. & Cleator, G. M. (1995). New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* **53**, 25–36.
- Chan, S. Y., Bernard, H. U., Ratterree, M., Birkebak, T. A., Faras, A. J. & Ostrow, R. S. (1997). Genomic diversity and evolution of papillomaviruses in rhesus monkeys. *J Virol* **71**, 4938–4943.
- Chen, Z., Terai, M., Fu, L., Herrero, R., DeSalle, R. & Burk, R. D. (2005). Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* **79**, 7014–7023.
- Chen, Z., DeSalle, R., Schiffman, M., Herrero, R. & Burk, R. D. (2009). Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *J Virol* **83**, 1443–1455.
- Cornet, I., Gheit, T., Franceschi, S., Vignat, J., Burk, R. D., Sylla, B. S., Tommasino, M., Clifford, G. M. & IARC HPV Variant Study Group (2012). Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol* **86**, 6855–6861.
- de Boer, M. A., Peters, L. A., Aziz, M. F., Siregar, B., Cornain, S., Vrede, M. A., Jordanova, E. S., Kolkman-Uljee, S. & Fleuren, G. J. (2004). Human papillomavirus type 16 E6, E7, and L1 variants in cervical cancer in Indonesia, Suriname, and The Netherlands. *Gynecol Oncol* **94**, 488–494.
- de Sanjose, S., Quint, W. G., Alemany, L., Geraets, D. T., Klaustermeier, J. E., Lloveras, B., Tous, S., Felix, A., Bravo, L. E. & other authors (2010). Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* **11**, 1048–1056.
- DeFilippis, V. R., Ayala, F. J. & Villarreal, L. P. (2002). Evidence of diversifying selection in human papillomavirus type 16 E6 but not E7 oncogenes. *J Mol Evol* **55**, 491–499.
- Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.
- Eriksson, A., Herron, J. R., Yamada, T. & Wheeler, C. M. (1999). Human papillomavirus type 16 variant lineages characterized by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J Gen Virol* **80**, 595–600.
- Ghittoni, R., Accardi, R., Hasan, U., Gheit, T., Sylla, B. & Tommasino, M. (2010). The biological properties of E6 and E7 oncoproteins from human papillomaviruses. *Virus Genes* **40**, 1–13.
- Grodzki, M., Besson, G., Clavel, C., Arslan, A., Franceschi, S., Birembaut, P., Tommasino, M. & Zehbe, I. (2006). Increased risk for cervical disease progression of French women infected with the human papillomavirus type 16 E6-350G variant. *Cancer Epidemiol Biomarkers Prev* **15**, 820–822.
- Halpern, A. L. (2000). Comparison of papillomavirus and immunodeficiency virus evolutionary patterns in the context of a papillomavirus vaccine. *J Clin Virol* **19**, 43–56.
- Junes-Gill, K., Sichero, L., Maciag, P. C., Mello, W., Noronha, V. & Villa, L. L. (2008). Human papillomavirus type 16 variants in cervical cancer from an admixed population in Brazil. *J Med Virol* **80**, 1639–1645.
- Kosakovsky Pond, S. L. & Frost, S. D. (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531–2533.
- Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* **23**, 1891–1901.
- Lee, K., Magalhaes, I., Clavel, C., Briolat, J., Birembaut, P., Tommasino, M. & Zehbe, I. (2008). Human papillomavirus 16 E6, L1, L2 and E2 gene variants in cervical lesion progression. *Virus Res* **131**, 106–110.
- Li, W., Wang, W., Si, M., Han, L., Gao, Q., Luo, A., Li, Y., Lu, Y., Wang, S. & Ma, D. (2008). The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* **134**, 1355–1361.
- Li, N., Franceschi, S., Howell-Jones, R., Snijders, P. J. F. & Clifford, G. M. (2011). Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: variation by geographical region, histological type and year of publication. *Int J Cancer* **128**, 927–935.
- Münger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., Grace, M. & Huh, K. (2004). Mechanisms of human papillomavirus-induced oncogenesis. *J Virol* **78**, 11451–11460.
- Muñoz, N. (2000). Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol* **19**, 1–5.
- Pande, S., Jain, N., Prusty, B. K., Bhamhani, S., Gupta, S., Sharma, R., Batra, S. & Das, B. C. (2008). Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J Clin Microbiol* **46**, 1060–1066.
- Puranen, M., Saarikoski, S., Syrjänen, K. & Syrjänen, S. (1996). Polymerase chain reaction amplification of human papillomavirus

DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* **40**, 391–395.

Quint, K. D., de Koning, M. N., van Doorn, L. J., Quint, W. G. & Pirog, E. C. (2010). HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix. *Gynecol Oncol* **117**, 297–301.

Sanchez, G. I., Kleter, B., Gheit, T., van Doorn, L. J., de Koning, M. N., de Sanjosé, S., Alemany, L., Bosch, X. F., Tommasino, M. & other authors (2011). Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *J Clin Virol* **51**, 165–169.

Schiffman, M., Rodriguez, A. C., Chen, Z., Wacholder, S., Herrero, R., Hildesheim, A., Desalle, R., Befano, B., Yu, K. & other authors (2010). A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. *Cancer Res* **70**, 3159–3169.

Sotlar, K., Diemer, D., Dethleffs, A., Hack, Y., Stubner, A., Vollmer, N., Menton, S., Menton, M., Dietz, K. & other authors (2004). Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol* **42**, 3176–3184.

Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E. & Pupko, T. (2007). Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res* **35** (Web Server issue), W506–W511.

Sun, M., Gao, L., Liu, Y., Zhao, Y., Wang, X., Pan, Y., Ning, T., Cai, H., Yang, H. & other authors (2012). Whole genome sequencing and evolutionary analysis of human papillomavirus type 16 in central China. *PLoS ONE* **7**, e36577.

Suzuki, Y. & Gojobori, T. (1999). A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* **16**, 1315–1328.

Swan, D. C., Rajeevan, M., Tortolero-Luna, G., Follen, M., Tucker, R. A. & Unger, E. R. (2005). Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* **96**, 695–700.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.

Tornesello, M. L., Losito, S., Benincasa, G., Fulciniti, F., Botti, G., Greggi, S., Buonaguro, L. & Buonaguro, F. M. (2011). Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* **121**, 32–42.

Tsakogiannis, D., Ruether, I. G., Kyriakopoulou, Z., Pliaka, V., Skordas, V., Gartzonika, C., Levidiotou-Stefanou, S. & Markoulatos, P. (2012). Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece. *Arch Virol* **157**, 1729–1739.

Wheeler, C. M., Yamada, T., Hildesheim, A. & Jenison, S. A. (1997). Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. *J Clin Microbiol* **35**, 11–19.

Wu, Y., Chen, Y., Li, L., Yu, G., He, Y. & Zhang, Y. (2006). Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China. *J Gen Virol* **87**, 1181–1188.

Yamada, T., Wheeler, C. M., Halpern, A. L., Stewart, A. C., Hildesheim, A. & Jenison, S. A. (1995). Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* **69**, 7743–7753.

Yamada, T., Manos, M. M., Peto, J., Greer, C. E., Munoz, N., Bosch, F. X. & Wheeler, C. M. (1997). Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* **71**, 2463–2472.

Zuna, R. E., Moore, W. E., Shanesmith, R. P., Dunn, S. T., Wang, S. S., Schiffman, M., Blakey, G. L. & Teel, T. (2009). Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. *Int J Cancer* **125**, 2609–2613.

zur Hausen, H. (1996). Papillomavirus infections—a major cause of human cancers. *Biochim Biophys Acta* **1288**, F55–F78.

Nucleotide polymorphisms of the human papillomavirus 16 E1 gene

D. Tsakogiannis · F. Darmis · P. Gortsilas ·
I. G. A. Ruether · Z. Kyriakopoulou ·
T. G. Dimitriou · G. Amoutzias · P. Markoulatos

Received: 7 February 2013 / Accepted: 7 June 2013
© Springer-Verlag Wien 2013

Abstract The E1 ORF is one of the most conserved regions in the human papillomavirus (HPV) genome. The complete E1 gene of the HPV16 genome was amplified with four overlapping primer sets in 16 high-grade (CIN II, III) and 13 low-grade cervical (CIN I) intraepithelial neoplasias as well as in one cervical cancer case. Sequence analysis of the E6 and E7 genes was also carried out in the same cervical samples in order to confirm the association between nucleotide sequence variations in the HPV16 E1 ORF and HPV16 variant lineages. Analysis of the E1 ORF revealed 27 nucleotide changes, and these changes were correlated with those found in HPV16 Asian American and African type II variants. Of these nucleotide variations, A1668G, G2073A, T2169C, T2189C, A2453T, C2454T, A2587T and G2650A were identified only in high-grade dysplasia cases. A phylogenetic tree of the E1 ORF and nucleotide sequence analysis of the E1, E6 and E7 genes revealed that intratypic nucleotide sequence polymorphisms located in the E1 ORF can be used to identify the major phylogenetic branch to which a HPV16 genome belongs. Moreover, amplification of the E1 ORF revealed a disruption between nucleotides 878 and 1523 in five high- and two low-grade cervical cases, indicating that integration of HPV DNA occurs at an early stage of viral infection.

Introduction

Human papillomaviruses (HPVs) are a divergent group of small, non-enveloped DNA viruses that replicate in the nucleus of squamous epithelial cells. Over 150 different HPV types have been characterized, and approximately 60 of them are detected in the cervical epithelia and are classified as members of the genus *Alphapapillomavirus* [1, 2]. Epidemiological studies have revealed that high-risk human papillomaviruses are associated with cervical cancer development, with HPV16 (alpha-9) and HPV18 (alpha-7) types being the most frequently identified HPV types in invasive cervical cancer [3–5]. The HPV16 genome is divided into the early region, comprising the early genes E1, E2, E4, E5, E6, and E7, the late region, comprising two late genes (L1, L2), the long control region (LCR), and a small non-coding region (NCR), which is located between the E5 and L2 genes.

Evolutionary analysis has provided evidence that papillomaviruses (PVs) coevolved with their hosts over millions of years [6]. However, Gottschling et al. [7] established that alternative events may contribute considerably to the evolution of PVs. Moreover, the viral genome is further classified into variant lineages within each type [1]. In particular, HPV16 variant lineages have been associated with the geographic origin of the populations in which they are prevalent [8, 9]. The HPV16 genome is classified into five major intratypic variant lineages known as HPV16 European, European Asian, African type II, African type I, and Asian American [10–12]. Nucleotide sequence analysis of the HPV16 variants in the E6, L1, L2, and LCR genomic regions has expanded this phylogenetic classification [9, 13]. In a previous study by Carvajal-Rodriguez et al. [14], it was proposed that classification of a given HPV16 genome in any of the major variant lineages

D. Tsakogiannis · F. Darmis · P. Gortsilas ·
I. G. A. Ruether · Z. Kyriakopoulou · T. G. Dimitriou ·
G. Amoutzias · P. Markoulatos (✉)
Department of Biochemistry and Biotechnology, School of
Health Sciences, University of Thessaly, Ploutonos 26 & Aiolou,
41221 Larissa, Greece
e-mail: markoulatos@bio.uth.gr

can be achieved according to E6 polymorphisms. Moreover, nucleotide polymorphisms in the E4 region can be used to classify the HPV16 variants into distinct phylogenetic branches [15]. However, recent studies have classified the common HPV16 variant lineages at a deeper level, referred to as “sublineages”. Variant sublineages constitute groups of sequences with 0.5 %-1.0 % divergence at the genome level [16–19].

In the papillomavirus genome, the E1 ORF is one of the most conserved regions [20]. The papillomavirus (PV) E1 protein is expressed during the early stage of the virus life cycle in order to maintain the viral DNA in an episomal form. The E1 protein displays both ATPase and 3'-5' helicase activities. Previous studies have suggested that the E1 protein of HPV11 is composed of three functional domains: an N-terminal regulatory region, a central origin-DNA-binding domain (OBD), and a C-terminal ATPase/helicase domain [21–24]. The PV E1 protein recognizes and binds to an AT-rich sequence within the PV DNA replication origin in cooperation with the PV E2 protein. The complex of PV E1 and E2 proteins binds to the viral replication origin with high affinity and initiates DNA replication via the E1 ATP-dependent helicase domain [20, 25]. In addition, the PV E1 helicase domain interacts with DNA polymerase α -primase, topoisomerase I and the single-stranded DNA binding protein RPA in order to recruit the cellular replication factors to the viral origin of replication, forming an active replication complex [26–30].

Previous studies have suggested that the integration of HPV16 DNA into the host genome usually disrupts the viral E1 and/or E2 ORFs [31, 32]. This disruption (of E1 and/or E2 ORFs) results in deregulation of viral gene expression and DNA replication control. This deregulation is associated with the progression of cervical intraepithelial neoplasia to invasive cancer [31–34].

Considering the essential role of the E1 protein in viral replication, the present study focused on the analysis of HPV16 E1 ORF. Thus, the complete HPV16 E1ORF was amplified from 30 cervical samples in order to identify nucleotide sequence variations and amino acid substitutions. Furthermore, sequence analysis of the E6 and E7 genes was also carried out from the same cervical samples in order to investigate the association between HPV16 variants and nucleotide sequence variations in the E1 ORF. In addition, sites of disruptions in the E1 region are reported.

Materials and methods

Cervical specimens

Thirty HPV16-positive specimens from Greek patients (5 paraffin-embedded cervical biopsies and 25 Thin Prep

samples) were diagnosed as cervical intraepithelial neoplasia stage I, II and III. Thin Prep samples were collected from patients who were having their annual Pap test, while the paraffin-embedded tissues were collected during biopsy from patients with a history of abnormal smears. The patients' ages ranged from 25 to 60 years old. Of the 30 samples, 16 were diagnosed as high-grade cervical intraepithelial neoplasias (CIN II, III), and thirteen were diagnosed as low-grade cervical intraepithelial neoplasias (CIN I). In addition, one cervical sample was diagnosed as cervical cancer.

DNA extraction

DNA from Thin Prep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) [35]. In addition, DNA from formalin-fixed, paraffin-embedded cervical tissues was extracted using the established proteinase K method [36], while DNA from stained cervical smears was extracted following the method described by Puramen et al. [37]. A 498-bp beta-actin fragment was amplified by PCR assay as a quality control in order to test the integrity of the extracted DNA and to test for the presence of PCR inhibitors [38]. Identification of HPV16 genomes was carried out using nested multiplex PCR [39].

Amplification of E1 ORF

The HPV16 E1 gene was amplified using the overlapping primer sets E1F1-E1R3, E1F4-E1R5, E1F6-E1R7, and E1F8-E1R9 in separate reactions [32]. The primers were selected to amplify subsequent fragments of the HPV16 E1 gene in order to obtain the complete sequence of E1 ORF and to locate sites of disruption within this ORF. In the present study, gene disruption events were considered only for those samples that were positive for the β -actin gene but gave no signal amplification with one of the primer pairs of HPV16 E1. The PCR reactions were performed in a final volume of 50 μ l. Each PCR mixture contained 50 pmol of each primer set (E1F1-E1R3, E1F4-E1R5, E1F6-E1R7, E1F8-E1R9), 5X Green GoTaq Flexi buffer (Promega), 2 mM MgCl₂, 1 mM dNTPs and 1.25 U of thermostable DNA polymerase (GoTaq® DNA Polymerase, Promega). PCR cycling conditions were as follows: an initial denaturation step at 95 °C for 2 min, then 50 cycles of 95 °C for 40 s, 50 °C for 30 s and 72 °C for 30 s. PCR reaction was ended with a 5 min incubation step at 72 °C.

In addition, when a site of disruption was identified within the E1 ORF, an auto-nested PCR was carried out in that particular viral region in order to confirm the result. Auto-nested PCR amplification was performed using as template 2 μ l of the first-round PCR and the corresponding primer set (E1F1-E1R3, E1F4-E1R5, E1F6-E1R7, E1F8-

E1R9). Auto-nested PCR amplification was conducted for 25 cycles under the conditions described above.

Amplification of the E6 and E7 ORFs

Sequence analysis of the HPV16 E6 and E7 genes was performed in order to correlate particular intratypic variations within those regions of E6 with nucleotide variations within the E1 gene. The E6-E7 region was amplified using the primer set HPV16 41/HPV16 757 as described previously [15].

Cloning and sequence analysis of HPV16 E1, E6 and E7 ORFs

Four distinct overlapping amplicons from E1 (645 bp, 488 bp, 533 bp and 405 bp length) and the E6, E7 genomic regions, derived from each individual cervical sample were subjected to cloning using the pGEM-T Easy Vector system (Promega, Madison, USA). Recombinant plasmid DNA was purified using a Nucleospin Plasmid Kit (Macherey-Nagel GmbH, Duren, Germany), and the plasmids were sequenced at Macrogen Inc., Korea. In order to confirm sequence variations that were detected within the E1, E6 and E7 ORFs, two independent PCR assays were performed for each viral genomic region, and amplicons from distinct PCR assays were cloned. For E1 overlapping PCR fragments, three clones from each individual amplicon and two clones from the E6-E7 genomic region were selected for sequence analysis. Sequencing reactions were performed in both DNA strands and were sequenced three times to insure the reliability of the results. However, only one cloned sequence was used for nucleotide sequence analysis of the E1 gene in cervical samples when the cloned sequences showed 100 % nucleotide identity. Multiple sequence alignment was performed using the Muscle algorithm in the MEGA v.5 software [40, 41], using the E1 and E6-E7 region sequences from the present study, the reference sequence of the HPV16 genome (HPV16R) available in the HPV16 Sequence Database (Los Alamos National Laboratory), and sequences from representative HPV16 lineages and sublineages available in GenBank. The accession numbers were as follows: European prototype 1 variant, HQ644236; European prototype 2 variant, AF536179; European-Asian variant, AF534061; African type I variant, AF472508; African type II variant AF472509; Asian American 1, HQ644247; Asian American 2, HQ644254; Asian American/North American, HQ644257.

Sequences between nucleotides 878 and 1523 of the E1 gene are available in GenBank under accession numbers JQ791065-JQ791098. The accession numbers are JQ768462-JQ768505 for nucleotides 1503 to 1991,

JQ791017-JQ791064 for nucleotides 1952 to 2485, and JQ779915 - JQ779969 for nucleotides 2443 to 2848.

Phylogenetic tree construction

A maximum-likelihood phylogenetic tree was constructed using E1 ORFs sequenced in the present study, the reference sequence of HPV16R, and those of representative sequences of HPV16 variant lineages and sublineages in order to investigate how the E1 gene performs as a marker for assigning a given HPV16 genome to any of the major HPV16 phylogenetic branches (Fig. 1). Cervical samples that showed a disruption within E1 ORF were not used in this phylogenetic tree. The phylogenetic tree (1000 bootstrap replicates) was constructed using the GTR + G substitution model according to the FindModel online software (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>).

Results

E1 gene disruption

Amplification of the HPV16 E1 region revealed an intact E1 gene in 11 out of 16 high-grade cases (68.95 %). The viral domain between nucleotides 878 and 1523 was not amplified in the high-grade cervical samples ATT9, D2, D13, D14, and D15 (31.05 %). In addition, 11 out of 13 low-grade cervical samples, (84.61 %) harbored the complete E1 ORF, while two of them (15.39 %) (samples 747 and D7) showed E1 gene disruption between nucleotides 878 and 1523. However, no evidence of gene disruption was found in the cervical cancer case PNL7. The above results show that the integration of the HPV16 genome through disruption of E1 gene might occur in both high- and low-grade cervical intraepithelial neoplasias. Furthermore, our results were also confirmed through real-time PCR by assaying the E1/E6 ratio. An experimental cutoff value was established in order to determine the physical state of HPV16 DNA as episomal, integrated or mixed (integrated and episomal forms). According to our results, it was concluded that integration occurs when the E1/E6 ratio is less than 0.003. Ratios less than 0.98, ranging from 0.003 to 0.98, indicated the presence of mixed forms of viral DNA, while ratios greater or equal to 0.98 indicated the presence of the episomal form only, (data not shown). The ratios for the high-grade cervical samples ATT9 (E1/E6=0.002), D2 (E1/E6=0), D13 (E1/E6=0), D14 0.001), and D15 (E1/E6=0) and for the low-grade cervical samples 747 (E1/E6=0.002) and D7 (E1/E6=0) were close to zero, indicating the presence of the integrated HPV16 DNA form (Table 1).

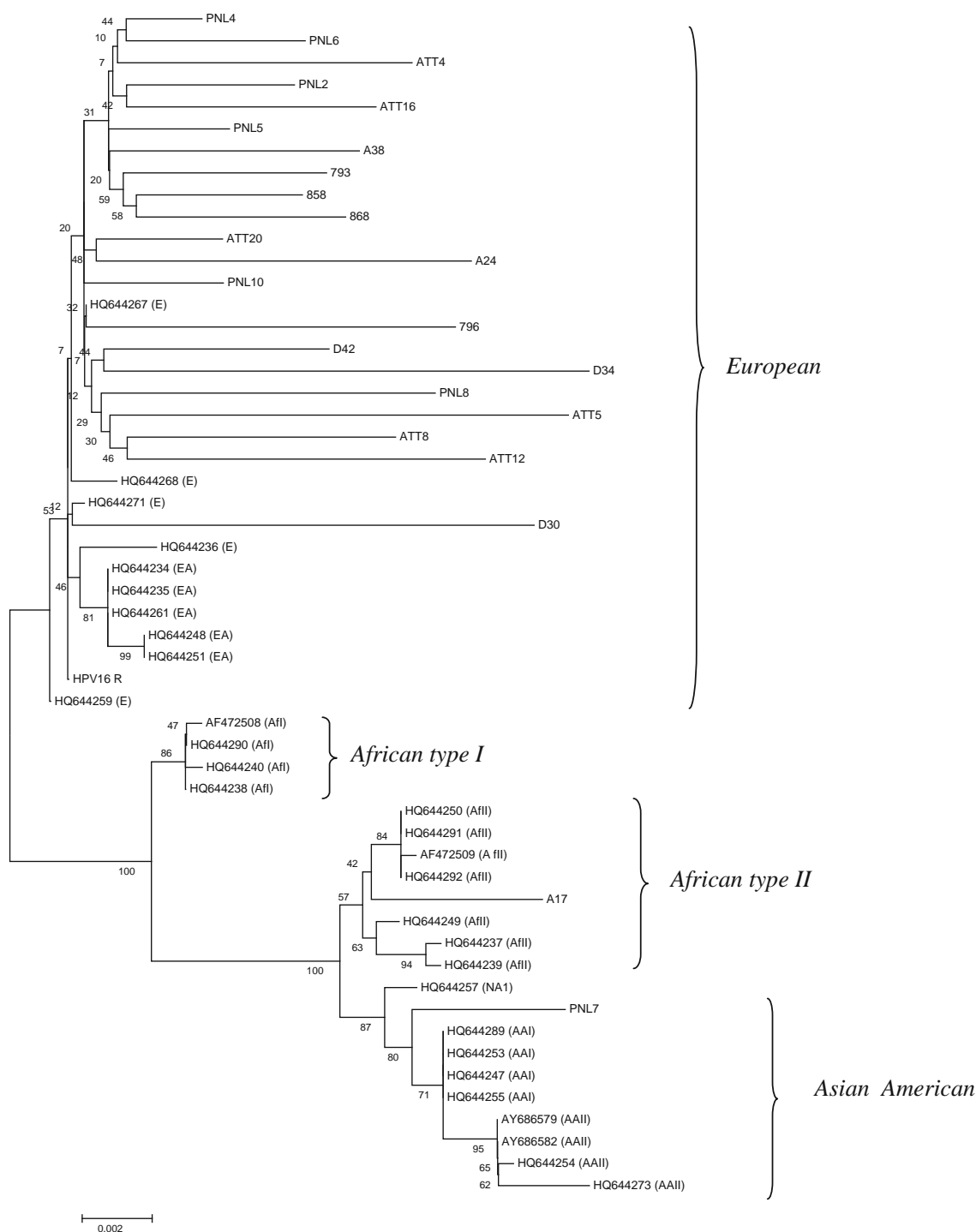


Fig. 1 Maximum-likelihood phylogenetic tree based on the complete E1 gene, derived from the four overlapping fragments of E1 gene of the investigated samples, the reference sequence of HPV16 genome, and the reference sequences of HPV16 variants retrieved from the GenBank sequence database (European prototype 1 variant, HQ644236, HQ644259, HQ644267, HQ644268, HQ644271; European-Asian variant, HQ644234, HQ644235, HQ644248, HQ644251, HQ644261; African type I variant, HQ644238, HQ644240, AF472508, HQ644290; African type II variant, HQ644237, HQ644239, HQ644249, HQ644250,

AF472509, HQ644291, HQ644292; Asian American 1, HQ644289, HQ644247, HQ644253, HQ644255; Asian American 2, HQ644254, AY686579, AY686582, HQ644273; Asian American/North American, HQ644257). The phylogenetic tree (1000 bootstrap replicates) was constructed using the GTR + G substitution model, calculated using the online software FindModel (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html). Using the E1 ORF sequence, the HPV16 variant lineages were divided into distinct phylogenetic branches with high bootstrap values (>80%)

Table 1 Four overlapping primer sets used to amplify the complete E1 ORF

Primer	Sequence (5'-3')	location	PCR product length
E1F1	CAGGTACCAATGGGAAGA	878–896	645 bp
E1R3	GAAAACTCACCCGTATAAC	1503–1523	
E1F4	GTTATACGGGTGAGTTTTTC	1503–1523	488 bp
E1R5	GCAATTCCTACTATCGTCTACTATG	1968–1991	
E1F6	GGGCCTACGATAATGACATAGTA	1952–1974	533 bp
E1R7	CCAATGCATTCTTAAATTGTC	2464–2485	
E1F8	CCCTGTTGGAACCTACATAGATGAC	2443–2466	405 bp
E1R9	GTCTATATGGTCACGTAGGTCTGTA	2824–2848	

Variant determination

The association between nucleotide changes within the E1 ORF and HPV16 variants was conducted through nucleotide sequence analysis of E6 and E7 ORFs on the same cervical samples. Intratypic polymorphisms in the E6 and E7 genes were used in order to classify the HPV16 genome into any of the major variant lineages [9, 13, 16]. The cloned sequences of the E6 and E7 genes were subjected to multiple sequence alignment with the reference sequences of HPV16 variant lineages and sublineages as described previously [16]. Nucleotide sequence analysis of the E6 and E7 genes revealed that 8 low-grade and 11 high-grade cervical samples were related to the HPV16 European variant and the reference sequence of HPV16 genome. However, the low-grade cervical sample A17 showed sequence similarity with the African type II variant, while the cancer case PNL7 showed higher sequence identity with the Asian American variant (Table 2). In addition, nine cervical samples (D2, D7, D13, D14, D15, D16, D30, D34, D42) were associated with the HPV16 European variant by nucleotide sequencing and phylogenetic analysis of the E2 gene, as has been reported previously [36] (Table 2). The phylogenetic tree that was constructed using the complete E1 gene further confirmed our results. According to sequence and phylogenetic tree analysis, it was concluded that the E1 ORF alone can be used to correctly classify the HPV16 genome in any of the major phylogenetic branches (Fig. 1).

E1 nucleotide sequence analysis

Multiple sequence alignment between the E1 amplicons, the genome of the HPV16 reference strain, and the sequences of HPV16 variants was conducted for the four consecutive fragments of the HPV16 E1 gene. Sequence analysis of the E1 region from nucleotides 858 to 1523 revealed the presence of nucleotide variations T921C, C1096G, G1163A, T1200C, T1366A, C1377T, C1426G,

T1486C in the low-grade cervical sample A17, the cancer sample PNL7, and the reference sequences of the HPV16 African type II and Asian American variants (Table 3). The nucleotide changes T921C, T1200C, C1377T and T1486C do not cause amino acid substitutions in the E1 protein, while the nucleotide changes C1096G, G1163A, T1366A and C1426G result in the amino acid changes Q78E, G100E, C168S and Q188E, respectively. However, the silent nucleotide variation A1041G was only identified in the cervical sample PNL7 and in the reference sequence of the Asian American variant, while the silent nucleotide change C1416T was only found in the cervical specimen A17 and in the reference sequence of the African type II variant (Table 3).

Sequence analysis of the E1 region from nucleotides 1503 to 1991 revealed the nucleotide variations C1624T, C1744A, A1842G, which were detected in high-grade cervical sample PNL7, in low-grade sample A17 and in the reference sequences of the African type II and Asian American variants (Table 4). The nucleotide change C1624T was a silent nucleotide variation, while C1744A and A1842G induced the amino acid substitutions L294M and I326M, respectively. In addition, the silent nucleotide variation, A1668G, was detected in the cervical sample PNL7, and in the reference sequence of the Asian American variant (Table 4). Moreover, between nucleotides 1952 and 2485, the nucleotide mutations G2073A, T2169C, T2189C, A2453T and C2454T were detected in high-grade malignancies (samples D2, D13, D15) (Table 5). In addition sequence analysis of the E1 region revealed the nucleotide variations C2041T, G2220C, C2237G, G2249A, C2262T, C2287T and C2344T, from which the nucleotide variations C2041T, G2220C, C2237G, and G2249A induced the amino acid changes L393I, E425D, T458S, and R462K, while the C2262T, C2287T, and C2344T changes were silent nucleotide variations. These nucleotide changes were detected in the cancer sample PNL7, in the low-grade sample A17, and in the reference sequences of the African type II and Asian American variants. Moreover, the silent

Table 2 Sequence analysis of the E6 and E7 genes of cervical samples in which the E1 ORF was also analyzed. Samples D2, D7, D13, D14, D15, D16, D30, D34 and D42 have previously been associated with the European variant according to nucleotide sequence analysis of the HPV16 E2 ORF [15]

Samples	E6 region											E7 region								
	1	1	1	1	1	2	2	3	3	4	5	6	7	7	7					
	3	3	4	4	7	8	8	3	5	0	3	4	3	8	9					
	1	2	3	5	8	6	9	5	0	3	2	7	2	9	5					
	A	G	C	G	T	T	A	C	T	A	A	A	T	T	T					
PNL2	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
PNL4	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
PNL5	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
PNL6	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
PNL7	-	-	-	T	-	A	G	T	G	-	G	-	C	C	G					
PNL8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
PNL10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
ATT4	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
ATT5	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
ATT8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
ATT9	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
ATT12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
ATT16	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
ATT20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
A17	-	T	G	T	-	A	G	T	-	G	-	G	-	C	G					
A24	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
A38	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
858	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
868	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
747	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
793	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-					
796	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
Lineage	Sublineage																			
European	European prototype I	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
	European prototype II	G	-	-	-	A	-	-	T	G	-	-	-	-	-					
Asian	European Asian	-	-	-	-	G / C	-	-	-	-	-	G	-	-	-					
Asian American	Asian American 1	-	-	-	T	-	A	G	T	G	-	G	-	C	C	G				
	Asian American 2	-	-	-	T	-	A	G	T	G	-	-/G	-	C	C	G				
	North American	-	-	-	T	-	A	G	T	G	-	-	-	-	C	G				
African type I	-	C	G	T	-	A	G	T	-	-	-	-	-	C	G					
African type II	-	T	G	T	-	A	G	T	-	G	-	G	-	C	G					
E2 region																				
Variant lineage	Accession number in Genbank					Samples					Variant lineage					Accession number in Genbank				
European	JQ422078					D2					European					HQ398221				
European	JQ422080					D7					European					JF422083				
European	JQ422082					D13					European					JQ864254				
European	JQ422083					D14					European					HQ398220				
Asian American	JQ422085					D15					European					JQ864253				
Prototype	JQ422087					D16					European					HQ398222				
Prototype	JQ422091					D30					European					HQ398216				
European	JQ422061					D34					European					JF422087				
European	JQ422063					D42					European					JF422091				

Table 2 continued

E2 region				
Variant lineage	Accession number in Genbank	Samples	Variant lineage	Accession number in Genbank
Prototype	JQ422065			
European	JQ422066			
Prototype	JX088614			
European	JQ422068			
Prototype	JQ422070			
African type II	JQ422056			
European	JQ422075			
European	JN205111			
European	JQ422101			
European	JQ422103			
European	JQ422116			
European	JQ422105			
Prototype	JQ422099			

nucleotide variation T2355C was only detected in the cervical sample A17 and in the reference sequence of the African type II variant (Table 5).

Finally, the genomic region between nucleotides 2443 and 2848 appears to be more conserved among the cloned E1 amplicons. The nucleotide variations A2587T and G2650A were found in high-grade samples (PNL7, D13, D14, and D30). Moreover, the silent nucleotide changes T2586C, A2608C and the T2631T were identified in the cervical samples A17 and PNL7 and in the reference sequences of the HPV16 African type II and Asian American variants (Table 6). However, the nucleotide change G2650A was only detected in the cancer sample PNL7 and in the reference sequence of the HPV16 Asian American variant (Table 6).

Discussion

The HPV16 E1 gene encodes an active DNA helicase that initiates the DNA replication of the episomal viral genome. The gene encoding the viral E1 protein is the largest viral ORF and is a highly conserved region among the various HPV types [20]. The present study aimed to investigate nucleotide sequence variations and amino acid substitutions in the HPV16 E1 ORF from Greek patients. Thus, the complete E1 gene was amplified, using four overlapping primer sets [32]. Nucleotide sequence analysis of the E6 and E7 genes from the same cervical samples was also conducted in the present study in order to investigate the association among E1 nucleotide variations and the various HPV16 variant lineages.

E1 gene disruption was identified in five high-grade and two low-grade cervical samples. It has been proposed that disruption of the E1 gene occurs mainly at the 3' end [31]. However, the present study found disruptions only between nucleotides 878 and 1523, located at the 5' end of the E1 gene [42]. The presence of a disrupted genome in low-grade samples reinforce the hypothesis that integration of HPV16 DNA occurs at an early stage of viral infection, leading to the progression of cervical intraepithelial neoplasia to invasive cancer [31, 33, 34, 43].

In the present study, nucleotide sequence analysis of the E6 and E7 genes from HPV16 genomes revealed that the most frequently identified HPV16 variants were the European variants and the HPV16 prototype strain. Previous studies have shown that HPV16 non-European variants are significantly more likely than European variants to cause high-grade cervical malignancies [44, 45]. In the present report, the European variant was detected both in high- and low-grade cervical intraepithelial neoplasia cases (Table 2). Moreover, the African type II variant was identified in the low-grade cervical case A17, while the Asian American variant was identified in the cervical cancer case PNL7. The Asian American variant has been strongly associated with cervical cancer development [46, 47].

In a previous report by Cornet et al. [16], it was proposed that nucleotide sequence polymorphisms within the HPV16 E6 and LCR regions may be used to further classify the major HPV16 variant lineages into HPV16 variant sublineages. However, a global alignment of cloned E1 sequences and HPV16 reference sequences that represented the majority of HPV16 diversity revealed that nucleotide polymorphisms in the E1 ORF might be sufficient to

Table 3 Sequence analysis of the E1 ORF between nucleotides 878 and 1523. Codes PNL4, PNL5, ATT4, ATT5, A17, A24, D34, D42, 747, 793, etc. correspond to the cervical samples. The number 645 corresponds to the size of the E1 amplicon, and the last letter (C1) corresponds to the number of the cloned E1 sequence. According to

sequence analysis, eight nucleotide variations were identified that were associated with HPV16 Asian American and African type II variants. However, the nucleotide variation A1041G was identified in the Asian American variant only, while the nucleotide change C1416T was detected only in the African type II variant

878–1523	1	1	1	1	1	1	1	1	1	1	Variant lineage	Accession number in GenBank
9	0	0	1	2	3	3	4	4	4			
2	4	9	6	0	6	7	1	2	8			
1	1	6	3	0	6	7	6	6	6			
T	A	C	G	T	T	C	C	C	T			
Clone (HG)												
PNL4 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791086
PNL5 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791089
PNL6 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791087
PNL6 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791088
PNL7 645 C1	C	G	G	A	C	A	T	-	G	C	Asian American	JQ791066
PNL8 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791067
PNL10 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791069
ATT4 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791090
ATT5 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791071
ATT12 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791075
ATT16 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791091
ATT20 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791073
D30 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791079
Clone (LG)												
PNL2 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791084
ATT8 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791070
A17 645 C1	C	-	G	A	C	A	T	T	G	C	African type II	JQ791076
A24 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791080
A38 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791092
793 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791095
796 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791082
858 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791094
868 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791097
D34 645 C2	-	-	-	-	-	-	-	-	-	-	European	JQ791078
D42 645 C1	-	-	-	-	-	-	-	-	-	-	European	JQ791077
Lineage												
European	-	-	-	-	-	-	-	-	-	-		
African type I	-	-	G	-	-	A	T	-	-	C		
African type II	C	-	G	A	C	A	T	T	G	C		
Asian American	C	G	G	A	C	A	T	-	G	C		
Asian	-	-	-	-	-	-	-	-	-	-		
Amino acid	F19F	L59L	Q78E	G100E	A112A	C168S	Y171Y	H184H	Q188E	L208L		

Table 4 Sequence analysis of E1 ORF between nucleotides 1503 and 1991. Codes PNL4, PNL5, ATT4, ATT5, A17, A24, D34, D42, 747, 793, etc. correspond to the cervical samples. The number 488 corresponds to the size of the E1 amplicon, and the last letter (C1) corresponds to the number of the cloned E1 sequence. According to

sequence analysis, three nucleotide variations were identified that were associated with Asian American and African type II variants. The nucleotide changes T1587C and A1668G were detected only in the Asian American variant

1503-1991	1	1	1	1	1	Variant lineage	Accession number in GenBank
	5	6	6	7	8		
	8	2	6	4	4		
	7	4	8	4	2		
	T	C	A	C	A		
Clone (HG)							
PNL4 488 C1	-	-	-	-	-	European	JQ768483
PNL5 488 C1	-	-	-	-	-	European	JQ768485
PNL6 488 C1	-	-	-	-	-	European	JQ768487
PNL7 488 C1	-	T	G	A	G	Asian American	JQ768488
PNL8 488 C1	C	-	-	-	-	European	JQ768490
PNL10 488 C1	-	T	G	A	G	European	JQ768491
ATT4 488 C1	-	-	-	-	-	European	JQ768492
ATT5 488 C1	C	-	-	-	-	European	JQ768504
ATT9 488 C1	C	-	-	-	-	European	JQ768496
ATT12 488 C1	C	-	-	-	-	European	JQ768498
ATT16 488 C1	-	-	-	-	-	European	JQ768500
ATT20 488 C1	-	-	-	-	-	European	JQ768502
D2 488 C1	-	-	-	-	-	European	JQ768462
D13 488 C1	-	-	-	-	-	European	JQ768476
D14 488 C1	-	-	-	-	-	European	JQ768477
D15 488 C1	-	-	-	-	-	European	JQ768478
D30 488 C1	-	-	-	-	-	European	JQ768463
Clone (LG)							
PNL2 488 C1	-	-	-	-	-	European	JQ768481
ATT8 488 C1	C	-	-	-	-	European	JQ768494
A17 488 C1	-	T	-	A	G	African type II	JQ768467
A24 488 C1	-	-	-	-	-	European	JX297818
A38 488 C1	-	-	-	-	-	European	JQ768468
747 488 C1	-	-	-	-	-	European	JQ768479
793 488 C1	-	-	-	-	-	European	JQ768470
796 488 C1	-	-	-	-	-	European	JQ768474
858 488 C1	-	-	-	-	-	European	JQ768472
868 488 C1	-	-	-	-	-	European	JQ768480
D7 488 C1	-	-	-	-	-	European	JQ768465
D34 488 C1	-	-	-	-	-	European	JQ768473
D42 488 C1	-	-	-	-	-	European	JQ768475
Lineage							
European	-	-	-	-	-		
African type I	-	T	-	T	G		
African type II	-	T	-	A	G		
Asian American	C	T	G	A	G		
Asian	-	-	-	-	G		
Amino acid	F241F	L254L	A268A	L294M	I326M		

Table 5 Sequence analysis of E1 ORF between nucleotides 1952 and 2485. Codes PNL4, PNL5, ATT4, ATT5, A17, A24, D34, D42, 747, 793, etc. correspond to the cervical samples. The number 533 corresponds to the size of the E1 amplicon, and the last letter (C1) corresponds to the number of the cloned E1 sequence. According to sequence analysis, seven nucleotide variations were identified that were associated with Asian American and African type II variants. The nucleotide change T2355C was identified only in the African type II variant

1952-2485	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	Variant lineage	Accession number in Gen Bank	
Clone (HG)																													
PNL4 533 C1	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791030		
PNL5 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791032		
PNL6 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791034		
PNL7 533 C1	T	-	-	-	C	G	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Asian American	JQ791035		
PNL8 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791037		
PNL10 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791036		
ATT4 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791017		
ATT5 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791019		
ATT9 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791021		
ATT12 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791022		
ATT16 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791024		
ATT20 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791026		
D2 533 C1	-	A	C	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791038		
D13 533 C1	-	A	C	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791053		
D14 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791044		
D15 533 C1	-	A	C	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791052		
D30 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791045		
Clone (LG)																													
PNL2 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791028		
ATT8 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791020		
A17 533 C1	T	-	-	-	C	G	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	African type II	JQ791062		
A24 533 C2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791061		
A38 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791063		
747 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791059		
793 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791050		
796 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791048		
858 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791057		
868 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791055		
D7 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791043		
D34 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791040		
D42 533 C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	European	JQ791047		
Lineage																													
European	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
African type I	T	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
African type II	T	-	-	-	C	G	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian American	T	-	-	-	C	G	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Amino acid	L393I	K403K	D435D	I442T	E425D	T458S	R462K	G466G	L475L	L494L	N530I	S497S	N530I	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 6 Sequence analysis of E1 ORF between nucleotides 878 and 1523. Codes PNL4, PNL5, ATT4, ATT5, A17, A24, D34, D42, 747, 793, etc. correspond to the cervical samples. The number 405 corresponds to the size of the E1 amplicon, and the last letter (C1) corresponds to the number of the cloned E1 sequence. According to

sequence analysis, three nucleotide variations were identified that were associated with the Asian American and African type II variants. Moreover, the nucleotide variation G2650A was only detected in the Asian American variant

2423-2848	2	2	2	2	2	Variant lineage	Accession number in Genbank
	5	5	6	6	6		
	8	8	0	3	5		
	6	7	8	1	0		
	T	A	A	T	G		
Clone (HG)							
PNL4 405 C1	-	-	-	-	-	European	JQ779917
PNL5 405 C1	-	-	-	-	-	European	JQ779919
PNL6 405 C1	-	-	-	-	-	European	JQ779921
PNL7 405 C1	C	-	C	A	A	Asian American	JQ779923
PNL8 405 C1	-	-	-	-	-	European	JQ779925
PNL10 405 C1	-	-	-	-	-	European	JQ779927
ATT4 405 C1	-	-	-	-	-	European	JQ779933
ATT5 405 C1	-	-	-	-	-	European	JQ779934
ATT9 405 C1	-	-	-	-	-	European	JQ779937
ATT12 405 C1	-	-	-	-	-	European	JQ779939
ATT16 405 C1	-	-	-	-	-	European	JQ779941
ATT20 405 C1	-	-	-	-	-	European	JQ779942
D2 405 C1	-	-	-	-	-	European	JQ779961
D13 405 C2	-	T	-	-	-	European	JQ779962
D14 405 C1	-	T	-	-	-	European	JQ779963
D15 405 C1	-	-	-	-	-	European	JQ779960
D30 405 C1	-	T	-	-	-	European	JQ779951
Clone (LG)							
PNL2 405 C1	-	-	-	-	-	European	JQ779915
ATT8 405 C1	-	-	-	-	-	European	JQ779936
A17 405 C1	C	-	C	A	-	African type II	JQ779944
A24 405 C2	-	-	-	-	-	European	JQ779946
A38 405 C1	-	-	-	-	-	European	JQ779947
747 405 C1	-	-	-	-	-	European	JQ779967
793 405 C1	-	-	-	-	-	European	JQ779929
796 405 C1	-	-	-	-	-	European	JQ779931
858 405 C1	-	-	-	-	-	European	JQ779965
868 405 C1	-	-	-	-	-	European	JQ779968
D7 405 C1	-	-	-	-	-	European	JQ779957
D34 405 C1	-	-	-	-	-	European	JQ779955
D42 405 C1	-	-	-	-	-	European	JQ779949
Lineage							
European	-	-	-	-	-		
African type I	C	-	-	A	-		
African type II	C	-	c	A	-		
Asian American	C	-	c	A	A		
Asian	-	-	-	-	-		
Amino acid	S574S	R575W	R582R	P589P	E596K		

classify the HPV16 genome in the major variant lineages. Moreover, a maximum-likelihood phylogenetic tree indicated that the HPV16 E1 ORF can be used to classify the HPV16 genome in distinct phylogenetic branches with high confidence (bootstrap values >80 %; Fig. 1). Furthermore, a comparative analysis of the E6, E7 and E1 regions from the same samples was done in the present study in order to associate E1 polymorphisms with particular intratypic polymorphisms within the E6 and E7 genes. In particular, within the E1 ORF, the nucleotide variations T921C, C1096G, G1163A, T1200C, T1366A, C1377T, C1426G, T1486C, C1624T, C1744A, A1842G, C2041T, G2220C, C2237G, G2249A, C2262T, C2287T, C2344T, T2586C, A2608C, T2631T were detected in both the HPV16 African type II and Asian American variant lineages (Fig. 1). These sequence changes resulted in the G145T, T286A, A289G, C335T, T789C, and T795G sequence variations that are encountered in the E6, E7 sequences of Asian American and African type II variant lineages. Moreover, the nucleotide variations A1041G, C1587T, A1668G, G2650A were detected only in the Asian American variant lineage (E6, E7: A532G and T732C), while the C1416T and T2355C variations were associated with the HPV16 African type II variant lineage (E6: A403G) (Fig. 1) [9, 13, 16, 17] (Tables 3, 4, 5 and 6).

We conclude that nucleotide changes within the E1 ORF can be used to correctly classify the HPV16 genome in any of the distinct variant lineages. To the best of our knowledge, this is the first report analyzing the HPV16 E1 ORF in order to identify nucleotide variations and amino acid substitutions in high- and low-grade cervical malignancies. However, the presence of an E1 gene disruption in low-grade malignancies indicates that the E1 region might be a noteworthy genomic region for the study of the physical state of the HPV16 genome in low-grade malignancies.

Acknowledgments This work was supported by research grants from the Postgraduate Programmes ‘Applications of Molecular Biology-Genetics. Diagnostic Biomarkers’, code 3817, and “Biotechnology”, code 3439, of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology.

Conflict of interest All authors declare that they have no conflicting or dual interests.

References

- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen Hz, de Villiers EM (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H (2004) Classification of papillomaviruses. *Virology* 324:17–27
- de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, Tous S, Felix A, Bravo LE, Shin HR, Vallejos CS, de Ruiz PA, Lima MA, Guimera N, Clavero O, Alejo M, Llobart-Bosch A et al (2010) Retrospective International Survey and HPV Time Trends Study Group. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 11:1048–1056
- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127:2893–2917
- Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM (2011) Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer* 128:927–935
- Bernard HU (2005) The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J Clin Virol* 32:S1–S6
- Gottschling M, Göker M, Stamatakis A, Bininda-Emonds OR, Nindl I, Bravo IG (2011) Quantifying the phylodynamic forces driving papillomavirus evolution. *Mol Biol Evol* 28:2101–2113
- Chan SY, Bernard HU, Ong CK, Chan SP, Hofmann B, Delius H (1992) Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 66:5714–5725
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jensen SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743–7753
- Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD (2005) Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 79:7014–7023
- Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P (1993) The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 67:6413–6423
- Ho L, Chan SY, Chow V, Chong T, Tay SK, Villa LL, Bernard HU (1991) Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol* 29:1765–1772
- Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM (1997) Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* 71:2463–2472
- Carvajal-Rodríguez A (2008) Detecting recombination and diversifying selection in human alpha-papillomavirus. *Infect Genet Evol* 8:689–692
- Tsakogiannis D, Ruether IGA, Kyriakopoulou Z, Pliaka V, Skordas V, Gartzonika C, Levidiotou-Stefanou S, Markoulatos P (2012) Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece. *Arch Virol* 157:1729–1739
- Cornet I, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, Tommasino M, Clifford GM, IARC HPV Variant Study Group (2012) Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol* 86:6855–6861
- Smith B, Chen Z, Reimers L, van Doorslaer K, Schiffman M, Desalle R, Herrero R, Yu K, Wacholder S, Wang T, Burk RD (2011) Sequence imputation of HPV16 genomes for genetic association studies. *PLoS One* 6:e21375
- Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, Sahasrabudhe VV, Gravitt PE, Hsing AW, Burk RD (2011)

- Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One* 6:e20183
19. Huertas-Salgado A, Martín-Gámez DC, Moreno P, Murillo R, Bravo MM, Villa L, Molano M (2011) E6 molecular variants of human papillomavirus (HPV) type 16: an updated and unified criterion for clustering and nomenclature. *Virology* 410:201–215
 20. García-Vallvé S, Alonso A, Bravo IG (2005) Papillomaviruses: different genes have different histories. *Trends Microbiol* 13:514–521
 21. Titolo S, Brault K, Majewski J, White PW, Archambault J (2003) Characterization of the minimal DNA binding domain of the human papillomavirus e1 helicase: fluorescence anisotropy studies and characterization of a dimerization-defective mutant protein. *J Virol* 77:5178–5191
 22. Titolo S, Pelletier A, Pulichino AM, Brault K, Wardrop E, White PW, Cordingley MG, Archambault J (2000) Identification of domains of the human papillomavirus type 11 E1 helicase involved in oligomerization and binding to the viral origin. *J Virol* 74:7349–7361
 23. Amin AA, Titolo S, Pelletier A, Fink D, Cordingley MG, Archambault J (2000) Identification of domains of the HPV11 E1 protein required for DNA replication in vitro. *Virology* 272:137–150
 24. White PW, Pelletier A, Brault K, Titolo S, Welchner E, Thauvette L, Fazekas M, Cordingley MG, Archambault J (2001) Characterization of recombinant HPV6 and 11 E1 helicases: effect of ATP on the interaction of E1 with E2 and mapping of a minimal helicase domain. *J Biol Chem* 276:22426–22438
 25. Stenlund A (2003) E1 initiator DNA binding specificity is unmasked by selective inhibition of non-specific DNA binding. *EMBO J* 22:954–963
 26. Clower RV, Hu Y, Melendy T (2006) Papillomavirus E2 protein interacts with and stimulates human topoisomerase I. *Virology* 348:13–18
 27. Loo YM, Melendy T (2004) Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *J Virol* 78:1605–1615
 28. Parker LM, Harris S, Gossen M, Botchan MR (2000) The bovine papillomavirus E2 transactivator is stimulated by the E1 initiator through the E2 activation domain. *Virology* 270:430–443
 29. Conger KL, Liu JS, Kuo SR, Chow LT, Wang TS (1999) Human papillomavirus DNA replication. Interactions between the viral E1 protein and two subunits of human dna polymerase alpha/primase. *J Biol Chem* 274:2696–2705
 30. Masterson PJ, Stanley MA, Lewis AP, Romanos MA (1998) A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit. *J Virol* 72:7407–7419
 31. Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbini M (2009) Disruption of HPV16 E1 and E2 genes in precancerous cervical lesions. *J Virol Methods* 158:180–183
 32. Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM (2006) Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 44:1755–1762
 33. Huang LW, Choa SL, Lee BH (2008) Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. *J Clin Pathol* 61:627–631
 34. Kulmala SM, Syrjänen SM, Gyllensten UB, Shabalova IP, Petrovichev N, Tosi P, Syrjänen KJ, Johansson BC (2006) Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 59:513–517
 35. Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25–36
 36. Tsakogiannis D, Ruether IG, Kyriakopoulou Z, Pliaka V, Theoharopoulou A, Skordas V, Panotopoulou E, Nepka C, Markoulatos P (2012) Sequence variation analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from women in Greece. *Arch Virol* 157:825–832
 37. Puranen M, Saarikoski S, Syrjänen K, Syrjänen S (1996) Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391–395
 38. Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D (2008) The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355–1361
 39. Sotlar K, Diemer D, Dethleffs A, Hack Y, Stubner A, Vollmer N, Menton S, Menton M, Dietz K, Wallwiener D, Kandolf R, Bültmann B (2004) Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol* 42:3176–3184
 40. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
 41. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
 42. Kalantari M, Karlsen F, Kristensen G, Holm R, Hagmar B, Johansson B (1998) Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *Int J Gynecol Pathol* 17:146–153
 43. Briolat J, Dalstein V, Saunier M, Joseph K, Caudroy S, Prétet JL, Birembaut P, Clavel C (2007) HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. *Int J Cancer* 121:2198–2204
 44. Schiffman M, Rodriguez AC, Chen Z, Wacholder S, Herrero R, Hildesheim A, Desalle R, Befano B, Yu K, Safaeian M, Sherman ME, Morales J, Guillen D, Alfaro M, Hutchinson M, Solomon D, Castle PE, Burk RD (2010) A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. *Cancer Res* 70:3159–3169
 45. Sicheo L, Ferreira S, Trottier H, Duarte-Franco E, Ferenczy A, Franco EL, Villa LL (2007) High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *Int J Cancer* 120:1763–1768
 46. Quint KD, de Koning MN, van Doorn LJ, Quint WG, Pirog EC (2010) HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix. *Gynecol Oncol* 117:297–301
 47. Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Greggi S, Buonaguro L, Buonaguro FM (2011) Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 121:32–42

Prevalence of HPV16 E1-1374^{63nt} Variants in Greek Women

D. Tsakogiannis,¹ Z. Kyriakopoulou,¹ F. Darmis,¹ I.G.A. Ruether,¹ T.G. Dimitriou,¹ G. Orfanoudakis,² E. Panotopoulou,³ and P. Markoulatos^{1*}

¹Department of Biochemistry & Biotechnology, Microbiology-Virology Laboratory, School of Health Sciences, University of Thessaly, Larissa, Greece

²Oncoprotein Group, University of Strasbourg, CNRS FRE 3211, The Biotechnology School of Strasbourg, ESBS, University of Strasbourg, Illkirch, France

³Papanicolaou Research Centre of Oncology and Experimental Surgery, Anticancer Oncology Hospital of Athens "St Savvas", Athens, Greece

Recent studies have focused on sequence variation of the HPV16 E1 gene. The present study investigates the prevalence of E1-1374^{63nt} duplication in the Greek population, and the sequence variation at the 5' end of the E1 and E6 genes from samples that harbored this genetic alteration. Fifty HPV16 positive cervical samples, derived from Greek patients were investigated. The 5' end of the E1 gene was amplified through PCR and the variant amplicons were cloned, sequenced, and bioinformatically analyzed for selective pressure. The E1-1374^{63nt} duplication was identified in 24% of the examined samples, with the same prevalence in both high and low-grade cervical malignancies. The E1-1374^{63nt} duplication was linked to the European variant lineage ($\chi^2=5.076$, $P<0.024$) and it was significantly associated with the nucleotide variation A1053C ($\chi^2=23.102$, $P<0.0001$). Molecular evolution analyses anticipate that the E1-1374^{63nt} duplication induces functional constraints on the 5' end of E1 gene, and it is proposed that this duplication might not affect negatively the function or structure of the E1 protein. The E1-1374^{63nt} duplication is prevalent in the Greek population, whereas the A1053C variation might constitute a significant marker for the characterization of the E1-1374^{63nt} variant in the Greek population, thus providing significant information about viral pathogenicity.

J. Med. Virol. © 2014 Wiley Periodicals, Inc.

KEY WORDS: HPV16; E1 gene; E1-1374^{63nt} duplication; T350G variation; A1053C variation; natural selection

INTRODUCTION

Human papillomaviruses (HPVs) are non-enveloped, epitheliotropic, double stranded, circular DNA viruses that infect cutaneous or mucosal epithelia [Muñoz, 2000]. HPVs are classified in different types, subtypes, and variants based on sequence identity of the L1 capsid gene [de Villiers et al., 2004; Bernard et al., 2010]. Human papillomavirus type 16 (HPV16) is identified most frequently among HPV types and is implicated in 50% of cervical cancer cases [De Sanjose et al., 2010; Ferlay et al., 2010; Li et al., 2011]. The HPV16 type is classified in five major intratypic variant lineages known as HPV16 European, European Asian, African type II, African type I, and Asian American [Ho et al., 1991, 1993; Chen et al., 2005]. Nucleotide polymorphism of HPV16 variants in genomic regions such as E6, L1, L2, and LCR expanded this phylogenetic classification [Yamada et al., 1995, 1997]. Previously it was demonstrated that intratypic variations within the E1 and E4 genes suitably

Grant sponsor: Postgraduate Programmes 'Applications of Molecular Biology-Genetics. Diagnostic Biomarkers', code 3817 and "Biotechnology", code 3439, of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology

Conflict of Interest: None.

Ethical Approval: All patients gave their written consent for each cervical specimen that was collected for HPV diagnostic and research purposes. The project was approved by the Ethical Board of Anticancer Oncology Hospital of Athens "St Savvas."

*Correspondence to: Professor P. Markoulatos, Department of Biochemistry & Biotechnology, School of Health Sciences, University of Thessaly, Ploutonos 26 & Aiolou, 41221 Larissa, Greece.

E-mail: markoulatos@bio.uth.gr

Accepted 14 January 2014

DOI 10.1002/jmv.23896

Published online in Wiley Online Library (wileyonlinelibrary.com).

classified HPV16 genomes in four distinct phylogenetic branches [Tsakogiannis et al., 2012, 2014]. Epidemiologic studies have revealed that different HPV16 variants vary in pathogenesis and progression of cervical malignancy in cancer disease [Quint et al., 2010; Sanchez et al., 2011; Tornesello et al., 2011].

Nucleotide analysis of the various HPV genes highlighted the E1 ORF as the most conserved. The E1 protein displays both ATPase and 3'-5' helicase activities and cooperates with the E2 protein in order to form an active replication complex, responsible for HPV DNA replication [Masterson et al., 1998; Conger et al., 1999; Parker et al., 2000; Loo and Melendy, 2004; Clower et al., 2006]. Nevertheless, sequence analysis of the E1 gene revealed that the region from nucleotide 1,311–1,374 was found to be duplicated in tandem. The duplication of 63 nucleotides (E1-1374^{63nt} duplication) was first reported by Dong et al. [2000], who investigated HPV16 positive cervical cancer cases in the Chinese Population. This intragenic duplication was also recently described in Croatia and Slovenia [Sabol et al., 2008, 2012; Bogovac et al., 2011]. However, previous reports demonstrated that the E1-1374^{63nt} duplication was linked with the European variant and was associated preferentially with low-grade cervical intra-epithelial neoplasia cases [Sabol et al., 2008, 2012; Bogovac et al., 2011]. The E1-1374^{63nt} duplication is placed at the N-terminal domain of the E1 protein, that is responsible for binding of the E1 protein–H1 histone [Swindle and Engler, 1998; Wilson et al., 2002]. The exact impact of this sequence duplication in E1 protein function and the molecular mechanism that induces this genetic alteration in HPV16 genome remains unknown.

Considering the presence of the E1-1374^{63nt} duplication in Croatia and Slovenia, the current study focused on prevalence of this duplication in Greek women. Thus the 5' end of the E1 gene was amplified through PCR, and sequence analysis was

performed in order to identify sequence variations in cervical samples that harbored the E1-1374^{63nt} duplication. In addition, sequence analysis of E6 and E7 genes in the same cervical samples was performed in order to investigate the association of E1-1374^{63nt} duplication with the nucleotide variation T350G. Furthermore, molecular evolutionary analysis was also performed in order to identify potential functional constraints acting on the E1 gene.

MATERIALS AND METHODS

Cervical Specimens

Fifty HPV16 positive cervical samples of Greek women were diagnosed as cervical intra-epithelial neoplasia stages I, II, and III. The cervical samples were derived from Thin Prep and cervical stained smears. Of the 50 samples, 27 were diagnosed as high-grade cervical intra-epithelial neoplasias and 23 as low-grade cervical intra-epithelial neoplasias.

DNA Extraction

DNA from Thin Prep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) [Casas et al., 1995], while DNA from cervical stained smears was extracted as previously described [Puranen et al., 1996]. The quality of extracted DNA was assayed by amplification of a 498 bp beta-actin fragment [Li et al., 2008]. The identification of HPV16 genome was performed through nested multiplex PCR [Sotlar et al., 2004].

Amplification of 5' End of the E1 Gene

The 5' end of the E1 gene was amplified using the primer set E1F1–E1R3 [Arias-Pulido et al., 2006] (Table I). PCR was performed in a final volume of 50 µl. PCR mixture contained 50 pmol of the primer set (E1F1–E1R3), 5× Green GoTaq Flexi buffer (Promega, Madison, WI), 2 mM MgCl₂, 1 mM dNTPs,

TABLE I. Distribution of the E1-1374^{63nt} Duplication in High (18.5%) and Low-Grade (30.4%) Malignancies and the Association With the T350G Sequence Variation ($\chi^2 = 5.076$, $P < 0.024$) With the Corresponding

Cervical samples (n = 50)	E1 gene E1-1374 ^{63nt} duplication/A1053C (E63D) GenBank accession number	E6 gene T350G (L83V) GenBank accession number	Variants
High-grade lesions (n = 27)			
ATT4	JQ791090	JQ422080	European
ATT16	JQ791091	JQ422082	European
PNL4	JQ791086	JQ422083	European
PNL5	JQ791065	JX291526	European
PNL6	JQ791087	JX291528	European
Low-grade lesions (n = 23)			
A38	JQ791092	JN205112	European
858	JQ791094	JQ422101	European
793	JQ791095	JQ422103	European
868	JQ791097	JQ422106	European
PNL2	JQ791084	JQ422078	European
ATT6	JX291529	JQ422061	European
ATT21	JX291531	JQ422068	European

and 1.25 U of thermostable DNA polymerase (GoTaq[®] DNA Polymerase, Promega). PCR cycling conditions were as follows: an initial denaturation step at 95°C for 2 min, then 40 cycles of 95°C for 40 sec, 50°C for 30 sec, and 72°C for 30 sec. PCR was ended with a 5 min incubation step at 72°C.

Amplification of the E6 and E7 Genes

E6 sequence analysis was performed in order to investigate for the association between the T350G sequence variation and the E1-1374^{63nt} duplication. The E7 gene was also analyzed in order to examine sequence variations that might be associated with the E1-1374^{63nt} duplication. Thus, the E6–E7 region was amplified using the primer set HPV16 41/HPV16 757 as has been recently reported [Tsakogianis et al., 2013].

Sequence Analysis of the E1, E6, and E7 Genes

The HPV16 E1 variant amplicons (of approximately 708 bp in size) and the E6, E7 genes derived from the same cervical samples were subjected to cloning using the pGEM T-easy vector system (Promega). Recombinant plasmid DNA was purified using the Nucleospin plasmid kit (Macherey-Nagel, Duren, Germany), and the plasmids were subjected to sequencing at Macrogen, Seoul, Korea. In order to confirm nucleotide variations in E1 amplicons, E6, and E7 genes, three independent PCR assays were performed and amplicons from distinct PCR assays were subjected to cloning. Two colonies were collected from each individual specimen and both strands were sequenced. The DNA strands were sequenced three times to ensure high-reliability of the results.

Multiple sequence alignment was carried out using the program MEGA v.5 [Tamura et al., 2007] and the Muscle algorithm [Edgar, 2004], between E1 cloned sequences that harbored the E1-1374^{63nt} duplication, and the reference sequence of HPV16 genome (HPV16R) available in the HPV16 Sequence Database (Los Alamos National Laboratory, Los Alamos, NM). However, in order to discern the association between the E1-1374^{63nt} duplication and HPV16 variants, multiple sequence alignments were performed along with E6 and E7 cloned sequences, the reference sequence of HPV16 genome, and those of representative sequences of HPV16 variants available online in GenBank sequence database, under accession numbers for European Prototype 1 variant, HQ644236, HQ644259, HQ644267, HQ644268, HQ644271, AY686581, AY686580, HQ644272, AY686584, AY686583, HQ644274 E, HQ644280, HQ644282, HQ644283, HQ644284, HQ644286, HQ644287, HQ644297; for European prototype 2 variant, AF536179; for European-Asian variant HQ644234, HQ644235, HQ644248, HQ644251, HQ644261, AF534061; for African type I variant HQ644238, HQ644240, AF472508, HQ644290, HQ644293, AF536180, HQ644296, HQ644298, HQ644299; for African type II

variant HQ644237, HQ644239, HQ644249, HQ644250, AF472509, HQ644291, HQ644292; for the Asian American 1, HQ644247, HQ644253, HQ644255, AF402678, HQ644265, HQ644269, HQ644276, HQ644278, HQ644285, HQ644288, HQ644289; for the Asian American 2, HQ644254, HQ644270, AY686579, AY686582, HQ644273, HQ644277, HQ644279, HQ644281; and for Asian American/North American, HQ644257.

Statistical Analysis

Statistical analysis was performed by Chi square test using GraphPad Prism version 4.00 for Windows, (GraphPad Software, San Diego, CA). *P*-values were considered significant at the 0.05 cut-off level.

Analysis of Functional Constraints in the E1 Gene

Natural selection acting on the E1 fragment with the E1-1374^{63nt} duplication was investigated by estimating the dN (non-synonymous changes per non-synonymous site) and dS (synonymous changes per synonymous site) rates, with the SLAC (single likelihood ancestor counting) method in Datamonkey [Suzuki and Gojobori, 1999; Pond and Frost, 2005]. In addition, maximum likelihood (ML) and empirical Bayesian (EB) methods were used in order to identify positively selected sites. According to the ML method, the fixed effects likelihood model (FEL) in Datamonkey server [Suzuki and Gojobori, 1999; Pond and Frost, 2005] was used to identify positive selected sites within the 5' end of the E1 gene that carried out the duplication of 63 nucleotides. Under the FEL method, *P*-value cut-off was set at the 0.05 level in order to assign a codon as no false positively selected. EB analysis was performed by the SELECTON server, using the highest precision level [Stern et al., 2007]. In addition, Bayesian analyses using Markov chain Monte Carlo (MCMC) was implemented on the E1 dataset with E1-1374^{63nt} duplication using BEAST v.1.6.2 in order to calculate the median mutation rate in the first, second, and third position of the codon [Drummond and Rambaut, 2007]. Bayesian analysis consisted of four chains with random starting trees. Results were examined using the program Tracer v1.5 from the BEAST package. Convergence was assessed with ESS (effective sample size) values, after a burning of 1 million steps. The TrN substitution model was used [Tamura and Nei, 1993]. The best nucleotide substitution model was selected according to Akaike information criteria (AIC) by the FindModel online program (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html).

RESULTS

E1 Duplication Analysis

The HPV16 variant amplicon of approximately 708 bp in size was detected in 24% (12/50) of the

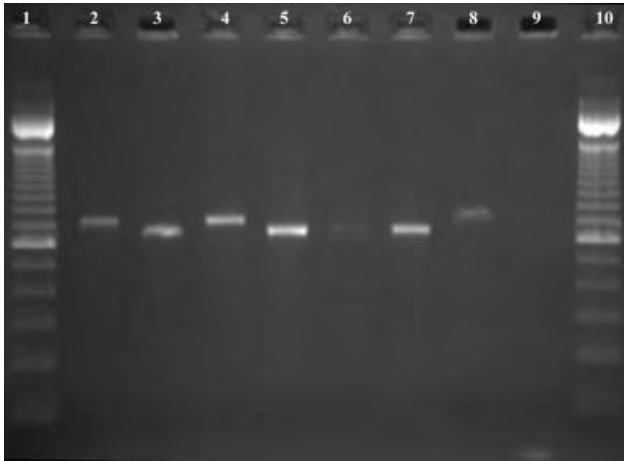


Fig. 1. Agarose gel electrophoresis of amplified fragments from the 5' end of E1 gene in seven cervical samples. Lines 1, 10, molecular weight marker, 100–1,500 bp in 100 bp increments and an additional fragment at 2,017 bp (100 bp DNA Ladder Invitrogen, Life Technologies, Carlsbad, CA); lines 2, 4, and 8, PCR products corresponding to variant amplicons of 708 bp in size, derived from cervical samples PNL4, PNL5, and PNL6, respectively; lines 3, 5, 6, and 7, PCR products corresponding to the normal amplicons of 645 bp in size; line 9 the negative control (ddH₂O).

examined cervical samples, while in 76% (38/50) of cervical specimens an ordinary amplicon of 645 bp in size was observed (Fig. 1). Sequence analysis of the E1 variant amplicons revealed that the region from nucleotide 1,311 to nucleotide 1,374 was found to be duplicated in tandem. In the present analysis, 7 out of 23 low-grade cervical samples (7/23, 30.4%) and 5 out of 27 high-grade cases (5/27, 18.5%) harbored the E1-1374^{63nt} duplication (Table I). In particular, the E1-1374^{63nt} duplication was detected in the low-grade samples A38, 858, 793, 868, PNL2, ATT6, ATT21, and in the high-grade cervical samples ATT4, ATT16, PNL4, PNL5, PNL6 (Table I). However, no statistically significant association among the E1-1374^{63nt} duplication and the severity of cervical malignancy was observed. According to amino acid analysis, it was observed that the E1-1374^{63nt} duplication does not interrupt the open reading frame of the subsequent region of the E1 gene, and causes the repetition of amino acids from 150 to 170.

Previous reports have demonstrated that the E1-1374^{63nt} duplication is linked with the nucleotide variation T350G within the E6 gene and the European variant lineage [Sabol et al., 2008, 2012; Bogovac et al., 2011]. Sequence analysis of the E6 gene was performed on all cervical samples that harbored the E1-1374^{63nt} duplication. According to E6 sequence analysis it was revealed that all cervical samples with the E1-1374^{63nt} duplication carried out the nucleotide variation T350G, and these samples presented highest sequence similarity with the European variant lineage (Table I). In addition, in order to

study the relation between the nucleotide variation T350G and the E1-1374^{63nt} duplication in the cervical specimens derived from Greek population, 21 cervical samples with regular E1 sequence (645 bp) were subjected to E6 sequence analysis. The T350G variation was detected in 12 cervical samples that were characterized as European variants (JQ422097, JQ422072, JQ422098, JQ422063, JQ422068, JX193802, JQ422089, JX193811, JX193810, JX193803, JQ422116, and JX193806), while it was not identified in seven cervical samples that had highest sequence similarity with the reference sequence of HPV16 genome (JQ422087, JQ422091, JQ422070, JQ422065, JQ422094, JQ422108, and JQ422117), and in two cervical samples that were characterized as African type I and African type II variants (JQ422059, JQ422056). Statistical analysis revealed that the E1-1374^{63nt} duplication was associated with the T350G sequence variation and the European variant lineage when compared with the presence of T350G nucleotide variation in cervical samples with regular sequence of E1 gene ($\chi^2 = 5.076$, $P = 0.024$). In addition, sequence analysis of the E7 gene in cervical samples with the E1-1374^{63nt} duplication revealed that there is no sequence variation associated with this duplication.

A previous study found that the nucleotide variation A1053C was present in all analyzed cervical samples that carried out the E1-1374^{63nt} duplication [Sabol et al., 2012]. The nucleotide variation A1053C induces the amino acid substitution E63D in the N-terminal domain of the E1 protein. In the present study, nucleotide analysis of the 5' end of the E1 gene revealed that the A1053C was only detected in cervical samples that harbored the E1-1374^{63nt} duplication. In order to investigate the association of A1053C with the duplication of 63 nucleotides, 15 cervical cases with regular amplicons were subjected to sequence analysis (JQ791078, JQ791077, JQ791079, JQ791076, JQ791080, JQ791082, JQ791066, JQ791067, JQ791069, JQ791071, JQ791070, JQ791075, JQ791074, JX297816, and JX297817). The sequence variation A1053C was not identified in the regular amplicons. Consequently, the A1053C sequence variation was significantly correlated with E1-1374^{63nt} duplication when compared with E1 regular sequences ($\chi^2 = 23.102$, $P < 0.0001$). In addition, in order to discern the presence of A1053C sequence variation in any of the distinct HPV16 variants, multiple sequence alignments were performed between the E1 variant sequences, and the reference sequences of HPV16 variants that were described above. Multiple sequence analysis revealed that the sequence variation A1053C was only related to cervical samples that carried out the E1-1374^{63nt} duplication.

Functional Constraints Analysis

Considering that the E1-1374^{63nt} duplication causes the repetition of amino acids from 150 to 170

within the E1 protein, the dN/dS ratio was calculated in order to investigate any selective pressure acting upon the N-terminal domain of the E1 protein and to investigate the functional constraints acting upon the E1 gene due to the E1-1374^{63nt} duplication. SLAC analysis estimated a dN/dS ratio of 0.56 and indicated that negative selective pressure is acting upon the N-terminal domain of the E1 protein that harbored the duplication of 21 amino acids (150–170). In addition, the median mutation rate in the first, second, and third codon position was calculated by Bayesian analyses, using MCMC and it was estimated that CP1 = 0.918, CP2 = 0.722, and CP3 = 1.36. Considering the results derived from the dN/dS ratio and the median mutation rate in the first, second, and third codon position, the evolution of these sequences was dominated by purifying selection it was clearly suggested that functional constraints dominated in the 5' region of the E1-1374^{63nt} variants. However, no evidence of positively selected sites was recorded in variant amplicons according to ML and EB methods.

DISCUSSION

The main objective of the present study was to investigate the prevalence of E1-1374^{63nt} duplication in cervical samples derived from Greek women. Previous reports proposed that the E1-1374^{63nt} duplication was linked with the European variant lineage [Dong et al., 2000; Sabol et al., 2008, 2012; Bogovac et al., 2011]. So far, this intragenic duplication has been described only in China, Croatia, and Slovenia [Dong et al., 2000; Sabol et al., 2008, 2012; Bogovac et al., 2011]. The presence of E1-1374^{63nt} duplication was confirmed in 8% and 10% of Slovenian and Croatian samples, respectively [Sabol et al., 2008, 2012; Bogovac et al., 2011]. In the present study, 24% of the investigated cervical samples derived from HPV16 positive Greek women, carried out the E1-1374^{63nt} duplication. The higher prevalence of the E1-1374^{63nt} variant in Greek population might be associated with an elevated geographic distribution of this variant in East-European countries. However, further samples derived from Greece and other European and Asian surrounding countries are required in order to confirm this hypothesis.

In addition, sequence analysis of the E6 gene in cervical cases that harbored this intragenic duplication revealed a significant association between the E1-1374^{63nt} duplication and the sequence variation T350G within E6 gene ($\chi^2 = 5.076$, $P = 0.024$). This result confirms that the E1-1374^{63nt} duplication is strongly associated with the European variant lineage [Sabol et al., 2008, 2012; Bogovac et al., 2011]. However, no statistically significant association among high and low-grade malignancies, and the E1-1374^{63nt} duplication was observed. The E1-1374^{63nt} duplication was prevalent in both high (18.5%) and low-grade (30.4%) dysplasias. However,

in a previous report by Sabol et al. [2012], the E1-1374^{63nt} duplication is linked with the E-G350 and the E-C109/G350 variants. In addition, it has been proposed that the E1-1374^{63nt} E-G350 variant was significantly associated with low-grade cervical malignancies, while the E1-1374^{63nt} E-C109/G350 was prevalent in both high and low-grade dysplasias, when compared with the European prototype strain [Sabol et al., 2012]. The prevalence of the E1-1374^{63nt} E-G350 and the E1-1374^{63nt} E-C109/G350 variants in Greek population among high and low-grade malignancies needs to be investigated further in order to better understand any association of the E1-1374^{63nt} variant with the severity of cervical malignancies.

Selective pressure analysis by estimating the dN/dS ratio revealed that negative selective pressure is acting upon coding sequences that harbor the E1-1374^{63nt} duplication (dN/dS = 0.56). These results indicate that functional constraints occur in the 5' region of the E1-1374^{63nt} variants [Vandepoele et al., 2005]. Consequently, it was assumed that the E1-1374^{63nt} duplication might not affect negatively the function or structure of the N-terminal domain of the E1 protein. Two potential scenarios might explain the presence of the E1-1374^{63nt} duplication. In the first scenario, since the E1 gene is highly conserved, frequent homologous recombination events between distinct HPV16 European variants may occur, thus resulting to the duplication of nucleotide sequences. In the second scenario, the association between the European variant and the intragenic duplication might indicate the presence of new HPV16 European sub-variant lineages [Sabol et al., 2008].

Nucleotide analysis of the E1 variant amplicons identified the reported A1053C variation [Sabol et al., 2012] and found a significant prevalence in all cervical samples that harbored the E1-1374^{63nt} duplication. The nucleotide variation A1053C induces the amino acid substitution E63D and is located at the N-terminal domain of the E1 protein. This protein domain is thought to be responsible for the binding of E1 protein to H1 histone [Swindle and Engler, 1998; Wilson et al., 2002]. A significant association among the nucleotide variation A1053C and the E1-1374^{63nt} duplication of ($\chi^2 = 23.102$, $P < 0.0001$) was detected in the present study. In addition, this mutation was neither identified in any of the regular amplicons that were sequenced in the present study, nor to the reference sequence of the five distinct HPV16 variant lineages. Taking these data into account the present study suggests that there is a strong relation among the E1-1374^{63nt} duplication and the nucleotide variation A1053C, and it was proposed that this particular mutation may predict the presence of E1-1374^{63nt} variant in the studied populations. However, previous three dimensional structure representations that predicted the E1 protein structure showed that both E1-1374^{63nt} duplication and E63D amino acid substitution were

sited on the surface of the E1 protein, and no drastic changes in the function or the structure of E1 protein occurred [Sabol et al., 2012]. These data reinforce the above hypothesis of this study, that functional constraints dominate the 5' region of the E1-1374^{63nt}, and might not affect negatively the function or the structure of the N-terminal domain of E1 protein.

In conclusion, little is known about sequence variation of the E1 gene. On the other hand, it was recently established that 27 nucleotide variations within E1 gene adequately classify the HPV16 genome in any of the four distinct HPV16 phylogenetic branches [Tsakogiannis et al., 2014]. Nucleotide analyses of the 5' end of the E1 gene revealed that the E1-1374^{63nt} duplication was present in 24% of the investigated cervical samples derived from Greek women. Furthermore, the E1-1374^{63nt} variant might have a different geographic distribution pattern in Europe. In addition, the E1-1374^{63nt} duplication seems to not affect negatively the function or structure of the N-terminal domain of the E1 protein, as it was concluded from molecular evolution analyses. Finally, it was concluded that the sequence variation A1053C was significantly associated with the E1-1374^{63nt} duplication and may characterize the E1-1374^{63nt} variant in the studied population, providing significant information about viral evolution, and oncogenicity.

REFERENCES

- Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM. 2006. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 44:1755–1762.
- Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79.
- Bogovac Z, Lunar MM, Kocjan BJ, Seme K, Jančar N, Poljak M. 2011. Prevalence of HPV 16 genomic variant carrying a 63bp duplicated sequence within the E1 gene in Slovenian women. *Acta Dermatovenerol Alp Pannonica Adriat* 20:135–139.
- Casas I, Powell L, Klapper PE, Cleator GM. 1995. New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 53:25–36.
- Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD. 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 79:7014–7023.
- Clower RV, Hu Y, Melendy T. 2006. Papillomavirus E2 protein interacts with and stimulates human topoisomerase I. *Virology* 348:13–18.
- Conger KL, Liu JS, Kuo SR, Chow LT, Wang TS. 1999. Human papillomavirus DNA replication. Interactions between the viral E1 protein and two subunits of human DNA polymerase alpha/primase. *J Biol Chem* 274:2696–2705.
- De Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, Tous S, Felix A, Bravo LE, Shin HR, Vallejos CS, de Ruiz PA, Lima MA, Guimera N, Clavero O, Alejo M, Llombart-Bosch A, Cheng-Yang C, Tatti SA, Kasamatsu E, Ijazovic E, Odida M, Prado R, Seoud M, Grce M, Usubutun A, Jain A, Suarez GA, Lombardi LE, Banjo A, Menéndez C, Domingo EJ, Velasco J, Nessa A, Chichareon SC, Qiao YL, Lerma E, Garland SM, Sasagawa T, Ferrera A, Hammouda D, Mariani L, Pelayo A, Steiner I, Oliva E, Meijer CJ, Al-Jassar WF, Cruz E, Wright TC, Puras A, Llave CL, Tzardi M, Agorastos T, Garcia-Barriola V, Clavel C, Ordi J, Andújar M, Castellsagué X, Sánchez GI, Nowakowski AM, Bornstein J, Muñoz N, Bosch FX. 2010. Retrospective International Survey and HPV Time Trends Study Group. Human papillomavirus genotype attribution in invasive cervical cancer: A retrospective cross-sectional worldwide study. *Lancet Oncol* 11:1048–1056.
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. 2004. Classification of papillomaviruses. *Virology* 324:17–27.
- Dong X, Zhou W, Pfister H. 2000. Presence of genetic rearrangements in E1/E2 regions of episomal HPV 16 isolates from cervical carcinomas. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi* 14:5–8 [In Chinese].
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
- Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008 Int J Cancer 127:2893–2917.
- Ho L, Chan SY, Chow V, Chong T, Tay SK, Villa LL, Bernard HU. 1991. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol* 29:1765–1772.
- Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P. 1993. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 67:6413–6423.
- Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM. 2011. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer* 128:927–935.
- Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D. 2008. The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355–1361.
- Loo YM, Melendy T. 2004. Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *J Virol* 78:1605–1615.
- Masterson PJ, Stanley MA, Lewis AP, Romanos MA. 1998. A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit. *J Virol* 72:7407–7419.
- Muñoz N. 2000. Human papillomavirus and cancer: The epidemiological evidence. *J Clin Virol* 19:1–5.
- Parker LM, Harris S, Gossen M, Botchan MR. 2000. The bovine papillomavirus E2 transactivator is stimulated by the E1 initiator through the E2 activation domain. *Virology* 270:430–443.
- Pond SL, Frost SD. 2005. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533.
- Puranen M, Saarikoski S, Syrjänen K, Syrjänen S. 1996. Polymerase chain reaction amplification of human papillomavirus DNA from archival, Papanicolaou-stained cervical smears. *Acta Cytol* 40:391–395.
- Quint KD, de Koning MN, van Doorn LJ, Quint WG, Pirog EC. 2010. HPV genotyping and HPV16 variant analysis in glandular and squamous neoplastic lesions of the uterine cervix. *Gynecol Oncol* 117:297–301.
- Sabol I, Matovina M, Gasperov NM, Grce M. 2008. Identification of a novel human papillomavirus type 16 E1 gene variant with potentially reduced oncogenicity. *J Med Virol* 80:2134–2140.
- Sabol I, Matovina M, Si-Mohamed A, Grce M. 2012. Characterization and whole genome analysis of human papillomavirus type 16 e1-1374^{63nt} variants. *PLoS ONE* 7:e41045.
- Sanchez GI, Kleter B, Gheit T, van Doorn LJ, de Koning MN, de Sanjosé S, Alemany L, Bosch XF, Tommasino M, Muñoz N, Quint WG. 2011. Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *J Clin Virol* 51:165–1169.
- Sotlar K, Diemer D, Dethleffs A, Hack Y, Stubner A, Vollmer N, Menton S, Menton M, Dietz K, Wallwiener D, Kandolf R, Bültmann B. 2004. Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol* 42:3176–3184.
- Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. 2007. Selecton: Advanced models for detecting positive

- and purifying selection using a Bayesian inference approach. *Nucleic Acids Res* 35:W506–W511.
- Suzuki Y, Gojobori T. 1999. A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* 16:1315–1328.
- Swindle CS, Engler JA. 1998. Association of the human papillomavirus type 11 E1 protein with histone H1. *J Virol* 72:1994–2001.
- Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526.
- Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA A4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599.
- Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Gregg S, Buonaguro L, Buonaguro FM. 2011. Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 121:32–42.
- Tsakogiannis D, Ruether IGA, Kyriakopoulou Z, Pliaka V, Skordas V, Gartzonika C, Levidiotou-Stefanou S, Markoulatos P. 2012. Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece. *Arch Virol* 157:1729–1739.
- Tsakogiannis D, Papadopoulou A, Kontostathi G, Ruether IG, Kyriakopoulou Z, Dimitriou TG, Orfanoudakis G, Markoulatos P. 2013. Molecular and evolutionary analysis of HPV16 E6 and E7 genes in Greek women. *J Med Microbiol* 62:1688–1696.
- Tsakogiannis D, Darmis F, Gortsilas P, Ruether IGA, Kyriakopoulou Z, Dimitriou TG, Markoulatos P. 2014. Nucleotide polymorphisms of the human papillomavirus 16 E1 gene. *Arch Virol* 159:51–63.
- Vandepoele K, Van Roy N, Staes K, Speleman F, van Roy F. 2005. A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution. *Mol Biol Evol* 22:2265–2274.
- Wilson VG, West M, Woytek K, Rangasamy D. 2002. Papillomavirus E1 proteins: Form, function, and features. *Virus Genes* 24:275–290.
- Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA. 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743–7753.
- Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM. 1997. Human papillomavirus type 16 sequence variation in cervical cancers: A worldwide perspective. *J Virol* 71:2463–2472.

Molecular Biology Reports

Sites of disruption within E1 and E2 genes of HPV16 and association with cervical malignancy. --Manuscript Draft--

Manuscript Number:	MOLE-7574
Full Title:	Sites of disruption within E1 and E2 genes of HPV16 and association with cervical malignancy.
Article Type:	Manuscript
Keywords:	HPV16, E1 disruption, E2 disruption, HPV16 integration
Corresponding Author:	Panayotis Markoulatos GREECE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Dimitris Tsakogiannis
First Author Secondary Information:	
Order of Authors:	Dimitris Tsakogiannis Paschalis Gortsilas Zaharoula Kyriakopoulou Irina Georgia Anna Ruether Tilemachos Dimitriou Georges Orfanoudakis Panayotis Markoulatos
Order of Authors Secondary Information:	
Suggested Reviewers:	Michel Favre Pasteur Institute, paris mfavre@pasteur.fr Gaelle Boulet University of Antwerp gaelle.boulet@ua.ac.be Maria-Paz Canadas General Lab, Spain pcc@general-lab.es Antonio Carvajal-Rodriguez University of Vigo, Spain acraaj@uvigo.es Brian Moris University of Sydney, Australia brianm@medsci.usyd.edu.au Marialuisa Zerbini University of Bologna, Italy marialuisa.zerbini@unibo.it

Sites of disruption within E1 and E2 genes of HPV16 and association with cervical malignancy.

D. Tsakogiannis¹, P. Gortsilas¹, Z. Kyriakopoulou¹, I.G.A. Ruether¹, T.G. Dimitriou¹, G. Orfanoudakis², P. Markoulatos^{1*}.

- 1) University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology, Microbiology-Virology Laboratory, Larissa , Greece
- 2) University of Strasbourg, Ecole Supérieure de Biotechnologie de Strasbourg, CNRS FRE 3211-Oncoprotein Group, 67412 Illkirch, France

*Corresponding Author: Professor P. Markoulatos

University of Thessaly, School of Health Sciences

Department of Biochemistry & Biotechnology

Ploutonos 26 & Aiolou

41221 Larissa, Greece

Tel.: +302410 565274, Fax: +302410 565294

E-Mail: markoulatos@bio.uth.gr

Abstract.

1
2 In the present study disruption events within E1 and E2 genes of HPV16 genome
3 were studied in 31 high grade (CIN II, III), 40 low grade (CIN I) cervical
4 intraepithelial neoplasia cases and 10 cancer cases, in order to associate the disruption
5 of E1, E2 genes with cervical malignancy. Complete E2 and E1 genes were amplified
6 using three and nine overlapping primer sets respectively, in order to determine sites
7 of disruption. Disruption within E2 gene was strongly associated with high grade
8 malignancies (29%, $P < 0.001$), while in low grade malignancies only disruption
9 within E1 gene was involved (15%, $P < 0.05$). Specifically, deletions within E2 gene
10 occurs only in high grade dysplasia, while disruptions within E1 gene occurs both in
11 high and low grade dysplasia and make up disruption events in E1 gene a significant
12 early marker towards low grade malignancies. However for E1 gene the most
13 prevalent site of disruption was located between nucleotides 1059 and 1323, while
14 extensive mapping analysis of E2 gene revealed that the most prevalent deleted region
15 was located on the E2 hinge region (from 3172 to 3649). Extensive mapping analysis
16 of sites of disruptions in E1 and E2 genes of HPV16 genome constitute a precious
17 molecular tool for the investigation of the pure integrated HPV16 DNA form towards
18 the prediction of cervical malignancy development.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 **Key words:** HPV16, E1 disruption, E2 disruption, HPV16 integration
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

Introduction

Cervical cancer is the second most common cancer among women worldwide. Epidemiologic studies have shown that persistent infection with oncogenic Human Papillomaviruses (HPVs) is the major cause of invasive cervical cancer development [1-2]. Human papillomavirus type 16 (HPV16) is a small DNA virus with an icosahedral capsid comprising a double-stranded circular DNA genome of about 8000bp in size [3]. HPV16 is the most frequently identified HPV type addressed in 50% of cervical cancer cases [1,4-5]. Extensive study of viral genome revealed that HPV16 is classified in five major intratypic variant lineages known as HPV16 European, European Asian, African type II, African type I, and Asian American [6-8]. HPV16 variant lineages have been associated with the geographic origin of the populations in which are prevalent [9-10]. However, epidemiologic studies revealed that different HPV16 variants vary in pathogenesis and progression of cervical malignancy in cancer disease [11-12]. Moreover, a significant event in viral infection that occurs at an early stage of viral life cycle and plays a major role in the pathogenesis of cervical cancer is the integration of HPV16 genome into the host chromosome [13-19].

HPV16 genome can be found either as extrachromosomal circular episome or in integrated form. Experimental data indicate that the integrated HPV16 genome, co-exist with the HPV episomal form in the same cell [14-16, 19-27]. The integration of the virus usually disrupts the E1 and/or the E2 open reading frames (ORFs), leading to transcriptional activation of the viral oncogenes E6 and E7 [15-16, 28-30]. The transactivation domain of E2 protein interacts with the helicase domain of E1 protein. The E2 protein recruits additional E1 molecules, leading to the formation of double E1 hexameric helicase. The resulting E1 hexamer is an active DNA helicase that initiates the DNA replication of the episomal viral genome [31]. Since the disruption of the viral genome occurs in E1 and /or E2 genes, the integration leads to the interference of viral DNA replication control. In addition, since E2 protein regulates the transcription of the E6 and E7 oncogenes, disruption into the E2 ORF increases the immortalization capacity of the infected epithelial cell by augmenting the level of stable integrant E6 and E7 transcripts [32].

Previous studies investigated the physical state of HPV16 genome through mapping of disrupted sites into E1 and E2 ORFs [15-16, 33-34] and disruption of the HPV16 E2 gene was associated with high grade cervical malignancies [15-16, 35]. A

1 previous study by Cricca et al. [15] reported that disruption of both E1 and E2 genes
2 occurred in high grade cervical intraepithelial malignancies while in low grade cases
3 only E1 gene disruption was involved. Moreover, reported studies varied on
4 frequencies and sites of gene deletions fact that might be explained by differences in
5 technical methods used or by differences on the studied population [16]. In the present
6 analysis sites of disruption within HPV16 E1 and E2 genes were investigated through
7 PCR assay in thirty one high grade, forty low grade cervical intraepithelial neoplasia
8 cases and and ten cervical cancer cases derived from Greek women, in order to
9 associate the severity of cervical dysplasia with the sites of disruption of the viral
10 genome.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

2. Materials and Methods.

2.1 Cervical specimens

Eighty one HPV16-positive specimens of Greek women (eighteen cervical biopsies paraffin embedded samples and sixty three Thin Prep samples) were diagnosed as cervical intraepithelial neoplasia of stage I, II, III. Of the eighty one samples thirty one were diagnosed as high grade cervical intraepithelial neoplasia (CIN II, III) and forty samples as low grade cervical intraepithelial neoplasia (CINI). Moreover, ten HPV16-positive cervical cancer cases were included in the present analysis (nine cervical biopsies paraffin embedded tissues and one Thin Prep sample).

2.2 DNA extraction

DNA from Thin Prep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) [36]. DNA from formalin-fixed, paraffin-embedded cervical tissues was extracted following the proteinase K extraction method as has been previously described [35]. PCR amplification of a 498 bp fragment of human β -actin gene was performed in order to test for the presence of PCR inhibitors and evaluate the quality of the extracted DNA [20]. The cervical samples were screened for presence of HPV DNA by Nested Multiplex PCR [37].

2.3 Amplification of E1 gene

In order to map the sites of disruptions within E1 gene the complete E1 gene was amplified using the primer sets that were designed in a previous study by Arias-Pulido et al. [16] in separate reactions (table 1).

PCR reactions were performed in a final volume of 50 μ l. For the amplification of E1 gene each PCR mixture contained 50 pmol of each primer set, 10X PCR buffer (Stratagene, La Jolla, USA) containing 2 mM MgCl₂, 0.25 mM from each dNTP, and 2.5 U of thermostable DNA polymerase (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA). PCR cycling conditions were as follows: an initial denaturation step at 95°C for 2 min and then 40 cycles of: 95°C for 20 sec, 55°C for 30 sec and 72°C for 10 sec. PCR reaction was ended with a 5 min incubation step at 72°C.

2.4 Amplification of E2 gene

The complete E2 gene was amplified using three overlapping primer sets (a1-a2, b1-b2, c1-c2) in separate reactions, in order to locate sites of disruption within E2 gene [20] (table 1). Each PCR reaction was performed in a final volume of 50µl. For the amplification of E2 gene each PCR mixture contained 50 pmol of each primer set (a1-a2, b1-b2, c1-c2), 10X PCR buffer (Stratagene, La Jolla, USA) containing 2 mM MgCl₂, 0.25 mM from each dNTP, and 2.5 U of thermostable DNA polymerase (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA). PCR cycling conditions were as follows: an initial denaturation step at 95°C for 2 min and then 40 cycles of 95°C for 20 sec, 50°C for 30 sec and 72°C for 20 sec. PCR reaction was ended with a 5 min incubation step at 72°C. Ten microliters of each amplicon were electrophoresed in a 2% agarose gel containing 1 µg/ml ethidium bromide in Tris-borate-EDTA buffer. A 100-bp DNA ladder (Invitrogen Life Technologies, Paisley, UK) was used as molecular weight marker.

In order to avoid mismatches between the primer and the DNA target due to intratypic polymorphisms the primer sets were selected in order to hybridize in conserved regions among the reference sequence of HPV16 genome and those of representative sequences of HPV16 variants. In consequence absence of amplification signal was not related with reduced binding of primers but with disruption of E1 and or E2 genes.

2.4 Amplification of E6 gene.

Moreover in order to look at the integrity of HPV16 genome (namely from formalin-fixed, paraffin-embedded cervical tissues) we amplified a partial fragment of E6 gene (354bp in size) in all cervical samples. In particular, the E6 fragment was amplified using the primer set HPV-16 1 5'-GCACAGAGCTGCAAACAACT-3' / HPV-16 353 5'-GGTCCACCGACCCCTTATATT-3'. The primers were designed in the present analysis, through Primer3 program (<http://frodo.wi.mit.edu/primer3/>) based upon the sequence of the prototype HPV16 strain (GenBank ID: NC001526).

PCR was performed in a final volume of 50µl. PCR mixture contained 50 pmol of each primer, 10X PCR buffer (Stratagene, La Jolla, USA) containing 2 mM MgCl₂, 0.25 mM from each dNTP, and 2.5 U of thermostable DNA polymerase (Paq5000TM DNA Polymerase, Stratagene, La Jolla, USA). The cycling conditions were as follows: 40 cycles of 40s at 95°C, 30s at 58°C and 20sec at 72°C. The first

1
2 cycle was proceeded by a 2 min denaturation step at 95°C and the last cycle was
3 followed by a 5 min elongation step at 72°C.
4
5

6 **2.5 PCR sensitivity.**

7
8 In order to examine the sensitivity of PCR assays for E1, E2 and E6
9 amplification, a plasmid was constructed containing the region between E6 and E2
10 genes of HPV16 genome. The sequence from E6 to E2 gene was amplified using
11 Caski cell line. In particular, this viral portion was amplified though PCR assay using
12 the primer set: HPV-16 41 5'-AGCGACCCAGAAAGTTACCA-3' / HPV-16 3912
13 GCACACAAAGCAAAGCAAAA. The primers were designed in the present
14 analysis, through Primer3 program (<http://frodo.wi.mit.edu/primer3/>) as it was
15 described above.
16
17
18
19
20
21

22
23 PCR was performed in a final volume of 50µl. PCR mixture contained 50 pmol
24 of each primer, 10X Taq buffer (GeneON), 2mM MgCl₂, 1.2mM dNTPs and 2.5U of
25 thermostable DNA Polymerase (Maximo Taq DNA Polymerase, GeneON). The
26 cycling conditions were as follows: 40 cycles of 30s at 95°C, 40s at 55°C and 4min at
27 72°C. The first cycle was proceeded by a 5 min denaturation step at 95°C and the last
28 cycle was followed by a 10 min elongation step at 72°C. The amplicon of 3791bp in
29 size was subjected to cloning using the StataClone PCR cloning kit (Stratagene,
30 Agilent technologies). The recombinant plasmid DNA was purified using the
31 Nucleospin plasmid kit (Macherey-Nagel GmbH, Duren, Germany) and the plasmids
32 were subjected to sequencing at Macrogen Inc., Korea in order to confirm the
33 integrity of HPV16 sequence of pE6-E2. However, serial 10-fold dilutions (from 10⁶
34 to 1) of plasmid pE6-E2 were used in triplicates in order to assess the sensitivity of
35 the primer sets that were used to amplify the complete E1 and E2 genes and the partial
36 fragment of E6 gene.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 **2.6 Statistical analysis.** The relation among the sites of disruption in E1 and E2 genes
53 and the grade of cervical malignancy was determined through Fisher's exact test,
54 using GraphPad Prism version 4.00 for Windows, GraphPad Software, San Diego,
55 CA. P- values were considered significant if less than 0.05.
56
57
58
59
60
61
62
63
64

3. Results.

In the present study gene disruption events were defined only for those samples that gave no signal after PCR assays in agarose gel electrophoresis with one of the overlapping primer pairs for E1 and/or for E2 genes but were positive for β -actin and E6 genes. The sensitivity of primer sets that were used in the present analysis in order to amplify the complete E1, E2 genes and a partial sequence of E6 gene was determined to be 10 copies/ reaction for each primer set.

Amplification of E2 and E1 gene was carried out in the same cervical samples. Amplification of the complete E2 gene was performed in a total number of eighty one cervical samples. Disruption within E2 region was identified in a total of seventeen cervical samples (21%). In particular, nine out of the thirty one high grade cervical samples (29%), revealed disruption within E2 gene, while twenty two high grade cervical samples (71%) harbored the complete E2 gene. However, no evidence of E2 gene disruption was recorded in low grade cervical intraepithelial neoplasia cases. According to our results it was concluded that E2 gene disruption was strongly associated with high grade dysplasia, when compared with low grade cervical malignancies (Fisher's exact test, $P < 0.001$). Moreover, of the ten cancer cases E2 gene was found to be disrupted in eight cancer cases. However, considering results derived from mapping analysis of E2 gene it was reported that, disruption from nucleotide 2735 to 3210 was detected in one high grade cervical intraepithelial neoplasia case (table 2). Moreover, disruption from nucleotide 3127 to 3649 was reported in six high grade and seven cancer cases, while disruption from nucleotide 3597 to 3873 was detected in two high grade and one cancer case (table 2). Taking these data into account it was concluded that the viral region from nucleotide 3172 to 3649 was the most frequently disrupted region of E2 gene although disruption pattern implicated in different degree the entire E2 gene (table 2).

Amplification of the complete E1 gene was also performed in eighty one cervical samples. E1 gene disruption was found in twenty cervical samples (24.7%). The most frequently disrupted region was located in the 5'end of E1 gene from nucleotide 1059 to 1323 (table 2). In particular, eight out of thirty one high grade cervical samples (25.8%), showed E1 gene disruption, while intact E1 gene was identified in twenty three (74.2), high grade cervical specimens (table 2). On the other hand, thirty four out of the forty low grade cervical samples (85%), harbored the complete E1 gene, while six (15%), showed E1 gene disruption. Moreover, of the ten

1 cervical cancer cases, seven cervical specimens revealed disruption between
2 nucleotides 1059 and 1323 (table 2). Extensive mapping analysis of E1 and E2 genes
3 into the same cervical samples, revealed that in low grade specimens only E1 gene
4 was disrupted when compared with E1 and E2 gene sites disruption in low and high
5 grade cervical malignancies (Fisher's exact test, $P < 0.05$). On the other hand
6 extensive mapping analysis within E1 and E2 genes in the same cervical samples
7 revealed that a disruption/deletion genome disruption pattern involve both E1 and E2
8 genes in the 11.1% of the examined cervical samples (table 2). In particular,
9 disruption in both genes was detected in four high grade cervical samples (D2, D13,
10 D14, D15) and in five cancer cases (D58, D62, D64, D65, D66) (table 2).
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

Discussion

1 Disruption events within E1 and E2 genes have been strongly associated with
2 the integration of HPV16 DNA in the cellular chromosome [15-16, 34]. The main
3 objective of the present report was to investigate the sites of disruption within E1 and
4 E2 genes of HPV16 genome in high and low grade cervical intraepithelia neoplasia as
5 well as in cervical cancer cases, in order to associate the sites of disruption within
6 viral genome with the severity of cervical dysplasia. However, disruption events were
7 defined in cervical samples that gave no signal in PCR amplification with one of
8 twelve primer sets that were used for extensive mapping analysis of E1 and E2 genes,
9 but these same samples were positive for E6 gene, confirming the integrity of HPV16
10 genome. By the design of the present study only the integrated or episomal form of
11 the viral genome can be detected as the assay can not identify the mixed DNA forms
12 (integrated and episomal viral DNA). Eighty one cervical samples derived from Greek
13 women were examined and HPV16 genome disruption was identified in twenty nine
14 cervical cases (table 2).
15
16
17
18
19
20
21
22
23
24
25
26

27 Previous studies have reported different sites of disruptions within E1 gene,
28 while it has been demonstrated that the most prevalent site of disruption of E2 gene
29 occurs in the domain that encompass the hinge region of E2 protein [15-16, 34]. In a
30 previous study by Cricca et al. [15] a disruption pattern that implicated the 3' portion
31 of E1 gene and the E2 hinge region was reported, while Kalantari et al. [34] proposed
32 that gene disruption occurs in the 5' end of both E1 and E2 genes in cervical cancer
33 cases. In the present study E1 gene disruption was also detected in the 5' end of the E1
34 gene from nucleotide 1059 to 1323, targeted by primers E1F2-E1R2 in eighteen
35 (22.2%) out of the eighty one examined specimens, in two samples from nucleotides
36 878 to 1118 (2.5%), while in one specimen from nucleotide 1503 to 1784 (1,23%)
37 (table 2). On the other hand, disruption of the E2 gene was detected through out the
38 entire E2 gene (table 2). In particular, the hinge region targeted by primers b1-b2 from
39 nucleotides 3172 to 3649 was disrupted in thirteen (16.04%) out of the eighty one
40 examined samples, in three samples from nucleotides 3597 to 3873 (3.70%) and in
41 one sample from nucleotide 2735 to 3210 (1,23%). [15-16, 34]. Moreover a distinct
42 deletion/disruption pattern that involves both E1 and E2 genes was recorded in the
43 11.1% of the examined cervical samples, including four high grade (D2, D13, D14,
44 D15) and five cervical cancer cases (D58, D62, D64, D65, D66) (table 2). Disruption
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

in both E1 and E2 genes has been also previously reported in severe dysplasia and cancer disease [15-16].

In addition, our data indicate that E2 gene disruption was exclusively involved in high grade dysplasia in nine out of thirty one samples (29%) and in eight out of ten cervical cancer samples (80%) (Fisher's exact test, $P < 0.001$). No evidence of disruptions were recorded within E2 gene among the low grade malignancies. However, the extensive mapping analysis of E1 gene showed that in six out of the forty examined low grade samples carried out disruption within E1 gene (15%), while E1 gene disruptions were recorded in eight out of thirty one high grade samples (25.8%) and in seven out of ten cervical cancer samples (70%). Taking these data into account it was concluded that in low grade malignancies only the E1 gene disruption was involved (Fisher's exact test, $P < 0.05$), while in high grade malignancies and cervical cancer cases deletions in both E1 and E2 genes occurs [15]. Considering the frequency of E1 gene disruption among low grade, high grade and cervical cancer cases, the extensive mapping analysis of E1 gene might constitute a significant molecular marker for the identification of genome disruption in low grade malignancies through the progression of cervical intraepithelial neoplasia from stage I to superior stages. However, the presence of E2 gene disruption only in high grade cervical malignancies reinforced earlier studies that correlated the association of E2 gene disruption with high grade cervical malignancies and cervical cancer [15, 20-21, 38]. However, it has been proposed that the observed differences on frequencies and sites of genes disruptions that have been recorded by previous analyses may be due to differences of technical approaches that were used or in differences on the studied populations [16]. In this way extensive mapping analysis of E1 and E2 genes is a prerequisite in order to further evaluate the physical state of HPV16 DNA form in distinct populations.

In conclusion, extensive mapping analysis of E1 and E2 genes in cervical samples derived from Greek women revealed that in low grade cervical cases only the disruption of E1 gene was involved rendering the disruption on the 5'end of E1 gene a significant early marker for the identification of genome disruption in low grade malignancies towards the progression of cervical intraepithelial neoplasia. On the other hand, the association of E2 gene disruption with more severe dysplasia make up the E2 region a valid marker, especially the E2 hinge region, for the viral integration into the host chromosome and the progression towards high grade cervical

1 lesions. Mapping of sites of disruptions in E1 and E2 genes of HPV16 genome
2 constitute a precious molecular tool for the investigation of the integrated or episomal
3 HPV16 DNA form towards the prediction of cervical malignancy development.
4 Finally, according to the hypothesis that frequencies and sites of disruptions may be
5 associated with differences on the studied populations, extensive mapping analysis of
6 E1 and E2 genes should be reinforced in order to detect in time course HPV16
7 genome integration events in different populations. All HPV-16 positive women
8 should be screened regularly for E1 and E2 sites of disruptions and monitored closely,
9 facts that will be significant for early diagnosis, prevention and finally treatment.
10
11
12
13
14
15
16
17
18
19

20 **Acknowledgements**

21
22 This work was supported by research grants of the Postgraduate Programmes
23 ‘Applications of Molecular Biology-Genetics. Diagnostic Biomarkers’, code 3817 and
24 “Biotechnology”, code 3439, of the University of Thessaly, School of Health
25 Sciences, Department of Biochemistry & Biotechnology.
26
27
28
29
30

31 All authors declare that they have no conflicting or dual interests.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127:2893-2917
2. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, Snijders PJ, Meijer CJ (2003) International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 348:518-527
3. Zur Hausen H (1996) Papillomavirus infections – a major cause of human cancers. *Biochim Biophys Acta* 1288:F55-78
4. Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM (2011) Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer* 128:927-935
5. De Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, Tous S, Felix A, Bravo LE, Shin HR, Vallejos CS, de Ruiz PA, Lima MA, Guimera N et al (2010) Retrospective International Survey and HPV Time Trends Study Group. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 11:1048-1056
6. Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD (2005) Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 79:7014-7023
7. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P (1993) The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 67:6413-6423
8. Ho L, Chan SY, Chow V, Chong T, Tay SK, Villa LL, Bernard HU (1991) Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol* 29:1765-1772
9. Chan SY, Bernard HU, Ong CK, Chan SP, Hofmann B, Delius H (1992) Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses. *J Virol* 66:5714-572
10. Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* 69:7743-7753
11. Sanchez GI, Kleter B, Gheit T, van Doorn LJ, de Koning MN, de Sanjosé S, Alemany L, Bosch XF, Tommasino M, Muñoz N, Quint WG (2011) Clinical evaluation of polymerase chain reaction reverse hybridization assay for detection and identification of human papillomavirus type 16 variants. *J Clin Virol* 51:165-169
12. Tornesello ML, Losito S, Benincasa G, Fulciniti F, Botti G, Greggi S, Buonaguro L, Buonaguro FM (2011) Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. *Gynecol Oncol* 121:32-42
13. Cañadas MP, Videla S, Darwich L, Tarrats A, Piñol M, García-Cuyás F, Llatjos M, Alcalde C, Fernández I, Sirera G, Clotet B (2010) Human papillomavirus HPV-16,

- 18, 52 and 58 integration in cervical cells of HIV-1-infected women. *J Clin Virol* 48:198-201
14. Huang LW, Choa SL, Lee BH (2008) Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. *J Clin Pathol* 61:627-631
15. Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbini M (2009) Disruption of HPV16 E1 and E2 genes in precancerous cervical lesions. *J Virol Methods* 158:180-183
16. Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM (2006) Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 44:1755-1762
17. Kulmala SMA, Syrjanen SM, Gyllensten UB, Shabalova IP, Petrovichev N, Tosi PT, Syrjanen KJ, Johansson BC (2007) Early integration of high copy HPV detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 2007; 59:513-517
18. Doorbar J (2007) Papillomavirus life cycle organization and biomarker selection. *Dis Markers* 23:297-313
19. Briolat J, Dalstein V, Saunier M et al. (2007) HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. *Int J Cancer* 121:2198–2204
20. Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D (2008) The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 134:1355-1361
21. Tonon SA, Picconi MA, Bos PD, Zinovich JB, Galuppo J, Alonio LV, Teyssie A (2001) Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J Clin Virol* 21:129-134
22. De Marco L, Gillio-Tos A, Bonello L, Ghisetti V, Ronco G, Merletti F (2007) Detection of human papillomavirus type 16 integration in preneoplastic cervical lesions and confirmation by DIPS-PCR and sequencing. *J Clin Virol* 38:7–13
23. Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B (2005) Type distribution, viral load and integration state of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer* 92:2195-2200
24. Hudelist G, Manavi M, Pischinger KID, Watkins-Riedel T, Singer CF, Kubista E, and Czerwenka KF (2004) Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol Oncol* 92:873–880
25. Pett MR, Herdman MT, Palmer RD, Yeo GSH, Shivji MK, Stanley MA, Coleman N (2006) Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *Proc Natl Acad Sci USA* 103:3822–3827
26. Matovina M, Sabol I, Grubisić G, Gasperov NM, Grce M (2009) Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecol Oncol* 113:120-127
27. Cheung JL, Lo KW, Cheung TH, Tang JW, Chan PK (2006) Viral load, E2 gene disruption status, and lineage of human papillomavirus type 16 infection in cervical neoplasia. *J Infect Dis* 194:1706-1712
28. Hamid NA, Brown C, Gaston K (2009) The regulation of cell proliferation by the papillomavirus early proteins. *Cell Mol Life Sci* 66:1700-1717
29. Kadaja M, Silla T, Ustav E, Ustav M (2009) Papillomavirus DNA replication – from initiation to genomic instability. *Virology* 384:360-368

- 1 30. Thiery F (2009) Transcription regulation of the papillomavirus oncogenes by
2 cellular and viral transcription factors in cervical carcinoma. *Virology* 384:375-379
- 3 31. D'Abramo CM, Archambault J (2011) Small molecule inhibitors of human
4 papillomavirus protein - protein interactions. *Open Virol J* 5:80-95
- 5 32. Boulet G, Horvath C, Vanden Broeck D, Sahebali S, Bogers J (2007) Human
6 papillomavirus: E6 and E7 oncogenes. *Int J Biochem Cell Biol* 39:2006-2011
- 7 33. Lukaszuk K, Liss J, Wozniak I, Emerich J, Wójcikowski C (2003) Human
8 papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex
9 PCR. *J Clin Microbiol* 41:608-612
- 10 34. Kalantari M, Karlsen F, Kristensen G, Holm R, Hagmar B, Johansson B (1998)
11 Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is
12 associated with poor prognosis. *Int J Gynecol Pathol* 17:146-153
- 13 35. Tsakogiannis D, Ruether IG, Kyriakopoulou Z, Pliaka V, Theoharopoulou A,
14 Skordas V, Panotopoulou E, Nepka C, Markoulatos P (2012) Sequence variation
15 analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from
16 women in Greece. *Arch Virol* 157:825-832
- 17 36. Casas I, Powell L, Klapper PE, Cleator GM (1995) New method for the extraction
18 of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain
19 reaction assay. *J Virol Methods* 53:25-36
- 20 37. Sotlar K, Diemer D, Dethleffs A, Hack Y, Stubner A, Vollmer N, Menton S,
21 Menton M, Dietz K, Wallwiener D, Kandolf R, Bültmann B (2004) Detection and
22 typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol*
23 42:3176-3184
- 24 38. Cricca M, Morselli-Labate AM, Venturoli S, Ambretti S, Gentilomi GA, Gallinella
25 G, Costa S, Musiani M, Zerbini M (2007) Viral DNA load, physical status and
26 E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical
27 lesions. *Gynecol Oncol* 106:549-557
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64

Primers	5'-3' sequence	location	PCR product	
E2 gene				
a1	AGGACGAGGACAAGGAAAA	2735	475 bp	95°C 2 min 95°C 20 sec 50°C 30 sec 72°C 5 min } 40 cycles
a2	ACTTGACCCTCTACCACAGTTACT	3210	477 bp	
b1	TTGTGAAGAAGCATCAGTAACT	3172	477 bp	
b2	TAAAGTATTAGCATCACCTT	3649	477 bp	
c1	GTAATAGTAACACTACACCCATA	3597	276 bp	
c2	GGATGCAGTATCAAGATTTGTT	3873	276 bp	
E1 gene				
E1F1	CAGGTACCAATGGGGAAGA	878	241bp	95°C 2 min 95°C 20 sec 55°C 30 sec } 40 cycles 72°C 10 sec 72°C 5 min
E1R1	ACCTGTACTGCATCTCTATGTTGT	1118	241bp	
E1F2	GACAGCACATGCGTTGTTTAC	1059	265bp	
E1R2	AGTCTCATGGCGCCCTTC	1323	265bp	
E1F3	CGCCATGAGACTGAAACAC	1312	212bp	
E1R3	GAAAAACTCACCCGTATAAC	1523	212bp	
E1F4	GTTATACGGGGTGAGTTTTTC	1503	282bp	
E1R4	ATCATCATAACATTGGAGACA	1784	282bp	
E1F5	GTGTGTCTCCAATGTGTATGATG	1760	232bp	
E1R5	GCAATTTCACTATCGTCTACTATG	1991	232bp	
E1F6	GGGCCTACGATAATGACATAGTA	1952	235bp	
E1R6	TGCTTCCAATCACCTCCAT	2186	235bp	
E1F7	GGAGGTGATTGGAAGCAA	2170	316bp	
E1R7	CCAATGCATTTCTTAAATTGTC	2485	316bp	
E1F8	CCCTGTTGGAACACTACATAGATGAC	2443	301bp	
E1R8	CCTCGTCCTCGTGCAAACCT	2743	301bp	
E1F9	CCAGTGTATGAGCTTAATGATAAGA	2662	187bp	
E2R9	GTCTATATGGTCACGTAGGTCTGTA	2848	187bp	

table1: Primers used for E1 and E2 gene amplification in order to map the sites of disruption within E1 and E2 genes.

				E1 gene						E2 gene					
		Position		E1F1-E1R1	E1F2-E1R2	E1F3-E1R3	E1F4-E1R4	E1F5-E1R5	E1F6-E1R6	E1F7-E1R7	E1F8-E1R8	E1F9-E1R9	a1-a2	b1-b2	c1-c2
		878-1118	1059-1323	1312-1523	1503-1784	1760-1991	1952-2186	2170-2485	2443-2743	2662-2848	2735-3210	3172-3649	3597-3873		
1															
2															
3															
4															
5															
6	CIN	source	Position												
7	LG	paraffin	D7	-											
8	LG	ThinPrep	747		-										
9	LG	ThinPrep	ATT6				-								
10	LG	ThinPrep	ATT78		-										
11	LG	ThinPrep	ATT91		-										
12	LG	ThinPrep	ATT98		-										
13	CIN														
14	HG	paraffin	D2		-								-		
15	HG	paraffin	D4								-				
16	HG	paraffin	D13		-								-		
17	HG	paraffin	D14		-								-		
18	HG	paraffin	D15		-										-
19	HG	paraffin	D30												-
20	HG	ThinPrep	ATT9		-										
21	HG	ThinPrep	ATT21										-		
22	HG	ThinPrep	ATT24										-		
23	HG	ThinPrep	PNL9		-										
24	HG	ThinPrep	PNL1		-										
25	HG	ThinPrep	PNL12	-											
26	HG	ThinPrep	ATT79										-		
27	HG	ThinPrep	ATT79										-		
28	HG	ThinPrep	ATT79										-		
29	cancer	ThinPrep	PNL7										-		
30	cancer	paraffin	D58		-								-		
31	cancer	paraffin	D59												-
32	cancer	paraffin	D60		-										
33	cancer	paraffin	D61		-										
34	cancer	paraffin	D62		-								-		
35	cancer	paraffin	D63										-		
36	cancer	paraffin	D64		-								-		
37	cancer	paraffin	D65		-								-		
38	cancer	paraffin	D65		-								-		
39	cancer	paraffin	D66		-								-		
40	cancer	paraffin	D66		-								-		

table 2: Sites of disruption placed within E1 and E2 genes according to overlapping primer pairs in high, low grade and cervical cancer samples.

█ Indicates the presence of nucleotide fragments

— Indicates sites of disruption

Determination of HPV16 physical status through E1/E6 and E2/E6 ratio analysis

D. Tsakogiannis¹, V. Diamantidou¹, G.D. Amoutzias¹, Z. Kyriakopoulou¹, I.G.A. Ruether¹, T.G. Dimitriou¹, C. Kotsovassilis², P. Markoulatos¹

- 1) University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology, Microbiology-Virology Laboratory, Larissa , Greece
- 2) General Hospital of Athens, Clinical Biochemistry Dpt, Athens, Greece

*Corresponding Author: Professor P. Markoulatos

University of Thessaly, School of Health Sciences

Department of Biochemistry & Biotechnology

Ploutonos 26 & Aiolou

41221 Larissa, Greece

Tel.: +302410 565274, Fax: +302410 565294

E-Mail: markoulatos@bio.uth.gr

Abstract

HPV16 genome integration into the host chromosome is a crucial event during viral life cycle and a major step towards carcinogenesis. The integration of HPV16 DNA promotes a constitutive high expression level of E6 and E7 oncoproteins, resulting to the extensive proliferation of the infected epithelial cells. The present report studied the physical status of HPV16 genome, through determination of E1/E6 and E2/E6 DNA copy number ratios in sixty one cervical samples of low and high grade malignancy and eight cervical cancer samples, all of them associated with HPV16 infection. The selection of E1, E2 and E6 amplification target regions was performed according to the most prevalent deleted/disrupted sites of E1 and E2 genes. For this target selection we also considered the most conserved regions of E1, E2 and E6 genes among the same HPV16 isolates that were recently reported by our group. The analysis of HPV16 DNA form revealed a significant association among the mixed DNA forms in low grade and high grade malignancies, (χ^2 , $P < 0.01$). The comparative analysis of E1/E6 and E2/E6 in the same cervical samples provides an accurate picture of HPV16 DNA form and may reveal whether different HPV16 DNA integrants coexist in the same cervical sample or not. This study proposes that E1/E6 and E2/E6 ratios determine with accuracy the HPV16 DNA integration pattern and may predict multiple integration events in the examined sample, thus providing significant information about the progression of cervical dysplasia.

Key words: HPV16 integration, E1 gene, E2 gene, E1/E6, E2/E6, HPV16 physical status

Introduction.

Cervical cancer is the third most common type of cancer among women worldwide, with a high mortality rate. The worldwide incidence of cervical carcinoma is more than 530,000 cases per year, whereas mortality reaches 275,000 deaths annually, of which approximately 85% occur in developing countries¹⁻². The etiologic agent for the development of high grade precancerous cervical lesions and invasive cervical cancer is the persistent infection with oncogenic Human Papillomavirus (HPV) types³. To date, more than 150 different HPV types have been characterized and about 40 of them are related with anogenital tract malignancy, grouped as high-risk (HR) or low risk (LR) genotypes⁴⁻⁶. Epidemiological studies revealed that HPV16 is the most commonly observed high-risk HPV type followed by HPV18, 31, 33 and 45⁷⁻⁸.

Persistent infection with high-risk HPV types is associated with an increasing risk of integration of viral circular genome (episome) into the host chromosomes, leading to cancer development. The circular HPV genome is then linearized but the Long Control Region (LCR) and the E6, E7 oncogenes are always retained intact⁹⁻¹¹. Viral integration appears to coincide with the development of high grade cervical intraepithelial neoplasia (CIN II, III) as a consequence of over expression of the E6 and E7 oncogenes¹². In particular, the integration of HPV16 genome usually disrupts the E2 and/or E1 regions found downstream of the early genes E6 and E7, thus resulting to functional inactivation of E1 and E2 proteins¹³⁻¹⁶. Disruption events within the E1 and /or E2 genes lead to interference of viral DNA replication control. Moreover, disruption of the E2 gene results in the elimination of the E2 viral repressor function that leads to deregulation of the E6/E7 promoter activity^{13,15-19}. In addition, integration of the HPV genome leads to the generation of viral - cellular fusion transcripts with increased stability that results to constitutive, high expression of the viral E6, E7 oncogenes²⁰. As a consequence, the integrated viral DNA stimulates cell cycle entry and provides a selective growth advantage to the infected cells. The extensive proliferation of infected epithelia is a key event for the accumulation of genetic errors and genome destabilization that finally results to cancer development²¹⁻²².

The present report focused on the determination of the physical status of HPV16 genome among low and high grade malignancies and cervical cancer cases.

Towards this end we took into consideration previous reports that the most frequently disrupted region of HPV16 genome during integration is the 3' end of viral DNA, between E1 and E2 genes. The ratio values of E1/E6 and E2/E6 DNA copy numbers for each individual cervical sample were determined in order to discern the precise physical state of HPV16 genome and the HPV16 integration pattern in host chromosomes. The overall goal was achieved through Real time PCR assays, using specific primer pairs hybridizing to E1 and E2 target regions and with an E6 specific primer pair hybridizing to the E6 ORF. The E1 and E2 target regions were previously determined to be the most frequently disrupted sites during HPV16 DNA integration, whereas the E6 region is always retained intact²³⁻²⁵.

2. Materials and Methods.

2.1 Cervical sample collection and DNA isolation

Sixty nine samples of Greek women positive for HPV16 DNA (55 from Thin Prep and 14 from cervical biopsies of paraffin embedded tissue), were diagnosed as cervical intraepithelial neoplasias of stage I, II, III and cervical cancer. Of these sixty nine samples, twenty seven were diagnosed as high grade cervical intraepithelial neoplasia (CIN II, III), thirty four as low grade cervical intraepithelial neoplasia (CIN I) and eight as cervical cancer cases.

Genomic DNA from Thin Prep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN)²⁶, while DNA from formalin-fixed, paraffin-embedded cervical tissues was extracted using the established proteinase K method²⁷. A 250bp GAPDH fragment was amplified by PCR assay as a quality control in order to assay for the integrity of the extracted DNA and for inhibition of PCR process²⁸. Identification of HPV16 genome was carried out through Nested Multiplex PCR as previously described²⁹.

2.2 Plasmids construction

Real time-PCR (RT-PCR) assays were conducted in order to determine the E2/E6 and E1/E6 DNA copy number ratios and thus understand the physical status of HPV16 DNA in all the investigated cervical samples. A plasmid containing the viral region from E6 to E2 genes (pE6-E2) was constructed in order to normalize the Real

Time - PCR assays, while a plasmid containing a partial sequence of GAPDH gene (pGAPDH) was constructed, in order to normalize for sample copy number by means of GAPDH quantification. Moreover, an additional E6 plasmid containing a partial sequence of the E6 ORF (pE6) was designed to perform reconstitution experiments for the determination of the ratio cut-off value.

2.2.1 pE6-E2

The standard curves for E1, E2 and E6 genes were constructed through plasmid pE6-E2, which contains the region from E6 to E2 gene. This viral region was amplified through PCR assay using the primer set: HPV-16 41 5'-AGCGACCCAGAAAGTTACCA-3' /HPV-16 3912 GCACACAAAGCAAAGCAAAA. The primers were designed with the Primer3 program (<http://bioinfo.ut.ee/primer3-0.4.0/>). The numbering of nucleotide positions was based upon the sequence of the prototype HPV16 strain (GenBank ID: NC001526).

PCR was performed in a final volume of 50µl. PCR mixture contained 50 pmol of each primer, 10X Taq buffer (GeneON), 2mM MgCl₂, 1.2mM dNTPs and 2.5U of thermostable DNA Polymerase (Maximo Taq DNA Polymerase, GeneON). The cycling conditions were as follows: 40 cycles of 30s at 95°C, 40s at 55°C and 4min at 72°C. The first cycle was preceded by a 5 min denaturation step at 95°C and the last cycle was followed by a 10 min elongation step at 72°C.

The amplicon of 3791bp in size was subjected to cloning using the StataClone PCR cloning kit (Stratagene, Agilent technologies). The recombinant plasmid DNA was purified using the Nucleospin plasmid kit (Macherey-Nagel GmbH, Duren, Germany) and the plasmids were subjected to sequencing at MacroGen Inc., Korea in order to confirm the integrity of HPV16 sequence of pE6-E2.

2.2.2 pGAPDH

The GAPDH plasmid was constructed through PCR amplification of a 250bp fragment in size. The PCR amplification was carried out using the primer set GAPDH F 5'-ACGGATTTGGTCGTATTGGG-3' / GAPDH R 5'-TGATTTTGGAGGGATCTCGC-3' as it was previously reported²⁸. The GAPDH plasmid construction was similar to that described for the construction of pE6-E2.

2.2.3 pE6

The E6 plasmid was constructed through PCR amplification of a 354bp fragment located in the E6 ORF in order to perform the reconstitution experiments. In particular, the E6 fragment was amplified using the primer set HPV-16 1 5'-GCACAGAGCTGCAAACAAC-3' / HPV-16 353 5'-GGTCCACCGACCCCTTATATT-3'. The primers were designed with the Primer3 program and the numbering of nucleotide positions was based upon the sequence of the prototype HPV16 strain (GenBank ID: NC001526).

PCR was performed in a final volume of 50µl. PCR mixture contained 50 pmol of each primer, 10X Taq buffer (GeneON), 1.5mM MgCl₂, 1mM dNTPs and 2.5U of thermostable DNA Polymerase (Maximo Taq DNA Polymerase, GeneON). The cycling conditions were as follows: 40 cycles of 40s at 95°C, 30s at 60°C and 20sec at 72°C. The first cycle was preceded by a 2 min denaturation step at 95°C and the last cycle was followed by a 5 min elongation step at 72°C. The amplicon of 354bp in size was subjected to cloning, as described above.

2.3. Quantitative Real time PCR assays

A total of sixty nine cervical samples were investigated in order to determine the integrated, mixed or episomal physical status of HPV16 genome. Real Time PCR assays were performed using the Mx3000P® instrument (Stratagene). To determine the physical state of viral DNA, the E1, E2 and E6 genes were quantified and both E1/E6 and E2/E6 ratios were calculated for each individual cervical sample. The principle of this approach is based on the assumption that E1, E2 and E6 genes are present in equal amounts in episomal HPV16 genome, while integration most often induces the deletion of the E2 and/or E1 target regions but not of E6.

Extensive mapping analysis of the E2 gene in the sixty nine cervical samples was conducted using three overlapping primer sets (a1-a2, b1-b2, c1-c2) through PCR assay, in order to locate sites of disruption into the E2 gene, as previously described²⁸. However, in order to map the sites of disruptions within E1 gene the complete E1 gene was amplified using the primer sets that were designed in a previous study in separate reactions¹³. In the present study the most commonly disrupted region of E2 ORF was that encoding for the hinge region of E2 protein (from 3172 to 3649), while for E1 gene the most prevalent region of disruption was located within the 5'end of E1

gene (from 1059 to 1323) (data not shown). The selection of primer pairs was conducted according to the following two criteria: the primer pairs must target the most frequently disrupted regions of E1 and E2 genes and must hybridize to conserved regions of E1, E2 and E6 genes. As a consequence, the amplification reaction for E2 target was carried out using the primer pair E2F12-E2R12 that targets the hinge region between nucleotides 3243 and 3539. For E1 target the primer pair E1F2-E1R2 was used, that targets the region between nucleotides 1059 and 1323. Moreover the amplification reaction for E6 target was performed using the primer pair HPV16 1- HPV16 353.

The construction of standard curves for E1, E2, E6 and GAPDH genes was carried out for each individual gene - target separately. Each Real Time PCR assay was performed in a final volume of 25 μ l. Real Time PCR mixture contained 10 pmol of each primer set, 2X KAPA SYBR[®] FAST qPCR Master Mix (KAPA SYBR[®] FAST qPCR Kit, Boston, Massachusetts, United States) and ROX Reference Dye. The cycling conditions for E1, E2, E6 and GAPDH genes were as follows: 40 cycles of 3s at 95 $^{\circ}$ C and 30s at 60 $^{\circ}$ C. The first cycle was preceded by a 1min denaturation step at 95 $^{\circ}$ C. Data acquisition at 510 nm was performed at 72 $^{\circ}$ C (table 1)

One standard curve for GAPDH was constructed through amplification of 10-fold dilution series of 10⁸ to 1 copy of pGAPDH. In addition, three distinct standard curves were constructed for E1, E2 and E6 targets using only the pE6-E2 plasmid construct that contains the viral fragment from E6 to E2 ORFs through primer sets E1F2-E1R2, E2F12-E2R12 and HPV16 1/ HPV16 353, respectively. The three standard curves were constructed through amplification of 10-fold dilution series ranging between 10⁸ to 1 copy of pE6-E2 per reaction. All standard curves were performed in triplicates in three independent Real Time PCR assays, in order to minimize unreliability of the quantification of results. All cervical samples were examined in triplicates. The copy number of E1, E2 and E6 genes was determined according to the number of cells (Real time PCR targeting GAPDH gene) and results were expressed as E1, E2, E6 copy number per 500 cells.

2.4 Cut- off value

An experimental cut-off value was established in order to determine the physical state of HPV16 DNA (episomal, integrated, mixed integrated and episomal

forms). The ratio cut-off was obtained by preparing DNA plasmid solutions containing a standard quantity of 100 copies of the pE6-E2 plasmid and different copy numbers of pE6 ranging from 0 to 1×10^5 copies (table 2). The plasmid solution that contained only the pE6-E2 was considered to be the episomal DNA, while the solution containing only the pE6 plasmid was considered to be the integrated DNA. However E1, E2 and E6 copy numbers were determined in all plasmid solutions in triplicate in three independent real time PCR assays and the E1/E6, E2/E6 ratios were calculated for each DNA solution (table 2).

2.5 Statistical analysis

The association among the physical status of HPV16 genome and the grade of cervical malignancy was determined through Chi-squared test. The statistical analysis was conducted with the GraphPad Prism version 4.00 software for Windows. P-values were considered significant if less than 0.05.

3. Results

In the present study a total of sixty nine cervical samples which were positive for HPV16 DNA were investigated in order to determine the physical state of HPV16 genome through E2/E6 and E1/E6 ratio analysis. An extensive mapping analysis of E1 and E2 genes was carried out in all cervical samples through PCR assay, using twelve overlapping primer sets^{13, 28}. According to our findings the most frequently disrupted site was located in the E2 hinge region (from nucleotides 3172 to 3649), while the most prevalent site of disruption within E1 gene was located between nucleotides 1059 and 1323. Moreover nucleotide analysis of the E1 and E2 genes in the same cervical samples that were derived from Greek women revealed that the region between nucleotides 3172 and 3649 (E2 gene) and the region between nucleotides 1059 and 1323 (E1 gene) is conserved among distinct HPV16 isolates that are circulating in the Greek population^{27,30-31}. Taking these data into consideration, we targeted the E2 domain between nucleotides 3243 and 3539 and the E1 domain between nucleotides 1059 and 1323, through Real Time PCR, in order to assign the copy numbers of E2 and E1 genes respectively.

Reconstitution experiments were performed using plasmid solutions containing from 0 to 1×10^5 copies of pE6 plasmid (integrated form of HPV16 DNA)

combined with 100 copies of pE2-E6 plasmid (episomal form of HPV16 genome) in order to determine the cut-off value and establish the physical state of HPV16 genome (table 2). The E1, E2 and E6 gene copy numbers were determined through quantitative Real Time PCR assay, using the standard curves that were constructed for this purpose. According to our findings the ratio values close to zero (E2/E6 ratio value < 0.001 , E1/E6 ratio value < 0.03) were observed when pE6 plasmid was in 100-fold excess of pE2-E6 plasmid. E2/E6 ratio value of 0.01 and the E1/E6 ratio value of 0.03 were observed in plasmid solutions that contained 10 fold excess of pE6/ pE2-E6 plasmid. The E2/E6 ratio value of 0.62 and the E1/E6 ratio value of 0.52 were observed in plasmid solutions that contained equal amounts of pE6-E2 and pE6 plasmids (100 copies from each). However when pE6-E2 plasmid was in 10-fold excess in plasmid solution the E2/E6 ratio value was 0.82, and the E1/E6 value was 0.7. Finally, the E2/E6 ratio value of 0.96 and the E1/E6 ratio value of 0.98 was encountered in plasmid solutions that contained only the pE6-E2 plasmid (episomal DNA) (table 2).

Taking these data into account it was concluded that E2/E6 ratio gave evidence for viral integration when its value was less than 0.001. However, E2/E6 ratios less than 0.96 (range, 0.001 to 0.96) indicated the presence of both integrated and episomal forms of viral DNA (mixed DNA forms), while ratio values equal to or greater than 0.96 indicated the presence of the episomal form of HPV16 genome (table 2). Similarly, E1/E6 ratio analysis revealed that viral integration occurs when the E1/E6 ratio value was less than 0.003. E1/E6 ratios ranging from 0.003 to 0.98 indicated the presence of mixed DNA forms, while ratio values greater or equal to 0.98 showed the presence of the episomal form of the viral genome (table 2).

In the present study E1/E6 and E2/E6 ratio analysis was performed in the same cervical specimens in order to identify the HPV16 DNA form. E1/E6 and E2/E6 ratio values analysis in the sixty nine cervical samples, showed that the pure episomal DNA form was present in six out of thirty four (6/34 - 17.6%) low grade cervical intraepithelial cases (table 3). However, no evidence of pure episomal DNA form was recorded in high grade malignancies and cervical cancer specimens. The mixed DNA form was identified in twenty seven out of thirty four (27/34 - 79.5%) low grade cases and in twenty four out of twenty seven (24/27 - 88.8%) high grade cervical intraepithelial neoplasia cases (table 3). Moreover, the mixed DNA form was

identified in two out of eight (2/8 - 25%) cervical cancer specimens. In addition, the pure integrated viral form was identified in three out of twenty seven (3/27 - 11.1%) high grade cases and in six out of eight (6/8 - 75%) cervical cancer samples (table 3). Finally, the pure integrated viral form was reported in the low grade sample 747 (2.9%) (table 4). Taking these data into account, it was concluded that the pure episomal DNA form is exclusively detected in low grade malignancies (17.4%), while the pure integrated viral form is preferentially encountered in cervical cancer cases (75%). However, a significant association among the high grade cervical intraepithelial neoplasias and the mixed DNA form was recorded, when compared with the episomal (0%) and the integrated viral form (11.1%) ($\chi^2 = 16.333$, $P < 0.0001$). In addition, a strong association among the low grade cases and the mixed DNA form was identified, when compared with the episomal (17.6%) and the integrated viral form (2.9%) ($\chi^2 = 11.7$, $P = 0.0006$).

Nevertheless, in seven low grade malignancies an important difference among E1/E6 and E2/E6 ratio values was recorded. In particular, in five low grade cervical cases (PNL3, ATT8, ATT18, ATT12, ATT46) the E2/E6 ratio values analysis showed the presence of mixed HPV16 DNA form, while the E1/E6 ratio values analysis revealed the presence of episomal HPV16 DNA form (table 4). In contrast, in two low grade cervical cases (PNL2, ATT54) the E2/E6 ratio analysis showed the episomal viral DNA form, while the E1/E6 ratio analysis revealed the mixed form of viral DNA (table 4). According to these results it was concluded that the mixed HPV16 DNA form is encountered in these cervical cases, suggesting that the integration of HPV16 DNA may occur through disruption/deletion of E2 and E1 genes. In consequence, we propose that both the E1/E6 and E2/E6 ratio analyses should be considered in order to identify the accurate viral form mainly in low grade malignancies, as HPV16 integration events also occur through disruption of E1 gene.

Discussion

HPV integration into the host chromosome is the key event for cervical cancer development and has been proposed to occur early in the development of cervical malignancy^{13,25,28,32-33}. In the present study a total of sixty one precancerous cervical samples and eight cervical cancer cases associated with HPV16 infection were investigated in order to study the physical status of HPV16 genome through quantitative Real Time PCR and E2/E6, E1/E6 ratio values analysis. However, in order to achieve more reliable results the selection of E1, E2 and E6 targets was based on the most frequently disrupted sites of E1 and E2 genes and according to the most conserved E6 regions among the same HPV16 isolates, that were recently reported by our group^{27, 30-31}.

Several reports have investigated the HPV16 genome status, through E2/E6 ratio analysis and demonstrated that this ratio is a significant molecular marker for evaluating the risk of progression of cervical intraepithelial neoplasia of stage I, II, III to cervical cancer^{25,28,32,34,35,37}. As a result, the HPV16 integration was determined only through disruption of E2 gene, while the integrity of E1 gene was not investigated. In addition, it has been demonstrated that disruption of both E1 and E2 genes is involved in high grade cervical intraepithelial neoplasia cases due to HPV16 DNA integration, while in low grade cases only the E1 gene is involved¹⁵. Considering the prevalence of E1 gene disruption in low grade cases, a comparative analysis along with E1/E6 and E2/E6 ratio values on the same cervical samples was performed in order to accurately determine the HPV16 DNA form and to explore for additional integration events that have been induced through disruption of E1 gene.

Previous analyses have reported that the integration of viral genome is exclusively identified in high grade lesions and cervical cancer cases^{32,39}, while other studies found early integration events in low grade malignancies^{24-25,38,40}. Our findings demonstrate that the pure integrated viral form is preferentially identified in high grade malignancies and cervical cancer cases, while the pure episomal DNA form is exclusively identified in low grade lesions. In addition, it was demonstrated that both high grade and low grade malignancies present significant relation with the mixed DNA viral form (χ^2 , $P < 0.001$)³⁸. The high prevalence of HPV16 integration as mixed DNA form in low grade dysplasia reinforces earlier studies who proposed that HPV16 integration occurs at the early stage of viral life

cycle^{24-25,38,40}. In addition, in one low grade sample the pure integrated viral form was identified, through complete disruption of E1 and E2 genes (E1/E6 and E2/E6 ratio values were close to zero), (table 4). Considering the implication of E1 gene disruption/deletion during HPV16 integration in low grade malignancies¹⁵ and the elevated prevalence of mixed DNA form in low grade cases, we conclude that single E2/E6 analysis might miss integration events through disruption of E1 gene. This assumption was confirmed in two low grade cases (PNL2, ATT54), in which E2/E6 ratio analysis showed the episomal viral DNA form, while the E1/E6 ratio analysis revealed the mixed form. Taking these data into account we propose that both E1/E6 and E2/E6 ratio value analysis should be calculated simultaneously, in order to determine the accurate form of HPV16 genome, and this is especially needed in low grade malignancies.

The outcome of E1/E6, E2/E6 ratio value analysis also demonstrates significant information about HPV16 DNA integration pattern. In particular, it was proposed that distinct HPV16 integrants coexist in the same cervical sample, when both E1/E6 and E2/E6 ratio values predict the presence of mixed DNA form. The HPV16 DNA episomes coexist with HPV16 DNA molecules that have been integrated into the host chromosome through disruption of E1 gene and with HPV16 DNA molecules that have been integrated into the host chromosome through disruption of E2 gene, respectively. These results were encountered in 58.8% of low grade, in 88.8% of high grade malignancies, and in 25% of cervical cancer cases. The identification of distinct HPV16 integrants through E1/E6, E2/E6 ratio value analysis constitutes a serious evidence for the presence of multiple integration sites in the investigated cervical cases. Multiple integration events have been recently reported in tumor cell lines and cervical cancer cases^{10-11, 41-42}. The identification of multiple integration events in low and high grade malignancies might provide a valuable information about the progression of cervical malignancy to cancer disease.

In conclusion, the present study suggests a combined analysis of both E1/E6 and E2/E6 ratio values, in order to accurately examine the physical state of HPV16 genome. We propose that E1/E6 and E2/E6 ratio value analysis comprises an important tool to investigate the multiple integration events in the selected samples. Considering that the integration of HPV16 genome into the host chromosome is a crucial step towards carcinogenesis, both E1/E6 and E2/E6 ratio value is a prospective

prognostic tool that may provide valuable information about HPV16 integration and HPV16 physical state in the investigated cervical samples.

References

1. Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, Vignat J, Ferlay J, Bray F, Plummer M, Franceschi S. Global burden of human papillomavirus and related diseases. *Vaccine* 2012;5:F12-23.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011;61:69-90.
3. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, Snijders PJ, Meijer CJ; International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003;348:518-27.
4. Bernard HU, Calleja-Macias IE, Dunn ST. Genome variation of human papillomavirus types: phylogenetic and medical implications. *Int J Cancer* 2006;118:1071-76.
5. Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010;401:70-79.
6. zur Hausen H. Papillomavirus infections – a major cause of human cancers. *Biochim Biophys Acta* 1996;1288:F55-78
7. Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer* 2011;128:927-935
8. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010;11:1048-56
9. Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res* 2004;64:3878-84
10. Xu B, Chotewutmontri S, Wolf S, Klos U, Schmitz M, Durst M, Schwarz E. Multiplex identification of human papillomavirus 16 DNA integration sites in cervical carcinomas. *PLoS ONE* 2013;8:e66693
11. Akagi K, Li J, Broutian TR, Padilla-Nash H, Xiao W, Jiang B, Rocco JW, Teknos TN, Kumar B, Wangsa D, He D, Ried T, Symer DE, Gillison ML. Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res* 2014 doi 10.1101/gr.164806.113
12. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, Stanley MA. The biology and life-cycle of human papillomaviruses. *Vaccine* 2012;5:F55-70
13. Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 2006;44:1755-62
14. Chen CM, Shyu MP, Au LC, Chu HW, Cheng WT, Choo KB. Analysis of deletion of the integrated human papillomavirus 16 sequence in cervical cancer: a rapid multiplex polymerase chain reaction approach. *J Med Virol* 1994;44:206-11
15. Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbini M. Disruption of HPV16 E1 and E2 genes in precancerous cervical lesions. *J Virol Methods* 2009;158:180-83
16. Kalantari M, Karlsen F, Kristensen G, Holm R, Hagmar B, Johansson B. Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *Int J Gynecol Pathol* 1998;17:146-53

17. Hamid NA, Brown C, Gaston K. The regulation of cell proliferation by the papillomavirus early proteins. *Cell Mol Life Sci* 2009;66:1700-17
18. Kadaja M, Silla T, Ustav E, Ustav M. Papillomavirus DNA replication – from initiation to genomic instability. *Virology* 2009;384:360-8
19. Thiery F. Transcription regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology* 2009;384:375-9
20. Jeon S, Lambert PF. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proc Natl Acad Sci USA* 1995;92:1654–8
21. Moody CA, Laimins LA. Human papillomavirus oncoproteins: Pathways to transformation. *Nat Rev Cancer* 2010;10:550–60
22. Peter M, Stransky N, Couturier J, Hupe P, Barillot E, de Cremoux P, Cottu P, Radvanyi F, Sastre-Garau X. Frequent genomic structural alterations at HPV insertion sites in cervical carcinoma. *J Pathol* 2010;221:320–30
23. Lukaszuk K, Liss J, Wozniak I, Emerich J, Wójcikowski C. Human papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex PCR. *J Clin Microbiol* 2003;41:608-612
24. Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B. Type distribution, viral load and integration state of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer* 2005;92:2195–2200
25. Cricca M, Morselli-Labate AM, Venturoli S, Ambretti S, Gentilomi GA, Gallinella G, Costa S, Musiani M, Zerbini M. Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions. *Gynecol Oncol* 2007;106:549-57
26. Casas I, Powell L, Klapper PE, Cleator GM. New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J Virol Methods* 1995;53:25-36
27. Tsakogiannis D, Ruether I, Kyriakopoulou Z, Pliaka V, Theoharopoulou A, Skordas V, Panotopoulou E, Nepka C, Markoulatos P. Sequence variation analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from women in Greece. *Arch Virol* 2012;157:825-32
28. Li W, Wang W, Si M, Han L, Gao Q, Luo A, Li Y, Lu Y, Wang S, Ma D. The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *J Cancer Res Clin Oncol* 2008;134:1355-61
29. Sotlar K, Diemer D, Dethleffs A, Hack Y, Stubner A, Vollmer N, Menton S, Menton M, Dietz K, Wallwiener D, Kandolf R, Bültmann B. Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol.* 2004;42:3176-84
30. Tsakogiannis D, Darmis F, Gortsilas P, Ruether IG, Kyriakopoulou Z, Dimitriou TG, Amoutzias G, Markoulatos P. Nucleotide polymorphisms of the human papillomavirus 16 E1 gene. *Arch Virol* 2014;159:51-63
31. Tsakogiannis D, Papadopoulou A, Kontostathi G, Ruether IG, Kyriakopoulou Z, Dimitriou TG, Orfanoudakis G, Markoulatos P. Molecular and evolutionary analysis of HPV16 E6 and E7 genes in Greek women. *J Med Microbiol* 2013;62:1688-96
32. Hudelist G, Manavi M, Pischinger KID, Watkins-Riedel T, Singer CF, Kubista E, and Czerwenka KF. Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol Oncol* 2004;92:873–80
33. Matovina M, Sabol I, Grubisić G, Gasperov NM, Grce M. Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecol Oncol* 2009;113:120-27

34. Briolat J, Dalstein V, Saunier M, Joseph K, Caudroy S, Prétet JL, Birembaut P, Clavel C. HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. *Int J Cancer* 2007;121:2198-2204.
35. Boulet GA, Benoy IH, Depuydt CE, Horvath CA, Aerts M, Hens N, Vereecken AJ, Bogers JJ. Human papillomavirus 16 load and E2/E6 ratio in HPV16-positive women: biomarkers for cervical intraepithelial neoplasia ≥ 2 in a liquid-based cytology setting? *Cancer Epidemiol Biomarkers Prev* 2009;18:2992-99
36. Manawapat A, Stubenrauch F, Russ R, Munk C, Kjaer SK, Iftner T. Physical state and viral load as predictive biomarkers for persistence and progression of HPV16-positive cervical lesions: results from a population based long-term prospective cohort study. *Am J Cancer Res* 2012;2:192-203
37. Theelen W, Litjens RJ, Vinokurova S, Haesevoets A, Reijans M, Simons G, Smedts F, Herrington CS, Ramaekers FC, von Knebel Doeberitz M, Speel EJ, Hopman AH. Human papillomavirus multiplex ligation-dependent probe amplification assay for the assessment of viral load, integration, and gain of telomerase-related genes in cervical malignancies. *Hum Pathol* 2013;44:2410-2418.
38. Kulmala SM, Syrjänen SM, Gyllenstein UB, Shabalova IP, Petrovichev N, Tosi P, Syrjänen KJ, Johansson BC. Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 2006;59:513-7
39. Tonon SA, Picconi MA, Bos PD, Zinovich JB, Galuppo J, Alonio LV, Teyssie A. Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J Clin Virol* 2001;21:129-134
40. Huang LW, Choa SL, Lee BH, Lee BH. Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. *J Clin Pathol* 2008;61:627-31
41. Yu T, Ferber MJ, Cheung TH, Chung TK, Wong YF, Smith DI. The role of viral integration in the development of cervical cancer. *Cancer Genet Cytogenet* 2005;158:27-34
42. Li H, Yang Y, Zhang R, Cai Y, Yang X, Wang Z, Li Y, Cheng X, Ye X, Xiang Y, Zhu B. Preferential sites for the integration and disruption of human papillomavirus 16 in cervical lesions. *J Clin Virol* 2013;56:342-47

Primers	Sequence (5'-3')	Position	Product size	References
E1 target				
E1F2	GACAGCACATGCGTTGTTTAC	1059	265bp	Arias Pulido et al., 2006
E1R2	AGTCTCATGGCGCCCTTC	1323		
E2 target				
E2F12	GAATACGAACATATTTTGTGCAG	3243	297 bp	Arias Pulido et al., 2006
E2R12	CTGAGTCTCTGTGCAACAACCTA	3539		
E6 target				
HPV16 1	GCACAGAGCTGCAAACAACT	150	353 bp	Present study
HPV16 353	GGTCCACCGACCCCTTATATT	502		
GAPDH target				
GAPDH F	ACGGATTTGGTCGTATTGGG		250bp	Li et al., 2008
GAPDH R	TGATTTTGGAGGGATCTCGC			

95°C 1 min
 95°C 3 sec
 60°C 20 sec
 60°C 10 sec
 55°C-95°C denaturation

} 40 cycles

table 1: Primer pairs and assay conditions used to amplify the target-regions for E1, E2, E6 and GAPDH genes with the qRT PCR.

	Plasmid Solutions		Ratio values		HPV16 DNA form
	pE6-E2 (episomal)	pE6 (integrated)	E1/E6	E2/E6	
Solution 1	100 copies	0 copies	0.98	0.96	episomal
Solution 2	100 copies	10 copies	0.7	0.82	mixed
Solution 3	100 copies	100 copies	0.52	0.62	mixed
Solution 4	100 copies	1000 copies	0.03	0.01	mixed
Solution 5	100 copies	10000 copies	0.003	0.001	integrated
Solution 6	100 copies	100000 copies	0	0	integrated

table 2. Plasmid solutions used in order to establish the cut-off value and determine the physical state of HPV16 genome. The plasmid solutions contain a standard quantity of 100 copies of the pE6-E2 plasmid and ranging amounts of pE6 plasmid from 0 to 10⁵ copies.

Status	Low grade grade N (%)	High grade N (%)	Cancer N (%)
Episomal	6 (17.6)	0 (0)	0 (0)
Mixed	27 (79.5)	24 (88.8)	2 (25)
Integrated	1 (2,9)	3 (11.1)	6 (75)
Total	34	27	8

table 3. HPV16 genome status according to patient cytological diagnosis

Sample	Source	CIN	E2	E1	E6	E2/E6	HPV16 DNA form (E2/E6)	E1/E6	HPV16 DNA form (E1/E6)
PNL2	ThinPrep	Low grade	800	721	820	0,98	episomal	0,88	mixed
PNL3	ThinPrep	Low grade	110	240	240	0,46	mixed	1	episomal
ATT8	ThinPrep	Low grade	60	260	265	0,23	mixed	0,98	episomal
ATT18	ThinPrep	Low grade	12	172	175	0,07	mixed	0,98	episomal
ATT12	ThinPrep	Low grade	150	4400	4500	0,03	mixed	0,98	episomal
ATT46	ThinPrep	Low grade	11	27	27	0,4	mixed	1	episomal
ATT54	ThinPrep	Low grade	208	163	214	0,97	episomal	0,76	mixed
747	ThinPrep	Low grade	12	7	86475	0	integrated	0	integrated

table 4. HPV16 genome physical state of eight low grade cases.