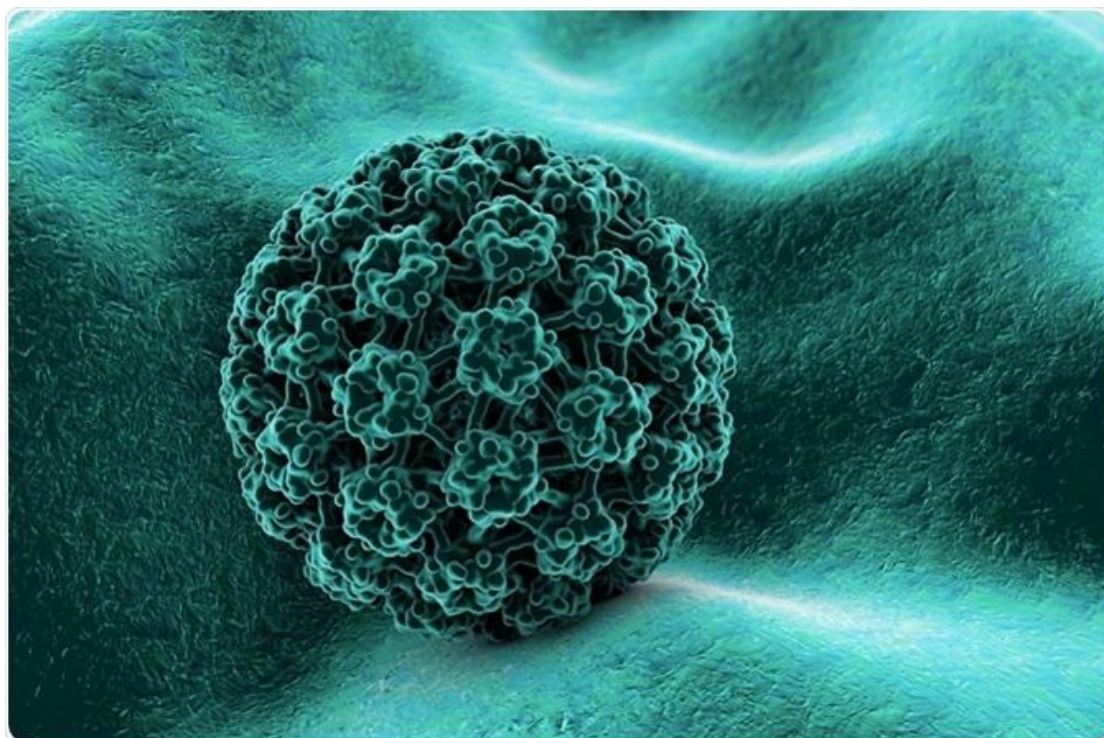




ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

Μπακούλας Αθανάσιος του Ηλία

**ΕΞΕΛΙΚΤΙΚΗ ΑΝΑΛΥΣΗ ΗΡV 16 ΜΕ
ΒΙΟΠΛΗΡΟΦΟΡΙΚΕΣ ΜΕΘΟΔΟΥΣ**



ΛΑΡΙΣΑ 2021

**«ΕΞΕΛΙΚΤΙΚΗ ΑΝΑΛΥΣΗ HPV 16 ΜΕ
ΒΙΟΠΛΗΡΟΦΟΡΙΚΕΣ ΜΕΘΟΔΟΥΣ»**

**«EVOLUTIONARY ANALYSIS OF HPV
16 USING BIOINFORMATIC METHODS»**

Η παρούσα διπλωματική εργασία εκπονήθηκε στο εργαστήριο Βιοπληροφορικής, του τμήματος Βιοχημείας και Βιοτεχνολογίας, Σχολής Επιστημών Υγείας του Πανεπιστημίου Θεσσαλίας (ΠΘ).

Επιβλέπων καθηγητής:

Αμούτζιας Γρηγόριος, Αναπληρωτής Καθηγητής Βιοπληροφορικής με έμφαση στην Μικροβιολογία, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας

Τριμελής Συμβουλευτική Επιτροπή:

Αμούτζιας Γρηγόριος, Αναπληρωτής Καθηγητής Βιοπληροφορικής με έμφαση στην Μικροβιολογία, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας

Μόσιαλος Δημήτριος, Αναπληρωτής Καθηγητής Βιοτεχνολογίας Μικροβίων, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας

Κοτταρίδη Χριστίνα, Επίκουρη Καθηγήτρια Μικροβιολογίας, Βιολογικό Τμήμα Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης.

ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά, θα ήθελα να ευχαριστήσω ιδιαίτερω τον υπεύθυνο καθηγητή μου κ. Αμούτζια Γρηγόριο για όλη τη βοήθεια που παρείχε στην εκπόνηση αυτής της πτυχιακής εργασίας, για όλα τα αντικείμενα που με δίδαξε και όλες τις συμβουλές που μου έδωσε αυτό το διάστημα που ήμουν μέλος του εργαστηρίου Βιοπληροφορικής. Επίσης θα ήθελα να ευχαριστήσω τον κ. Νικολαΐδη Μάριο που βοήθησε τα μέγιστα κάθε φορά που χρειάστηκαν οι δεξιότητες και οι γνώσεις του. Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για την τεράστια υποστήριξή τους.

ΠΕΡΙΛΗΨΗ

Οι ιοί των θηλωμάτων (PVs) είναι μια ομάδα μικρών ιών DNA που απέκτησαν την ικανότητα να μολύνουν ένα ευρύ φάσμα σπονδυλωτών, συμπεριλαμβανομένων των ανθρώπων (Human PVs). Η λοίμωξη με υψηλού κινδύνου (High Risk; HR) HPVs είναι η βασική αιτία του καρκίνου του τραχήλου της μήτρας. Συγκεκριμένα, ο HPV16 έχει ταυτοποιηθεί σε περίπου δύο τρίτα των καρκίνων του τραχήλου της μήτρας και είναι ο πιο διαδεδομένος τύπος HPV παγκοσμίως. Ο στόχος της συγκεκριμένης εργασίας είναι η εύρεση παραγόντων που οδηγούν στην εξέλιξη του HPV16. Έτσι αξιοποιήσαμε 1534 γονιδιωματικές αλληλουχίες του HPV16 οι οποίες ομαδοποιήθηκαν σε 180 ομάδες/αντιπροσώπους. Η δημιουργία δέντρων Neighbour Joining ολόκληρου γονιδιώματος αλλά και για κάθε γονίδιο ξεχωριστά μας επέτρεψε να βρούμε 5 συμβάντα ανασυνδυασμού μεταξύ των δύο μεγάλων εξελικτικών κλάδων. Οι ανασυνδυασμοί αυτοί φαίνεται να μην κυκλοφορούν στον γενικό πληθυσμό και εμφανίζονται με συχνότητα 0.3% (~1/300).

ABSTRACT

Papillomaviruses (PVs) are a group of small DNA viruses that have acquired the ability to infect a wide range of vertebrates, including humans (Human PVs). Infection with high-risk (High Risk; HR) HPV is the major cause of cervical cancer. Specifically, HPV16 has been identified in about two-thirds of cervical cancers and is the most common type of HPV worldwide. The aim of this study is to find factors that lead to the evolution of HPV16. Thus, we utilized 1534 HPV16 genomic sequences which were grouped into 180 representatives. The calculation of Neighbour Joining trees for complete genome and for each gene separately allowed us to find out 5 incongruent sequences which constitute evidence for recombinations between the two major clades. These recombinations occur with an estimated frequency of 0.3% (~1/300) and do not appear to be circulating in the general population.

Περιεχόμενα

1. ΕΙΣΑΓΩΓΗ.....	8
1.1. Ταξινόμηση	8
1.2. Γονιδιωματική οργάνωση	10
1.3. Δομή ιοσωμάτιου.....	11
1.4. Κύκλος ζωής των ΡV.....	12
1.4.1. Προσκόλληση και είσοδος.....	12
1.4.2. Αντιγραφή και μεταγραφή	13
1.4.3. Απελευθέρωση	13
1.5. Προέλευση και εξέλιξη των ιών του θηλώματος	14
1.6. Ανασυνδυασμός.....	15
1.7. Στέλεχος HPV 16.....	19
1.8. Εμβόλιο κατά των HPV	23
2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ.....	24
2.1. Λογισμικό	24
2.1.1 Linux Ubuntu 20	24
2.1.2 Python 3.6.....	24
2.2. Λήψη γονιδιωμάτων.....	25
2.2.1 Ομαδοποίηση γονιδιωμάτων.....	25
2.2.2 Υπολογισμός δέντρων	25
3. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ.....	26
4. ΒΙΒΛΙΟΓΡΑΦΙΑ.....	37

Κατάλογος Πινάκων

Πίνακας 1. Οι λειτουργίες των πρωτεϊνών του γονιδιώματος των HPV.....	11
Πίνακας 2. Ταξινόμηση των παραλλαγών του HPV16.	20

Κατάλογος Εικόνων

Εικόνα 1. Σχεδιάγραμμα ταξινόμησης των ιών των θηλωμάτων σύμφωνα με το ICTV, 2020. Οι παρενθέσεις υποδηλώνουν τον αριθμό των ειδών κάθε γένους.	9
Εικόνα 2. Πίνακας ταξινόμησης τύπων HPV. Το κόκκινο πλαίσιο χαρακτηρίζει τους HR και το κίτρινο τους LR.	9
Εικόνα 3. Κατανομή γονιδίων στο DNA των HPVs (Antonishyn N A, 2014)..	10
Εικόνα 4. (Αριστερά) Ατομική απόδοση του καψιδίου του PV. (Κέντρο) Σχηματικό διάγραμμα που αντιπροσωπεύει τα 72 καψιδιομερή ενός καψιδίου PV. Η εικοσαεδρική δομή αποτελείται αποκλειστικά από πενταμερή καψιδιομερή για συνολικά 360 καψιδιακές πρωτεΐνες. (Δεξιά) Ηλεκτρονική μικρογραφία αρνητικής αντίθεσης ισωμάτων του HPV1. Η ράβδος αντιπροσωπεύει 100 nm (Van Doorslaer et al., 2018).	12
Εικόνα 5. Οργάνωση του κύκλου ζωής του ιού.....	14
Εικόνα 6. Φυλογενετική ασυμφωνία μεταξύ πρώιμων και όψιμων γονιδίων στους Alpharparillomavirus του ανθρώπου.....	17
Εικόνα 7. Φυλογενετική ανάλυση ολόκληρων γονιδιωμάτων όλων των Alpha-PVs.....	18
Εικόνα 8. Κατανομή υπογενεαλογιών σε θετικά δείγματα HPV16, ανά γεωγραφική περιοχή.	21
Εικόνα 9. Φυλογενωμικό δέντρο των γενεαλογιών και υπογενεαλογιών του HPV16.....	22
Εικόνα 10. Φυλογενωμικό δέντρο πλήρους γονιδιώματος.....	28
Εικόνα 11. Φυλογενετικό δέντρο γονιδίου E6.....	29
Εικόνα 12. Φυλογενετικό δέντρο γονιδίου E7.....	30
Εικόνα 13. Φυλογενετικό δέντρο γονιδίου E1.....	31
Εικόνα 14. Φυλογενετικό δέντρο γονιδίου E2.....	32
Εικόνα 15. Φυλογενετικό δέντρο γονιδίου E4.....	33
Εικόνα 16. Φυλογενετικό δέντρο γονιδίου E5.....	34
Εικόνα 17. Φυλογενετικό δέντρο γονιδίου L2.....	35
Εικόνα 18. Φυλογενετικό δέντρο γονιδίου L1.....	36

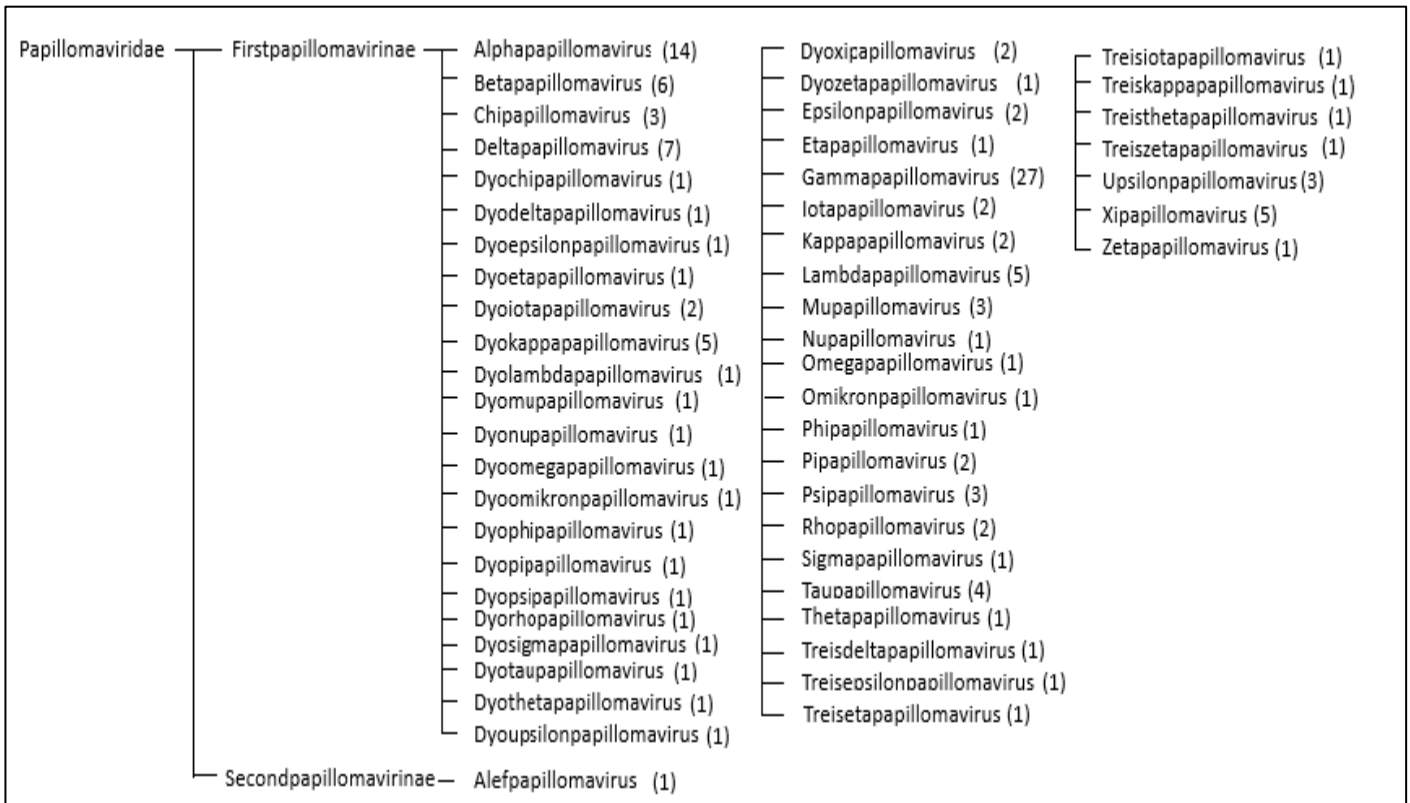
1. ΕΙΣΑΓΩΓΗ

1.1. Ταξινόμηση

Οι ιοί των θηλωμάτων (PV) είναι μια ομάδα μικρών ιών που ανήκουν στην οικογένεια των *Papillomaviridae*. Οι τύποι του ιού που προσβάλλουν τον άνθρωπο ονομάζονται Human Papillomavirus (HPV) και ανήκουν στην υποοικογένεια των *Firstpapilomavirinae*, σύμφωνα με τη Διεθνή Επιτροπή Ταξινόμησης των ιών (ICTV) (Εικόνα 1).

Οι HPV κατανέμονται σε 5 γένη: *Alphapapillomavirus*, *Betapapillomavirus*, *Gammapapillomavirus*, *Muapapillomavirus* και *Nuapapillomavirus* (Εικόνα 2). Το ICTV είναι υπεύθυνο για την ονοματολογία των ιών σε επίπεδο ειδών ενώ κάτω από το επίπεδο αυτό, το Διεθνές Κέντρο Αναφοράς HPV στη Στοκχόλμη της Σουηδίας εκχωρεί μοναδικούς αριθμούς στους τύπους HPV αφού προηγηθεί η αλληλούχιση, η κλωνοποίηση και η επιβεβαίωση του πλήρους γονιδιώματός τους από το Κέντρο (Murahwa et al., 2019). Οι διαφορετικοί τύποι HPV καθορίζονται με βάση την αλληλουχία του ανοιχτού πλαισίου ανάγνωσης (ORF) του πολύ καλά συντηρημένου γονιδίου L1. Έτσι οι HPV που παρουσιάζουν ταύτιση νουκλεοτιδικής αλληλουχίας μικρότερη από 90% στο γονίδιο L1 θεωρούνται διαφορετικά είδη, οι HPV με ταύτιση αλληλουχίας 90-98% στο γονίδιο L1 είναι διαφορετικοί υπότυποι και οι HPV με ταύτιση αλληλουχίας μεγαλύτερη από 98% στο γονίδιο L1 θεωρούνται μεταλλάγματα του ίδιου υπότυπου (Zhang et al., 2018). Μέχρι στιγμής πάνω από 200 τύποι HPV έχουν χαρακτηριστεί πλήρως (Chen et al., 2019).

Ένας άλλος διαχωρισμός που επικρατεί για τους HPV στηρίζεται στην ογκογόνο δράση τους (Εικόνα 2). Έτσι διακρίνουμε τους HPV υψηλού κινδύνου (HR) που μπορεί να οδηγήσουν σε επεμβατικό καρκίνο του τραχήλου της μήτρας και τους HPV χαμηλού κινδύνου (LR) που προκαλούν ήπια κονδυλώματα (McBride AA., 2017).



Εικόνα 1. Σχεδιάγραμμα ταξινόμησης των ιών των θηλωμάτων σύμφωνα με το ICTV, 2020. Οι παρενθέσεις υποδηλώνουν τον αριθμό των ειδών κάθε γένους.

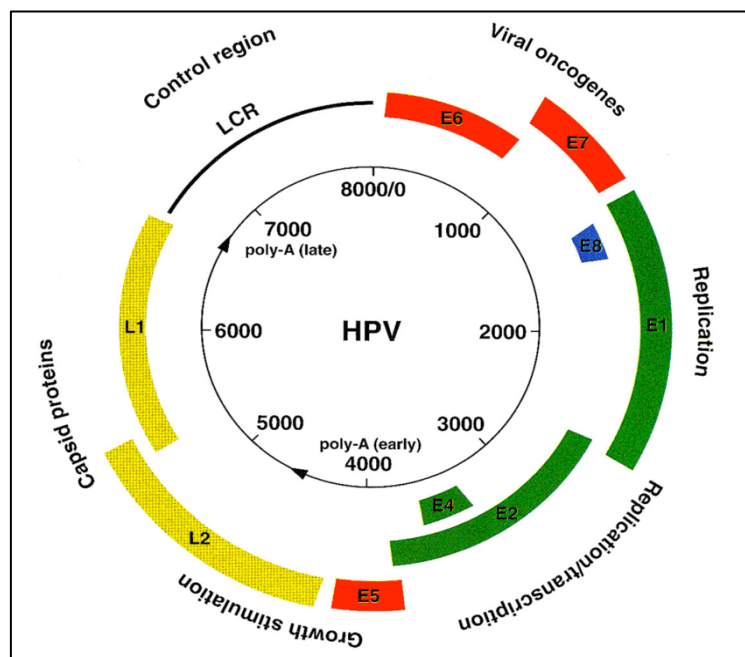
Alphapapillomavirus	HPV 2	HPV 16	HPV 31	HPV 42	HPV 54	HPV 66	HPV 73	HPV 84	HPV 94	HPV 160
	HPV 3	HPV 18	HPV 32	HPV 43	HPV 56	HPV 67	HPV 74	HPV 85	HPV 97	
	HPV 6	HPV 26	HPV 33	HPV 44	HPV 57	HPV 68	HPV 77	HPV 86	HPV 102	
	HPV 7	HPV 27	HPV 34	HPV 45	HPV 58	HPV 69	HPV 78	HPV 87	HPV 106	
	HPV 10	HPV 28	HPV 35	HPV 51	HPV 59	HPV 70	HPV 81	HPV 89	HPV 114	
	HPV 11	HPV 29	HPV 39	HPV 52	HPV 61	HPV 71	HPV 82	HPV 90	HPV 117	
	HPV 13	HPV 30	HPV 40	HPV 53	HPV 62	HPV 72	HPV 83	HPV 91	HPV 125	
Betapapillomavirus	HPV 5	HPV 19	HPV 36	HPV 80	HPV 104	HPV 118	HPV 151			
	HPV 8	HPV 20	HPV 37	HPV 92	HPV 105	HPV 120	HPV 152			
	HPV 9	HPV 21	HPV 38	HPV 93	HPV 107	HPV 122	HPV 159			
	HPV 12	HPV 22	HPV 47	HPV 96	HPV 110	HPV 124	HPV 174			
	HPV 14	HPV 23	HPV 49	HPV 98	HPV 111	HPV 143				
	HPV 15	HPV 24	HPV 75	HPV 99	HPV 113	HPV 145				
	HPV 17	HPV 25	HPV 76	HPV 100	HPV 115	HPV 150				
Gammapapillomavirus	HPV 4	HPV 101	HPV 121	HPV 131	HPV 138	HPV 147	HPV 157	HPV 166	HPV 173	HPV 199
	HPV 48	HPV 103	HPV 123	HPV 132	HPV 139	HPV 148	HPV 158	HPV 167	HPV 175	HPV 200
	HPV 50	HPV 108	HPV 126	HPV 133	HPV 140	HPV 149	HPV 161	HPV 168	HPV 178	HPV 201
	HPV 60	HPV 109	HPV 127	HPV 134	HPV 141	HPV 153	HPV 162	HPV 169	HPV 179	HPV 202
	HPV 65	HPV 112	HPV 128	HPV 135	HPV 142	HPV 154	HPV 163	HPV 170	HPV 180	HPV 205
	HPV 88	HPV 116	HPV 129	HPV 136	HPV 144	HPV 155	HPV 164	HPV 171	HPV 184	
	HPV 95	HPV 119	HPV 130	HPV 137	HPV 146	HPV 156	HPV 165	HPV 172	HPV 197	
Mupapillomavirus	HPV 1									
	HPV 63									
	HPV 204									
Nupapillomavirus	HPV 41									

Εικόνα 2. Πίνακας ταξινόμησης τύπων HPV. Το κόκκινο πλαίσιο χαρακτηρίζει τους HR και το κίτρινο τους LR.

1.2. Γονιδιωματική οργάνωση

Οι HPV είναι ιοί μικρού μεγέθους και περιέχουν μη τυλιγμένο, κυκλικό δίκλωνο DNA μεγέθους 7-8 kbp. Το γονιδίωμα τους κωδικοποιεί οκτώ ORFs, τα οποία είναι οργανωμένα σε δύο γενικές περιοχές. Επίσης υπάρχει μια ρυθμιστική περιοχή περίπου 800 νουκλεοτιδίων. Η ρυθμιστική περιοχή (URR) βρίσκεται ανοδικά του E6 και περιέχει στοιχεία που ελέγχουν τη μεταγραφή και την αντιγραφή. Η περιοχή πρώιμων γονιδίων κωδικοποιεί τις πρωτεΐνες E6, E7, E1, E2, E4 και E5 (E από τη λέξη Early). Η όψιμη περιοχή κωδικοποιεί τις δομικές πρωτεΐνες L2 και L1(L; Late) (Εικόνα 3).

Το μικρό γονιδίωμα είναι γεμάτο με επικαλυπτόμενα ανοιχτά πλαίσια ανάγνωσης και cis ρυθμιστικά στοιχεία (McBride AA., 2017). Οι πρωτεΐνες E6, E7, E1, E2, E4 και E5 εκφράζονται αμέσως μετά την αρχική μόλυνση ενός κυττάρου ξενιστή, ενώ οι δομικές πρωτεΐνες L2 και L1 είναι απαραίτητες για το σχηματισμό του καψιδίου του ιού. Η L1 αποτελεί την κύρια πρωτεΐνη και η L2 τη δευτερεύουσα. Η περιοχή URR βρίσκεται μεταξύ των L1 και E6 ORFs και περιέχει τη θέση έναρξης της αντιγραφής (ori) καθώς και θέσεις σύνδεσης για μεταγραφικούς παράγοντες (Zheng and Baker, 2006). Οι λειτουργίες τους φαίνονται περιληπτικά στον Πίνακα 1.1. Μεταξύ των γονιδίων του ιού, τα E6 και E7 είναι καλά συντηρημένα στους τύπους HPV που προκαλούν λοιμώξεις και βλάβες υψηλού βαθμού. Όλα τα ORFs του ιού κωδικοποιούνται σε έναν κλώνο DNA. Επίσης σε όλους τους HPVs συναντούμε το εξόνιο E8, που είναι ενσωματωμένο στο γονίδιο E1, και κωδικοποιεί την πρωτεΐνη E8 ^ E2 η οποία καταστέλλει τη μεταγραφή και την αντιγραφή του ιού, αποφεύγοντας έτσι την ανίχνευση από το ανοσοποιητικό σύστημα του ξενιστή (McBride AA., 2017).



Εικόνα 3. Κατανομή γονιδίων στο DNA των HPVs (Antonishyn N A, 2014).

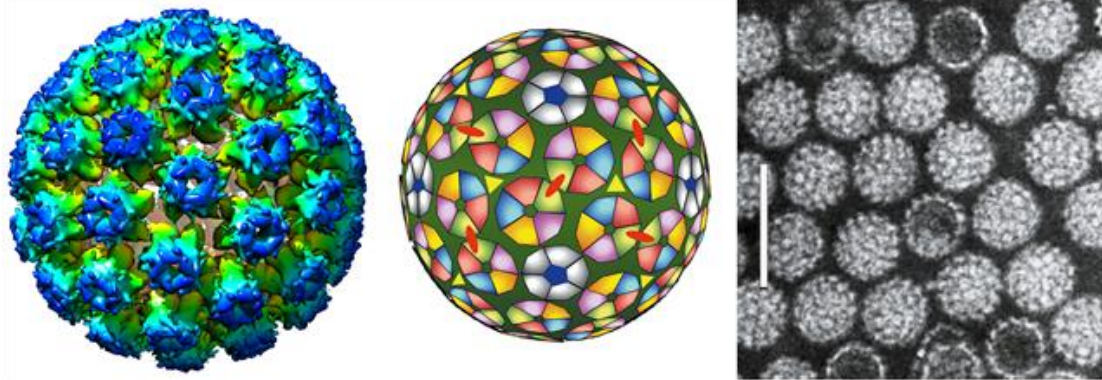
Πρωτεΐνες	Λειτουργία
E1	Συνδέεται στη θέση έναρξης της αντιγραφής, έχει δραστικότητα ελικάσης.
E2	Κύριος ρυθμιστής μεταγραφής, διευκολύνει τη δέσμευση του E1 στη θέση ορί, προσδένει το γονιδίωμα του ιού σε κυτταρικά χρωμοσώματα, χρησιμεύει ως αρνητικός ρυθμιστής έκφρασης των E6 και E7 σε μολυσμένα κερατινοκύτταρα της βασικής στιβάδας.
E4	Διευκολύνει την απελευθέρωση ιοσωμάτων στο περιβάλλον διαταράσσοντας τα ενδιάμεσα νημάτια του κυτταροσκελετού των κερατινοκυττάρων.
E5	Αποσταθεροποιεί τη λειτουργία πολλών μεμβρανικών πρωτεϊνών στο μολυσμένο κύτταρο, ενεργοποιεί την ανάπτυξη κυττάρων, αποτρέπει την εξάλειψη του μολυσμένου κυττάρου από τα φονικά T κύτταρα.
E6	Μεσολαβεί στην αποδόμηση της p53 (μείζον πρωτεΐνη καταστολής όγκων), αλληλεπιδρά με περιοχές PDZ, αυξάνει τα επίπεδα τελομεράσης.
E7	Αδρανοποιεί μέλη της οικογένειας πρωτεϊνών pRb, συμμετέχει στην αθανатоποίηση των μολυσμένων κυττάρων ενεργοποιώντας την κυτταρική τελομεράση.
L1	Κύρια πρωτεΐνη καψιδίου.
L2	Δευτερεύουσα πρωτεΐνη καψιδίου, συνοδεύει το γονιδίωμα μέσω των ενδοκυτταρικών οδών στον πυρήνα, συσκευάζει το DNA του ιού στο ιικό σωματίο.

Πίνακας 1. Οι λειτουργίες των πρωτεϊνών του γονιδιώματος των HPV (McBride AA., 2017).

1.3. Δομή ιοσωμάτιου

Οι ιοί των θηλωμάτων δεν διαθέτουν λιπιδικό έλυτρο. Η L1, είναι απαραίτητη για το σχηματισμό του καψιδίου που έχει διάμετρο 55-60 νανομέτρων και αποτελείται από 72 καψιδιομερή διατεταγμένα σε σχήμα αστεριού (Rayment et al., 2014). Όπως και οι περισσότεροι ιοί χωρίς έλυτρο, το καψίδιο είναι γεωμετρικά κανονικό και παρουσιάζει εικοσαεδρική συμμετρία (Εικόνα 4).

Κάθε καψίδιο συσκευάζει ένα μόνο αντίγραφο του κυκλικού δίκλωνου DNA του ιού. Αν και δεν είναι σαφής ο τρόπος με τον οποίο είναι διατεταγμένη η L2 μέσα στο ιικό σωματίο, εκτελεί πολλές σημαντικές λειτουργίες, συμπεριλαμβανομένης της διευκόλυνσης της συσκευασίας του γονιδιώματος του ιού στα νέα ιοσωμάτια καθώς και της μολυσματικής εισόδου του ιού σε νέα κύτταρα ξενιστές (McBride AA., 2017).



Εικόνα 4. (Αριστερά) Ατομική απόδοση του καψιδίου του PV. (Κέντρο) Σχηματικό διάγραμμα που αντιπροσωπεύει τα 72 καψιδιομερή ενός καψιδίου PV. Η εικοσαεδρική δομή αποτελείται αποκλειστικά από πενταμερή καψιδιομερή για συνολικά 360 καψιδιακές πρωτεΐνες. (Δεξιά) Ηλεκτρονική μικρογραφία αρνητικής αντίθεσης ιοσωμάτων του HPV1. Η ράβδος αντιπροσωπεύει 100 nm (Van Doorslaer et al., 2018).

1.4. Κύκλος ζωής των PV

1.4.1. Προσκόλληση και είσοδος

Οι ιοί των θηλωμάτων αποκτούν πρόσβαση στα βλαστοκύτταρα των κερατινοκυττάρων μέσω μικρών πληγών - μικροτραυμάτων, στην επιφάνεια του δέρματος ή του βλεννογόνου. Αρχικά το καψίδιο του ιού αλληλεπιδρά με πρωτεογλυκάνες θειικής ηπαρίνης στη βασική μεμβράνη, η οποία επάγει μια αλλαγή στη διαμόρφωση που επιτρέπει στον ιό να συνδεθεί με έναν δευτερεύοντα υποδοχέα, πιθανώς την ιντεγκρίνη $\alpha_6\beta_4$, στην επιφάνεια των βασικών κερατινοκυττάρων. Στη συνέχεια ο ιός μεταφέρεται σε μεμβρανικά κυστίδια, τα ενδοσώματα, προς το δίκτυο Golgi. Η L2 πρωτεΐνη του καψιδίου διαταράσσει τη μεμβράνη του ενδοσώματος μέσω ενός πεπτιδίου που δεισδύει στα κύτταρα, επιτρέποντας στο γονιδίωμα του ιού να διαφύγει. Εν συνεχεία το γονιδίωμα του ιού μαζί με την L2, εισέρχεται στον πυρήνα όπου εγκλείεται σε ένα μεμβρανικό κυστίδιο. Το κύτταρο πρέπει να υποβληθεί σε μίτωση και διάσπαση του πυρηνικού φακέλου για να επιτρέψει στο σύμπλοκο L2-γονιδιώματος να έχει πρόσβαση στον πυρήνα. Έπειτα, το σύμπλεγμα L2-γονιδιώματος παρατηρείται δίπλα στα πυρηνικά σώματα ND10, τα οποία είναι σημαντικά για την εγγενή ανοσολογική άμυνα. Οι ιοί συχνά αναδιοργανώνουν τα συστατικά των σωμάτων ND10, και η πρωτεΐνη L2 αντικαθιστά την πρωτεΐνη Sp100 και στρατολογεί την Daxx, για να δημιουργήσει το κατάλληλο περιβάλλον για την έναρξη της αντιγραφής και της μεταγραφής του ιού (McBride AA., 2017).

1.4.2. Αντιγραφή και μεταγραφή

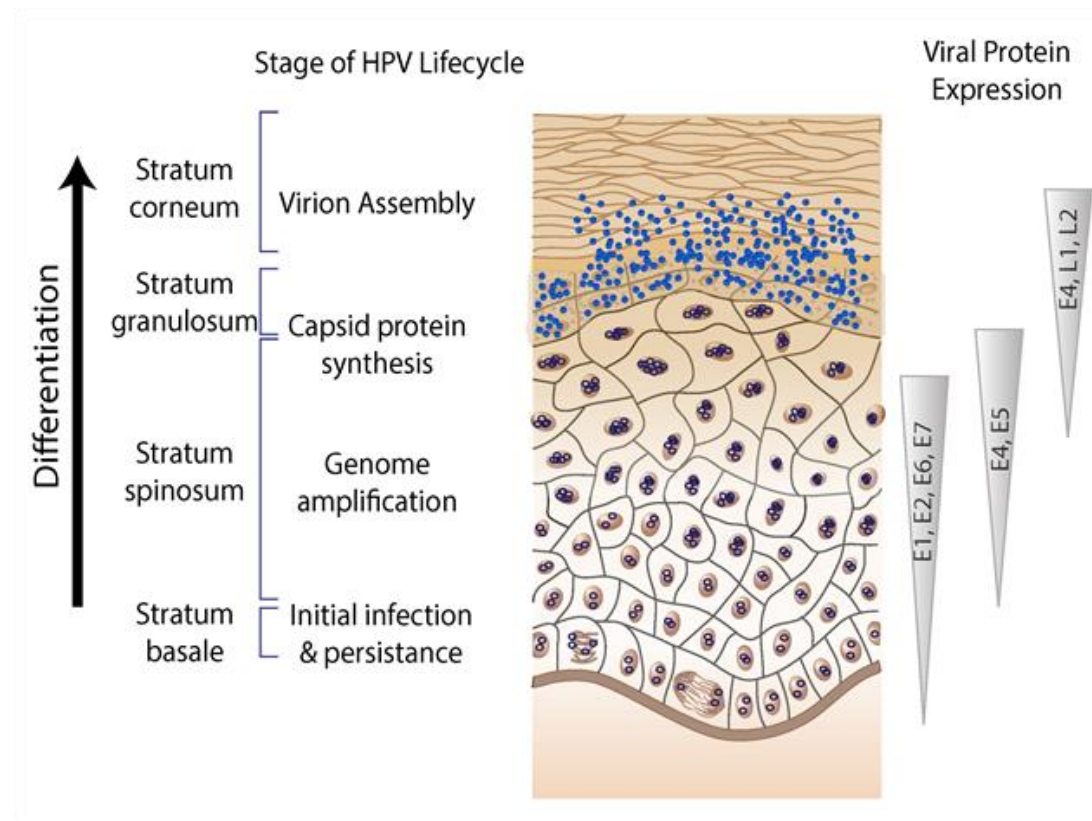
Μετά τη μόλυνση ενός κερατινοκυττάρου, ο ιός εκφράζει τις πρωτεΐνες E1 και E2, οι οποίες προορίζονται για διπλασιασμό και διατήρηση του DNA του ιού. Η E2 στρατολογεί την ελικάση E1 στη θέση έναρξης της αντιγραφής του ιού. Υπάρχει όμως και εναλλακτική περίπτωση όπου ο ιός βασίζεται σε κυτταρικές πρωτεΐνες για να αντιγράψει το DNA του. Συγκεκριμένα, υπάρχουν τρεις φάσεις αντιγραφής στον κύκλο ζωής του ιού. Στην πρώτη φάση, όταν ο ιός μολύνει για πρώτη φορά το κύτταρο υπάρχει περιορισμένη ενίσχυση του DNA. Στη συνέχεια υπάρχει η αντιγραφή της συντήρησης, όταν το γονιδίωμα του ιού αντιγράφεται σε έναν σταθερό αριθμό αντιγράφων στα πολλαπλασιαζόμενα κύτταρα μιας πληγής. Και τέλος υπάρχει η ενίσχυση που εξαρτάται από τη διαφοροποίηση, όταν το DNA του ιού αντιγράφεται σε υψηλό αριθμό αντιγράφων για τη δημιουργία ιοσωμάτων. Τόσο η πρώιμη όσο και η όψιμη ενίσχυση του DNA του ιού εμπλέκει την απόκριση σε βλάβη του κυτταρικού DNA (DDR) για να υποστηρίξει τη σύνθεση του. Κατά τη φάση της συντήρησης, η πρωτεΐνη E2 διασφαλίζει ότι ο χαμηλός αριθμός αντιγράφων κατανέμεται αποτελεσματικά στα θυγατρικά κύτταρα με την πρόσδεσή τους στη χρωματίνη του ξενιστή. Για να επιτευχθεί αυτό, η E2 περιέχει τόσο μία περιοχή δέσμευσης DNA που αλληλεπιδρά με διατηρημένες θέσεις στο γονιδίωμα του ιού, όσο και μία περιοχή που αλληλεπιδρά με την χρωματίνη του ξενιστή (McBride AA., 2017).

Η μεταγραφή του ιού πραγματοποιείται σε τρία κύματα, τα οποία εξαρτώνται από την κατάσταση διαφοροποίησης του κυττάρου ξενιστή. Η πρώιμη μεταγραφή ξεκινά από τους πρώιμους υποκινητές που βρίσκονται ακριβώς πριν την πρώιμη περιοχή και τερματίζεται στη θέση της πρώιμης πολυαδενυλίωσης. Η ενδιάμεση μεταγραφή ξεκινά από τον όψιμο υποκινητή και μεταγράφει υψηλά επίπεδα των πρωτεϊνών E1 και E2, αλλά εξακολουθεί να τερματίζεται στον πρώιμο υποκινητή. Τέλος η όψιμη μεταγραφή χρησιμοποιεί τόσο τον όψιμο υποκινητή όσο και την όψιμη θέση πολυαδενυλίωσης και οδηγεί σε έκφραση υψηλών επιπέδων των L1 και L2 (McBride AA., 2017).

1.4.3. Απελευθέρωση

Στα ανώτερα στρώματα της μολυσμένης πληγής, το DNA του ιού ενισχύεται σε μεγάλο αριθμό αντιγράφων. Αυτή η φάση απαιτεί την επαγωγή του DNA Damage Response (DDR), το οποίο στρατολογεί παράγοντες επιδιόρθωσης μέσω των οποίων ο ιός μπορεί να εισβάλει και να αντιγράψει το DNA του. Η πρωτεΐνη E4 εκφράζεται επίσης άφθονα στα ανώτερα στρώματα της βλάβης, όπου αναδιοργανώνει το δίκτυο νημάτων κερατίνης για να διευκολύνει την απελευθέρωση και τη μετάδοση του ιού. Τα υψηλά επίπεδα των L1 και L2 έχουν ως αποτέλεσμα τη σύνθεση καψιδίων που

ενσωματώνουν το DNA του ιού. Τελικά τα επιφανειακά κύτταρα απελευθερώνουν τις συστοιχίες των νέων ισωμάτων (Εικόνα 5) (McBride AA., 2017).



Εικόνα 5. Οργάνωση του κύκλου ζωής του ιού. Τα διαφορετικά στρώματα του επιθηλίου φαίνονται στα αριστερά. Τα σχετικά επίπεδα πρωτεΐνης συνοψίζονται χρησιμοποιώντας τρίγωνα (Van Doorslaer et al., 2018).

1.5. Προέλευση και εξέλιξη των ιών του θηλώματος

Εκτιμάται πως οι PVs εμφανίστηκαν πριν από 250-150 εκατομμύρια χρόνια στη Μεσοζωική εποχή με τον πιο πρόσφατο κοινό πρόγονο (MRCA) των *Alpha*, *Beta*, *Gamma*, *Mu* και *Nu* HPV να εμφανίζεται μεταξύ 30 και 50 εκατομμύρια χρόνια πριν (Murahwa et al., 2019). Θεωρείται ότι οι PV είχαν συν εξελιχθεί με τους ξενιστές τους επειδή η φυλογένεση τους έμοιαζε με εκείνη των ξενιστών τους. Ωστόσο, διάφοροι παράγοντες όπως η κυτταρική ή χυμική ανοσοαπόκριση του ξενιστή μπορεί να επηρεάσουν την εξέλιξη των HPV με την πάροδο του χρόνου. Επιπλέον, τα φυλογενετικά δέντρα των PVs των πρώιμων και όψιμων περιοχών παρουσιάζουν διαφορετικές τοπολογίες, γεγονός που υποδηλώνει ανασυνδυασμό. Αυτά τα δεδομένα υποστηρίζουν ότι πολλές εξελικτικές δυνάμεις όπως η μόλυνση μεταξύ διαφορετικών ειδών, ο ανασυνδυασμός και η αντιγραφή των ιών οδήγησαν στη διαφοροποίηση τους (Cristina M. de Oliveira, 2018).

Έχει αποδειχθεί ότι τα γονίδια των PVs έχουν διαφορετικούς ρυθμούς εξέλιξης. Τα νέα γονίδια E5, E6 και E7, που κωδικοποιούν ογκοπρωτεΐνες, αποκλίνουν ταχύτερα από τα παλαιότερα τέσσερα γονίδια E1, E2, L1 και L2. Επιπλέον, ο αριθμός των μεταλλάξεων είναι υψηλότερος στα E5, E6 και E7 ογκογονίδια από ό, τι στα υπόλοιπα. Προτείνεται λοιπόν ότι η ιστορία των PVs πραγματοποιήθηκε σε διαφορετικά στάδια, με το πρώτο στάδιο να αντιπροσωπεύει την αρχική εμφάνιση ενός πρωτοτύπου PV με τα βασικά γονίδια (E1, E2, L2 και L1) που βρίσκονται σε όλα τα PVs. Στη συνέχεια, το δεύτερο στάδιο περιλάμβανε περαιτέρω εξελικτικές διεργασίες που οδήγησαν στην απόκτηση των ογκογονιδίων E5, E6 και E7, και αυτά εξελίχθηκαν περίπου δύο φορές ταχύτερα από την υπόλοιπη περιοχή του γονιδιώματος (Murahwa et al., 2019).

1.6. Ανασυνδυασμός

“Ο αυστηρός ορισμός του ανασυνδυασμού ενσωματώνει την παλινδρόμηση, που σημαίνει ότι ο παραλήπτης ενός γονιδιακού τμήματος ενεργεί ως δότης του αντικατασταθέντος τμήματος στην πηγή, κάτι το οποίο δεν συμβαίνει με την περίπτωση του ανασυνδυασμού των PVs” (Pérez-Losada et al., 2015). Ο ανασυνδυασμός των PVs θα μπορούσε να ονομαστεί καταλληλότερα, μετατροπή γονιδίου αλλά ο ορισμός αυτός έχει χρησιμοποιηθεί ευρέως και η αλλαγή του θα προκαλούσε σύγχυση (Alltalents T. Murahwa, 2020). Οι διάφορες έρευνες για τον ανασυνδυασμό αποδεικνύουν το σημαντικό αντίκτυπο του στην εξέλιξη των PVs.

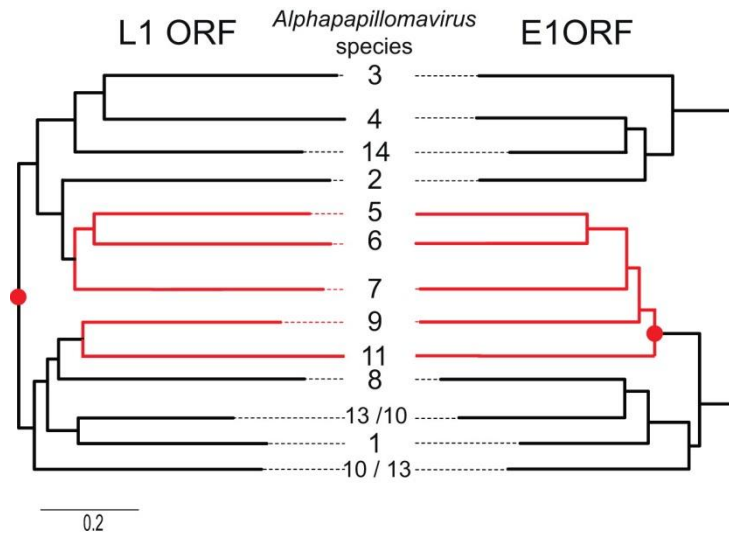
Μία στατιστική ανάλυση του 2006 η οποία αφορούσε τοπολογικές διαφορές στα φυλογενετικά δέντρα διαφόρων γονιδίων των α-PVs, παρείχε πειστικές ενδείξεις για ανασυνδυασμό των PVs. Η χρήση μεθόδων ανίχνευσης ανασυνδυασμού επέτρεψε την ακριβέστερη αναγνώριση του τμήματος του γονιδίου L2 όπου συνέβη ο ανασυνδυασμός. Οι φυλογενετικές αυτές αναλύσεις αποκάλυψαν επίσης ότι ο ανασυνδυασμός πιθανώς συνέβη πριν τη διαφοροποίηση των κύριων ομάδων των PVs. Αυτό σημαίνει ότι οι βιώσιμοι ανασυνδυασμένοι ιοί, όταν σχηματιστούν, μπορούν να εξαπλωθούν και να αναπτυχθούν. Η μελέτη επίσης συσχέτισε τη μόλυνση από HPV με τη μόλυνση από HIV και υποστήριξε ότι ενώ οι λοιμώξεις με HPVs σε ασθενείς με HIV είναι πιθανό να ευνοούν τη δημιουργία ανασυνδυασμών, η ανοσολογική δυσλειτουργία που προκαλεί ο HIV μπορεί να διευκολύνει τόσο τη βραχυπρόθεσμη επιβίωση όσο και την ενδεχόμενη προσαρμογή των ανασυνδυασμών σε νέες θέσεις στον ανθρώπινο οργανισμό (A. Varsani, 2006).

Έρευνα του 2008 για τον χαρακτηρισμό νέων PVs δελφινιών παρείχε ενδείξεις για ανασυνδυασμό των ιών. Η φυλογενετική ανάλυση που πραγματοποιήθηκε μεταξύ του δέντρου του E1 ORF και εκείνου του L1 ORF έδειξε διαφορετική ομαδοποίηση των υπό μελέτη PVs κάτι το οποίο

υποδεικνύει πιθανό συμβάν ανασυνδυασμού. Η εκτιμώμενη θέση ανασυνδυασμού εντοπίζεται κοντά στο τέλος του γονιδίου E2 και στην αρχή του L2. Το συμπέρασμα των αναλύσεων είναι ο ανασυνδυασμός μεταξύ της πρώιμης περιοχής του προγόνου ενός είδους δελφινιών και της όψιμης περιοχής του προγόνου ενός εκ των νέων PVs. Γίνεται φανερός λοιπόν ένας εξελικτικός μηχανισμός στον οποίο οι κασέτες πρώιμων και όψιμων γονιδίων των PVs είναι σχετικά ανεξάρτητες οντότητες που μπορούν να εναλλάσσονται με ανασυνδυασμό (A. Rector et al., 2008).

Το 2009 μελέτη για την εξέταση της γονιδιωματικής ποικιλότητας των παραλλαγών του HPV16 ανέδειξε την παρουσία ενδοτυπικού ανασυνδυασμού σε φυσική λοίμωξη. Η ενδοτυπική ποικιλότητα του γονιδιώματος των HPV γενικά τεκμηριώνεται ως σημειακή μετάλλαξη, εισαγωγή και / ή διαγραφή. Στην εν λόγω μελέτη βρέθηκαν ανασυνδυασμένες αλληλουχίες της Ευρωπαϊκής καθώς και της Αφρικάνικης 2 παραλλαγής σε ένα δείγμα, οι οποίες διέφεραν στον αριθμό των αντιγράφων. Οι ερευνητές κατέληξαν στο γεγονός ότι ο ανασυνδυασμός μπορεί να συμβεί μετά τη διαφοροποίηση των παραλλαγών των HPVs και να οδηγήσει εν τέλει στην εξέλιξη των τρεχόντων τύπων και παραλλαγών των HPVs. Αυτό μπορεί εν μέρει να εξηγήσει την εξαιρετική ποικιλία αυτής της οικογένειας ιών (M. Jiang, 2009).

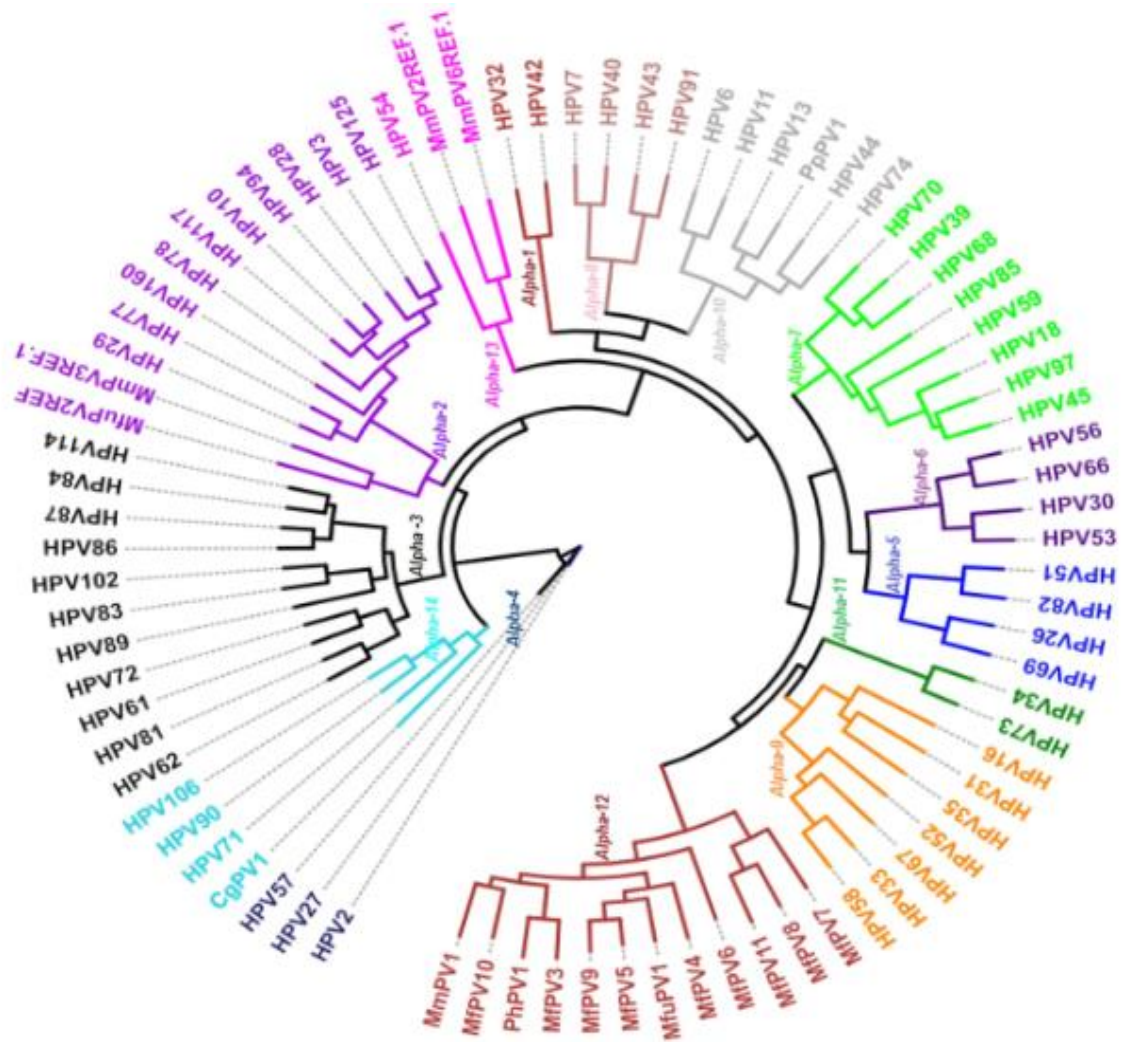
Σε έρευνα του 2013, η ανάλυση υποδηλώνει ένα πιθανό συμβάν ανασυνδυασμού που χαρτογραφείται κοντά στο τέλος του E2 γονιδίου και στην αρχή του L2. Μια δεύτερη πιθανή θέση ανασυνδυασμού εκτιμάται ότι βρίσκεται εντός της LCR περιοχής του γονιδιώματος του ιού. Επιπλέον η συγκεκριμένη έρευνα κάνει λόγο για το γένος *Alpha*-PV που περιέχει όλους τους τύπους PV που προκαλούν καρκίνο, υποδηλώνοντας έτσι την ύπαρξη ενός κοινού προγόνου για τους ογκογόνους HPVs. Ωστόσο, σε ξεχωριστά δέντρα που κατασκευάστηκαν από τα πρώιμα ORFs (E1 και E2) και τα όψιμα ORFs (L1, L2), το καρκινογόνο δυναμικό ταιριάζει με την πρώιμη και όχι με την όψιμη περιοχή του γονιδιώματος των PVs (Εικόνα 6). Ένα συμβάν ανασυνδυασμού στη ρίζα του *Alphapapillomavirus* μπορεί να εξηγήσει αυτήν την παρατήρηση. Τέλος αναφέρεται πως μέσα από τη διαδικασία της συν εξέλιξης, κάθε μεμονωμένος τύπος PV βελτιστοποίησε το οπλοστάσιό του για να μολύνει με επιτυχία μια συγκεκριμένη θέση στον ξενιστή του για εκατομμύρια χρόνια. Επομένως, ο ανασυνδυασμός παρέχει στον ιό ένα πλεονέκτημα έναντι των ιών που δεν ανασυνδυάζονται (K. Van Doorslaer, 2013).



Εικόνα 6. Φυλογενετική ασυμφωνία μεταξύ πρώιμων και όψιμων γονιδίων στους *Alphapapillomavirus* του ανθρώπου. Με κόκκινο χρώμα επισημαίνεται το κλαδί που περιέχει τους ογκογόνους ιούς (K. VanDoorslaer, 2013).

Η έκθεση μιας πιο πρόσφατης μελέτης του 2016 που ασχολήθηκε με το γένος *Gamma papillomavirus* έδειξε την παρουσία δύο συμβάντων ανασυνδυασμού. Οι ανασυνδυασμοί αυτοί εντοπίστηκαν εντός του E1 ORF σε αντίθεση με τους ανασυνδυασμούς των *Alpha-PVs* που εντοπίζονται εντός του γονιδίου L2. Το πρώτο συμβάν αντιπροσώπευε έναν ανασυνδυασμό μεταξύ ειδών ενώ το δεύτερο μεταξύ τύπων ενός είδους (E.M. Bolatti et al., 2016).

Η πρώτη απόδειξη αλληλεπιδράσεων μεταξύ PVs διαφορετικών ξενιστών λαμβάνει χώρα σε έρευνα του 2020. Οι αναλύσεις έδειξαν την ύπαρξη συμβάντων ανασυνδυασμού μεταξύ των *Alpha-PVs* υψηλού κινδύνου, μεταξύ HPVs και PVs μη πρωτευόντων καθώς και μεταξύ των PVs κητοειδών (Εικόνα 7). Αποδεικνύεται έτσι για πρώτη φορά ανασυνδυασμός μεταξύ ξενιστών καθώς και μεταξύ γενών. Ωστόσο οι παραδοχές της μελέτης δεν μπορούν απαραίτητα να γενικευτούν και να εφαρμοστούν σε όλους τους PVs (Alltalents T. Murahwa et al., 2020).



Εικόνα 7. Φυλογενετική ανάλυση ολόκληρων γονιδιωμάτων όλων των Alpha-PVs (Alltalents T. Murahwa et al., 2020).

Σε πρόσφατη έρευνα που αφορά τους PVs κητοειδών συσχετίζονται οι φυλογενετικές σχέσεις με το περιεχόμενο του γονιδιώματος και παρατηρούνται σημαντικές διαφορές μεταξύ ανασυνδυασμένων και μη PVs του συγκεκριμένου είδους. Οι αναλύσεις που παρουσιάζονται δείχνουν ότι οι ανασυνδυασμένοι PVs των κητοειδών περιέχουν ένα μοναδικό σύνολο μοτίβων στη ρυθμιστική περιοχή, υποδεικνύοντας ότι κατά τον ανασυνδυασμό αυτοί οι ιοί έχουν ακολουθήσει μια συγκεκριμένη εξελικτική πορεία. Πιθανώς, τα μοτίβα αυτά εξελίχθηκαν ως προσαρμοστική απόκριση στην ανάγκη επιπρόσθετης ρύθμισης για αποτελεσματική έκφραση ή αντιγραφή των γονιδίων των χιμαιρικών αυτών γονιδιωμάτων (F. Borveto et al., 2020).

1.7. Στέλεχος HPV 16

Οι περισσότεροι ογκογόνοι ή υψηλού κινδύνου (HR) τύποι HPV που σχετίζονται με τον καρκίνο του τραχήλου της μήτρας συγκεντρώνονται σε ένα κλαδί των *Alpha*-HPV που περιέχει τα είδη *alpha*-5, *alpha*-6, *alpha*-7 και *alpha*-9 και αντιπροσωπεύουν το 90% όλων των καρκίνων του τραχήλου της μήτρας παγκοσμίως (R.D. Burk et al., 2013). Ο Διεθνής Οργανισμός Έρευνας για τον Καρκίνο θεωρεί καρκινογόνους 12 τύπους HPV, με τον 16 των ανθρώπινων θηλωμάτων 16 να είναι ο πιο καρκινογόνος, αντιπροσωπεύοντας περισσότερους από τους μισούς από τους εκτιμώμενους 620.000 καρκίνους του τραχήλου της μήτρας που διαγιγνώσκονται παγκοσμίως και περισσότερους από 200.000 θανάτους ετησίως. Ωστόσο, οι περισσότερες μολύνσεις του τραχήλου της μήτρας από HR-HPV είναι καλοήθειες και υποχωρούν αυθόρμητα (G.M. Clifford et al., 2019). Μόνο ένα μικρό ποσοστό (<5%) των γυναικών που έχουν μολυνθεί με έναν από αυτούς τους τύπους θα αναπτύξει προ-καρκίνο του τραχήλου της μήτρας (Mirabello et al., 2018).

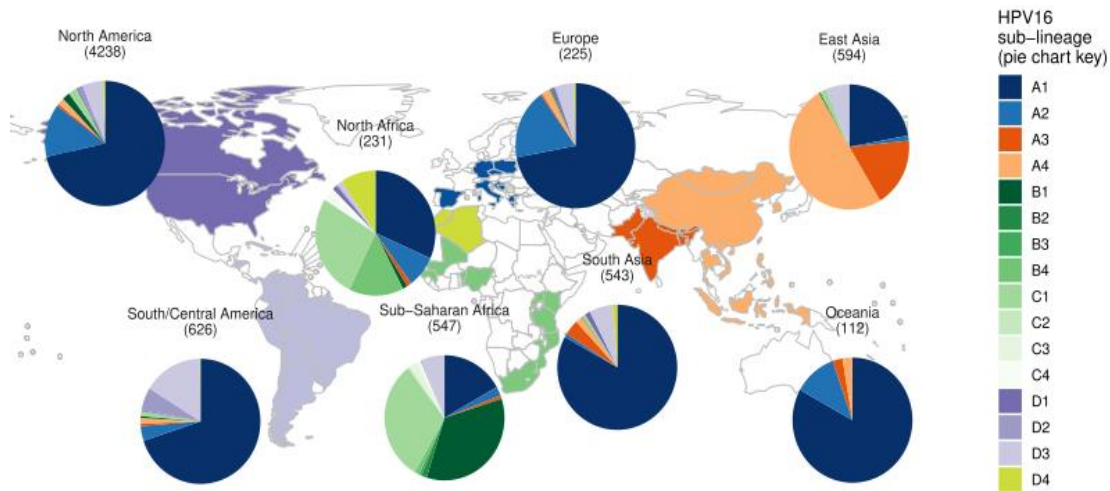
Απομονώσεις του ίδιου τύπου HPV αναφέρονται ως παραλλαγές ή υπότυποι όταν οι νουκλεοτιδικές αλληλουχίες του L1 ORF διαφέρουν λιγότερο από 10%. Αντίθετα η ταξινόμηση των παραλλαγών των γονιδιωμάτων του ίδιου τύπου HPV βασίζεται στο πλήρες γονιδίωμα. Έτσι ο προσδιορισμός των γενεαλογιών βασίζεται στη διαφορά νουκλεοτιδικής αλληλουχίας εύρους 1,0% – 10% και ο προσδιορισμός των υπογενεαλογιών στη διαφορά νουκλεοτιδικής αλληλουχίας εύρους 0,5% – 1,0% (Chen et al., 2011). Ο HPV16 μπορεί να ταξινομηθεί σε τέσσερις γενεαλογίες (A, B, C και D), και 16 υπογενεαλογίες (A1–4, B1–4, C1–4 και D1-4), οι οποίες ποικίλλουν στη γεωγραφική κατανομή και στους κινδύνους ασθένειας (Πίνακας 2) (R.D. Burk et al., 2013). Η κατανομή των γενεαλογιών του HPV16 εξαρτάται από τον πληθυσμό και θεωρείται ότι οφείλεται στην αρχαία συν εξέλιξη των HPVs και των ανθρώπων, σε συνδυασμό με πιο πρόσφατα πρότυπα μετανάστευσης (Εικόνα 8). Υπάρχουν δεδομένα που υποστηρίζουν ότι οι παραλλαγές του HPV16 επηρεάζουν την “επιμονή” του ιού, την ανάπτυξη καρκίνου και τον ιστολογικό τύπο καρκίνου του τραχήλου της μήτρας (G.M. Clifford, et al., 2019). Αν και μόνο μια μειονότητα των λοιμώξεων από τον HPV16 “επιμένει” πέραν του 1 έτους, ένα μεγάλο ποσοστό αυτών που παραμένουν θα συνεχίσουν να αναπτύσσουν προ-καρκίνο.

ΕΙΔΟΣ	ΤΥΠΟΣ	ΓΕΝΕΑΛΟΓΙΑ	ΥΠΟΓΕΝΕΑΛΟΓΙΑ	ΕΝΑΛΛΑΚΤΙΚΟ ΟΝΟΜΑ
Alpha-9	HPV 16	A	A1	European (E)
			A2	
			A3	
			A4	
		B	B1	African-1 (Af-1)
			B2	
			B3	
			B4	
		C	C1	African-2 (Af-2)
			C2	
			C3	
			C4	
		D	D1	Asian–American (AA)
			D2	
			D3	
			D4	

Πίνακας 2. Ταξινόμηση των παραλλαγών του HPV16 (R.D. Burk et al., 2013).

Η γενεαλογία A ονομάστηκε Ευρωπαϊκή (E), η B Αφρικάνικη-1 (Af-1), η C Αφρικάνικη-2 (Af-2) και η D Ασιατική-Αμερικάνικη (AA). Μία παρατήρηση σχετικά με τις παραλλαγές του HPV16 είναι η παρουσία γενετικής ετερογένειας στις μη ευρωπαϊκές γενεαλογίες σε αντίθεση με τη σχετική ομοιογένεια της ευρωπαϊκής γενεαλογίας. Οι περισσότερες μελέτες εμπλέκουν τις μη Ευρωπαϊκές γενεαλογίες (B / C / D) ως πιο παθογόνες σε σύγκριση με αυτές της Ευρωπαϊκής (A) (R.D. Burk et al., 2013).

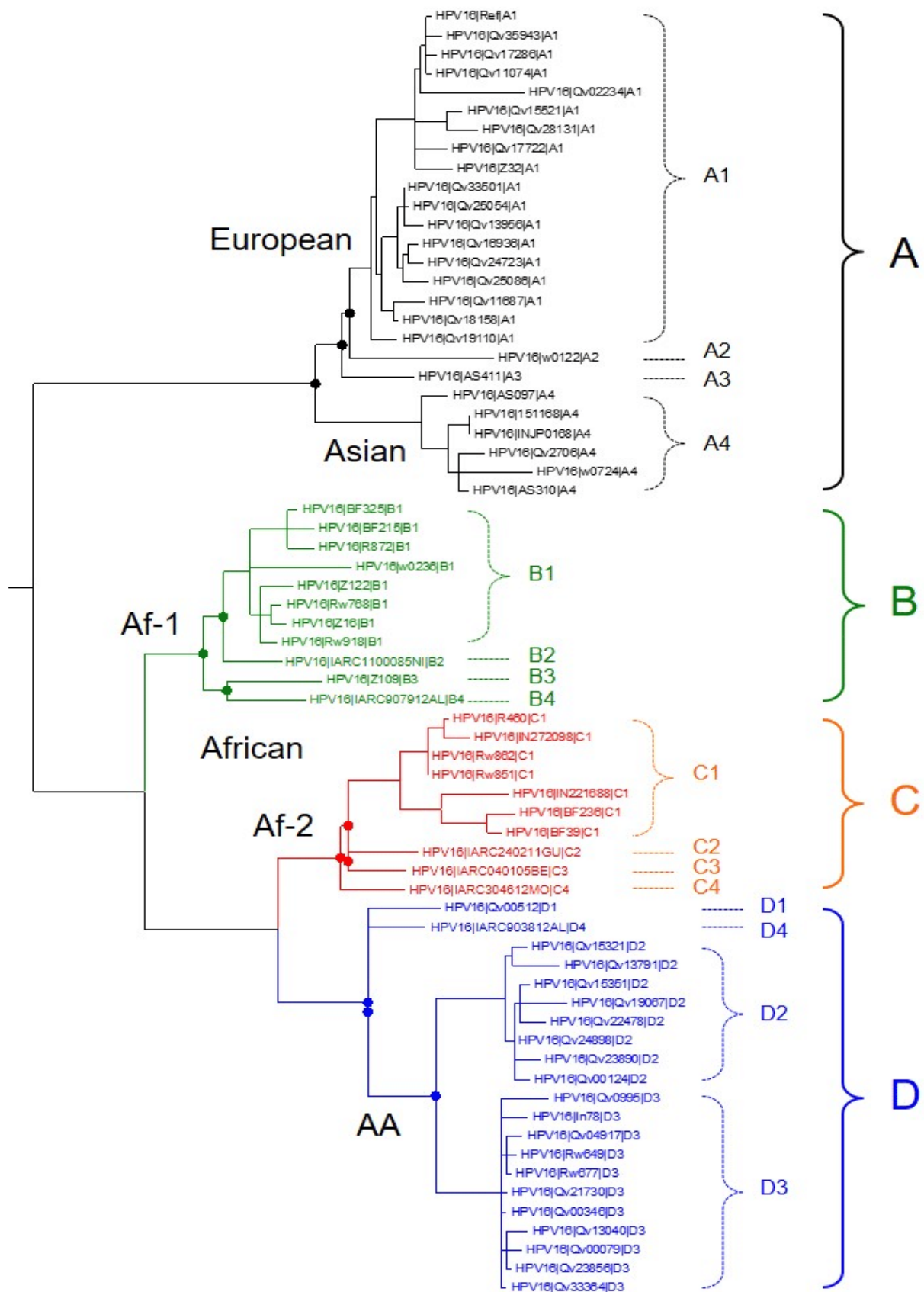
Μελέτες έδειξαν ότι, μεταξύ της γενεαλογίας A, οι A3 και A4 υπογενεαλογίες έχουν αυξημένο κίνδυνο καρκίνου έναντι της A1, η οποία είναι η πιο διαδεδομένη, και μεταξύ των γενεαλογιών B, C και D, μόνο η D σχετίζεται με υψηλότερο κίνδυνο καρκίνου. Πιο συγκεκριμένα, για τις υπογενεαλογίες D2, D3 αυτός ο κίνδυνος φαίνεται να είναι ιδιαίτερα αυξημένος για το αδενοκαρκίωμα (ADC) (G.M. Clifford, et al., 2019). Οι όγκοι της D2 εμφανίζουν υψηλότερο ρυθμό ενσωμάτωσης, έχουν πρώιμη ηλικία διάγνωσης και χαμηλότερο ποσοστό σωματικής μετάλλαξης, ενώ οι αντίστοιχοι της D3 είναι λιγότερο πιθανό να ενσωματωθούν, έχουν μεταγενέστερη ηλικία διάγνωσης και εμφανίζουν υψηλότερο ποσοστό σωματικής μετάλλαξης (Lou et al., 2020). Ωστόσο, όλες οι υπογενεαλογίες του HPV16 εντοπίζονται σε καρκίνο του τραχήλου της μήτρας σε όλο τον κόσμο και αναμένεται να είναι πιο καρκινογόνες από οποιονδήποτε άλλο τύπο HR-HPV (G.M. Clifford, et al., 2019).



Εικόνα 8. Κατανομή υπογενεαλογιών σε θετικά δείγματα HPV16, ανά γεωγραφική περιοχή (G.M. Clifford, et al., 2019).

Το ογκογόνο δυναμικό των HPVs οφείλεται σε μεγάλο βαθμό στις πρωτεΐνες E6 και E7, οι οποίες είναι απαραίτητες για τον κύκλο ζωής του ιού. Κατά τη διάρκεια μόλυνσης από HPV, παράγονται αρκετά αντίγραφα που περιλαμβάνουν ένα μετάγραφο που παράγει την πρωτεΐνη E6 και ένα ματισμένο μετάγραφο E6 που δημιουργεί την πρωτεΐνη E7. Πρόσφατες μελέτες δείχνουν υψηλότερα ποσοστά έκφρασης του ματισμένου μεταγράφου κατά την ενσωμάτωση του HPV16 και ότι η αλληλουχία E7 είναι συντηρημένη σε μεγάλο βαθμό σε καρκινικά κύτταρα. Η αυστηρή συντήρηση των 98 αμινοξέων της E7, που διαταράσσει τη λειτουργία της πρωτεΐνης Rb, είναι κρίσιμη για την καρκινογένεση. Η ανάλυση των περιπτώσεων καρκίνου του τραχήλου της μήτρας από όλο τον κόσμο δείχνει ότι η γενετική μεταβολή του γονιδίου E7 μειώνει σημαντικά τον κίνδυνο καρκίνου (Mirabello et al., 2018).

Η ενσωμάτωση του DNA του HPV συμβαίνει σε πολλές θέσεις στο ανθρώπινο γονιδίωμα, με προτίμηση κυρίως στις μεταγραφόμενες περιοχές. Για παράδειγμα, ενώ οι θέσεις ενσωμάτωσης HPV16 είναι ευρέως κατανομημένες στο ανθρώπινο γονιδίωμα, στην περίπτωση ενός όγκου, ο ιός ενσωματώθηκε στον υποκινητή του γονιδίου του ρυθμιστικού παράγοντα IFN 4, το οποίο παίζει σημαντικό ρόλο στη ρύθμιση της απόκρισης του IFN σε ιογενή λοίμωξη (Lou et al., 2020).



Εικόνα 9. Φυλογενωμικό δέντρο των γενεαλογιών και υπογενεαλογιών του HPV16 (Mirabello et al., 2018).

1.8. Εμβόλιο κατά των HPV

Η ανακάλυψη ότι το κύριο αντιγόνο L1 του καψιδίου των HPVs θα μπορούσε να αυτοσυναρμολογηθεί σε κενά σωματίδια που μοιάζουν με ιούς (VLPs) οδήγησε στην έγκριση αρκετών VLP εμβολίων HPV για την πρόληψη του καρκίνου του τραχήλου της μήτρας. Η πρόληψη της λοίμωξης από HR HPV είναι σημαντική και για τα δύο φύλα, καθώς ευθύνεται για τον καρκίνο του πρωκτού και του στοματοφάρυγγα στα δύο φύλα και του πέους στους άνδρες (Roden and Stern, 2019).

Τα σωματίδια που μοιάζουν με ιούς (VLPs) είναι κενά κελύφη ιού που προέρχονται από την έκφραση, σε κατάλληλο κύτταρο ξενιστή, δομικών πρωτεϊνών ιού όπως πρωτεΐνες καψιδίου ή επικάλυψης. Η υπερβολική έκφραση αυτών των πρωτεϊνών τους επιτρέπει να αυτοσυναρμολογούνται αυθόρμητα σε VLPs. Είναι μορφολογικά και δομικά παρόμοια με τους ιούς από τους οποίους προέρχονται οι πρωτεΐνες της επικάλυψης ή του καψιδίου, εκτός από το γεγονός ότι δεν διαθέτουν το γονιδίωμα του ιού. Είναι εξαιρετικά ανοσογόνα, ακόμη και σε μικρές δόσεις αντιγόνων (R. Yadav et al., 2019).

Τα τρία εγκεκριμένα εμβόλια - ένα δισθενές (2νHPV) εμβόλιο τεχνολογίας L1-VLP από τους τύπους HPV-16 και HPV-18 που διατίθενται στο εμπόριο ως Cervarix, ένα τετρασθενές (4νHPV) από τους HPV-6, HPV-11, HPV-16 και HPV-18 που διατίθενται στο εμπόριο ως Gardasil και ένα εννιάσθενές (9νHPV) από τους HPV-6, HPV-11, HPV-16, HPV-18, HPV-31, HPV-33, HPV-45, HPV-52 και HPV-58 που διατίθεται ως Gardasil 9 - παράγουν 2 έως 3 φορές περισσότερα αντισώματα από εκείνα που παρατηρήθηκαν σε φυσική λοίμωξη. Έρευνες αποδεικνύουν πως η προστασία που παρέχεται από τα παραπάνω εμβόλια διαρκεί για τουλάχιστον μια δεκαετία. Παρά την ισχυρή ανοσοποίηση που προσφέρουν, η προστασία από εμβόλια L1-VLP περιορίζεται σε συγκεκριμένους τύπους HPV και προσφέρουν ελάχιστη διασταυρούμενη προστασία έναντι άλλων τύπων HPV. Πιθανότατα για να επιτευχθεί ευρεία προστασία, ένα εμβόλιο HPV θα πρέπει να περιέχει L1 VLPs από διάφορους βασικούς τύπους του ιού (Roden and Stern, 2019).

Ο εμβολιασμός με L1-VLP δεν παρέχει θεραπευτικό όφελος. Η ανοσοποίηση προκαλεί μια ισχυρή απόκριση των CD8+ T κυττάρων, αλλά στα επιφανειακά στρώματα διαφοροποιημένων κερατινοκυττάρων όπου εντοπίζεται ο ιός δεν εκφράζεται η πρωτεΐνη L1 και επομένως ξεφεύγουν από αυτήν την απόκριση (Roden and Stern, 2019).

Η πρωτεΐνη L1 δεν είναι συντηρημένη μεταξύ των διαφόρων τύπων HPV. Από την άλλη πλευρά, η δευτερεύουσα πρωτεΐνη καψιδίου L2, διατηρείται σε μεγάλο βαθμό μεταξύ των τύπων HPV και υπήρξε ένα εναλλακτικό αντιγόνο-στόχος για την ανάπτυξη ενός ευρέως προστατευτικού εμβολίου HPV. Ωστόσο, η L2 σε αντίθεση με την L1 δεν μπορεί να σχηματίσει VLPs και ως εκ τούτου είναι λιγότερο ανοσογόνα. Οι τρέχουσες μελέτες

στοχεύουν στην ανάπτυξη εμβολίων HPV χρησιμοποιώντας L2 πεπτίδια σε VLPs που προέρχονται από βακτηριοφάγους και ευκαρυωτικούς ιούς. Πρόσφατα δεδομένα δείχνουν ότι ένα μονοσθενές L1-VLP καθώς και δισθενή VLP από βακτηριοφάγο MS2 με L2 πεπτίδια από HPV προκαλούν ισχυρά προστατευτικά αντισώματα έναντι διαφορετικών τύπων HPV που σχετίζονται με καρκίνους και κονδυλώματα των γεννητικών οργάνων. Τέτοιες ενδείξεις καθιστούν τα εμβόλια L2-VLP αρκετά υποσχόμενα και περισσότερο ευνοϊκά σε σχέση με τα τρέχοντα εμβόλια HPV που βασίζονται στην πρωτεΐνη L1 (R. Yadav et al., 2019).

Ο σκοπός της παρούσας διπλωματικής εργασίας είναι να μελετήσει τα ελεύθερα διαθέσιμα γονιδιώματα HPV16 για ύπαρξη ανασυνδυασμών και να υπολογίσει αυτή την συχνότητα όπως επίσης και να εντοπίσει τις θέσεις που συμβαίνουν οι ανασυνδυασμοί. Έτσι, θα μπορέσουμε να εκτιμήσουμε σε ποιο βαθμό οι ανασυνδυασμοί μπορεί να δημιουργήσουν προβλήματα σε εξελικτικές αναλύσεις και εάν μπορεί να αποτελέσουν πρόβλημα για τα υπάρχοντα εμβόλια.

2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

2.1. Λογισμικό

2.1.1 Linux Ubuntu 20

Το Linux είναι ένα λειτουργικό σύστημα ανοιχτού κώδικα και δωρεάν λογισμικού για τους χρήστες του. Το Ubuntu είναι μια έκδοση του λειτουργικού συστήματος (διανομή) που έχει τον πυρήνα του Linux ως βάση του. Βασίζεται στο λογισμικό Debian και το προεπιλεγμένο σύστημα γραφικών του είναι το GNOME. Είναι το ασφαλέστερο λειτουργικό σύστημα και ενδείκνυται για χρήση μεγάλου όγκου δεδομένων κάτι που το καθιστά ιδιαίτερα ανταγωνιστικό σε σχέση με το λειτουργικό σύστημα Windows και Mac.

2.1.2 Python 3.6

Η Python είναι μια αντικειμενοστρεφής γλώσσα προγραμματισμού υψηλού επιπέδου. Υποστηρίζει πολλαπλά πρότυπα προγραμματισμού, συμπεριλαμβανομένου δομημένου και λειτουργικού προγραμματισμού. Οι ενσωματωμένες βάσεις δεδομένων υψηλού επιπέδου σε συνδυασμό με τη δυναμική πληκτρολόγηση το καθιστούν πολύ ελκυστικό για την ταχεία ανάπτυξη εφαρμογών.

2.2. Λήψη γονιδιωμάτων

Για την ανάκτηση των γονιδιωματικών αλληλουχιών του HPV16, το γονιδίωμα αναφοράς του HPV16 με αριθμό καταχώρησης NC_001526 χρησιμοποιήθηκε ως ακολουθία επερώτησης στον αλγόριθμο NCBI BlastN (Camacho et al., 2009) έναντι της βάσης δεδομένων NCBI nr τον Φεβρουάριο του 2021. Έπειτα θέσαμε ως όριο το $1e^{-5}$ για την τιμή e-value, το ποσοστό νουκλεοτιδικής ταυτότητας μεγαλύτερο από 97%, και κρατήσαμε τις αλληλουχίες με μήκος τουλάχιστον 7500 νουκλεοτιδίων, όπου το ποσοστό των άγνωστων νουκλεοτιδίων σε ολόκληρο το γονιδίωμα ήταν λιγότερο από 1%. Αυτή η αναζήτηση έδωσε ως αποτέλεσμα 1534 ακολουθίες. Τα όρια των γονιδίων αυτών των 1534 αλληλουχιών προσδιορίστηκαν με BlastN έναντι των χαρακτηρισμένων κωδικών περιοχών του γονιδιώματος αναφοράς του HPV16 με αριθμό NC_001526.

2.2.1 Ομαδοποίηση γονιδιωμάτων

Προκειμένου να απορρίψουμε τις περιττές γονιδιωματικές αλληλουχίες με υψηλά ποσοστά νουκλεοτιδικής ταύτισης, τα 1534 γονιδιώματα ομαδοποιήθηκαν με τον αλγόριθμο Uclust του Usearch (Edgar, 2010) θέτοντας ως κατώφλι νουκλεοτιδικής ταύτισης το 99,6% στο 99,6% του μήκους επερώτησης. Αυτή η ανάλυση έδωσε ως αποτέλεσμα 193 ομάδες. Για κάθε ομάδα επιλέχθηκε μια αλληλουχία αναφοράς/αντιπρόσωπος, ενώ διατηρήσαμε κάθε ένα από τα 16 γονιδιώματα αναφοράς (A1-A4, B1-B4, C1-C4, D1-D4) που περιγράφονται στο Mirabello et al., 2018. Με βάση αυτό το σύνολο των 193 γονιδιωμάτων, καθένα από τα γονίδια του HPV-16 (E6, E7, E1, E2, E4, E5, L2, L1) εξήχθησαν και στοιχήθηκαν ως κωδικόνια χρησιμοποιώντας το MUSCLE (Edgar, 2004) το οποίο είναι ενσωματωμένο στο πρόγραμμα Seaview (Gouy et al., 2010). Η χειροκίνητη επιθεώρηση των στοιχίσεων αποκάλυψε 13 αλληλουχίες με πολλά άγνωστα νουκλεοτίδια στο γονίδιο E5 και κατά συνέπεια αφαιρέθηκαν από ολόκληρη την ανάλυση, με αποτέλεσμα να προκύψουν 180 συνολικά αντιπροσωπευτικές γονιδιωματικές αλληλουχίες. Έπειτα, πραγματοποιήσαμε στοιχίσεις ολόκληρου γονιδιώματος των 180 αντιπροσωπευτικών γονιδιωμάτων με το mafft v7.471 (παραμέτροι: G-INSI) (Katoh, 2002).

2.2.2 Υπολογισμός δέντρων

Υπολογίσαμε τα δέντρα Neighbour Joining για την πολλαπλή στοίχιση του πλήρους γονιδιώματος και για κάθε γονίδιο ξεχωριστά, με το μοντέλο Poisson, χρησιμοποιώντας 500 bootstrap επαναλήψεις, που είναι ενσωματωμένα στο πρόγραμμα SeaView (Gouy et al., 2010). Τα φυλογενετικά δέντρα οπτικοποιήθηκαν και χαρακτηρίστηκαν στο Treedyn (Chevenet et al., 2006) και ελέγχθηκαν χειροκίνητα για μη συμβατές αλληλουχίες.

3. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ

Στη φυλογενωμική ανάλυση πλήρους γονιδιώματος (**Εικόνα 10**) εντοπίσαμε δύο κύριες ομάδες όπως έχουν αναφερθεί προηγουμένως στην ανάλυση των Mirabello et al., 2018. Η πρώτη ομάδα, η A (γνωστή και ως Ευρωπαϊκή) αποτελείται από τις υπογενεαλογίες A1-4 (bootstrap 100), ενώ η άλλη ομάδα περιλαμβάνει τις γενεαλογίες B, C και D (bootstrap 100) και αναφέρεται ως BCD.

Με εξαίρεση πέντε αναγνωρισμένες ανασυνδυασμένες αλληλουχίες, σε όλα τα μεμονωμένα γονίδια βλέπουμε μια συνολική ομαδοποίηση, παρόμοια με το δέντρο πλήρους γονιδιώματος, όπου η Ευρωπαϊκή ομάδα A ξεχωρίζει με υψηλές τιμές bootstrap (72 - 100) με εξαίρεση τα δέντρα E4, E5 και E7, που είναι μικρού μήκους αλληλουχίες και αποτελεί έναν κλάδο, ενώ οι άλλες τρεις γενεαλογίες σχηματίζουν μία δεύτερη μεγάλη ομάδα. Οι γενεαλογίες B, C και D δεν εμφανίζονται ταυτόχρονα ως μονοφυλετικές σε κάποιο δέντρο. Η γενεαλογία B είναι μονοφυλετική στα γονίδια L2 (bootstrap 95; **Εικόνα 17**) και L1 (bootstrap 79; **Εικόνα 18**), ενώ η C είναι μονοφυλετική στο γονίδιο L2 (bootstrap 97; **Εικόνα 17**). Επιπλέον, οι ομάδες C και D σχηματίζουν μία μεγαλύτερη ομάδα στα E1, E2 και L1 και είναι αδελφές ομάδες στο E2. Είναι ενδιαφέρον ότι οι γενεαλογίες C και D διαπλέκονται στο γονίδιο E5, το οποίο υποδηλώνει είτε την παρουσία πολλών γεγονότων ανασυνδυασμού μεταξύ αυτών των δύο γενεαλογιών, είτε ότι το φυλογενετικό σήμα δεν είναι αρκετά ισχυρό για να τα αναδείξει ως μονοφυλετικές ομάδες. Η φυλογενετική ανάλυση πραγματοποιήθηκε για τον εντοπισμό φυλογενετικά ασύμβατων αλληλουχιών που αποτελούν σαφή σημάδια ανασυνδυασμού.

Η ανάλυσή μας εντόπισε τουλάχιστον 5 συμβάντα ανασυνδυασμού. Το πρώτο συμβάν είναι η αλληλουχία με τον αριθμό καταχώρησης MT316214, που ανακτήθηκε από δείγματα καρκίνου του τραχήλου της μήτρας από τη Γουατεμάλα, όπως περιγράφεται στο Lou et al., 2020. Η αλληλουχία σε αυτή τη μελέτη πραγματοποιήθηκε με το Ion Torrent. Με βάση τη φυλογενετική ανάλυση η αλληλουχία αυτή είναι ένας ανασυνδυασμός της ευρωπαϊκής γενεαλογίας (A1-4) που έχει αποκτήσει το μεγαλύτερο μέρος του γονιδίου E1 και ένα τμήμα του E2 από την ομάδα BCD και πιο συγκεκριμένα από τη D γενεαλογία.

Η δεύτερη ανασυνδυασμένη αλληλουχία είναι η KU641509, η οποία απομονώθηκε από μια βιοψία καρκινώματος τραχήλου της μήτρας στην Ινδία (Mandal et al., 2016). Είναι μέλος της υπογενεαλογίας D1 που έχει αποκτήσει ένα μικρό μέρος της περιοχής E1 από άλλο άγνωστο μέλος της ευρύτερης ομάδας BCD.

Η τρίτη ανασυνδυασμένη αλληλουχία είναι η MT316301, που απομονώθηκε από δείγματα καρκίνου του τραχήλου της μήτρας (Lou et al., 2020). Είναι μέλος της υπογενεαλογίας D3, που έχει παραλάβει μέρος του E2 γονιδίου της από την Ευρωπαϊκή γενεαλογία.

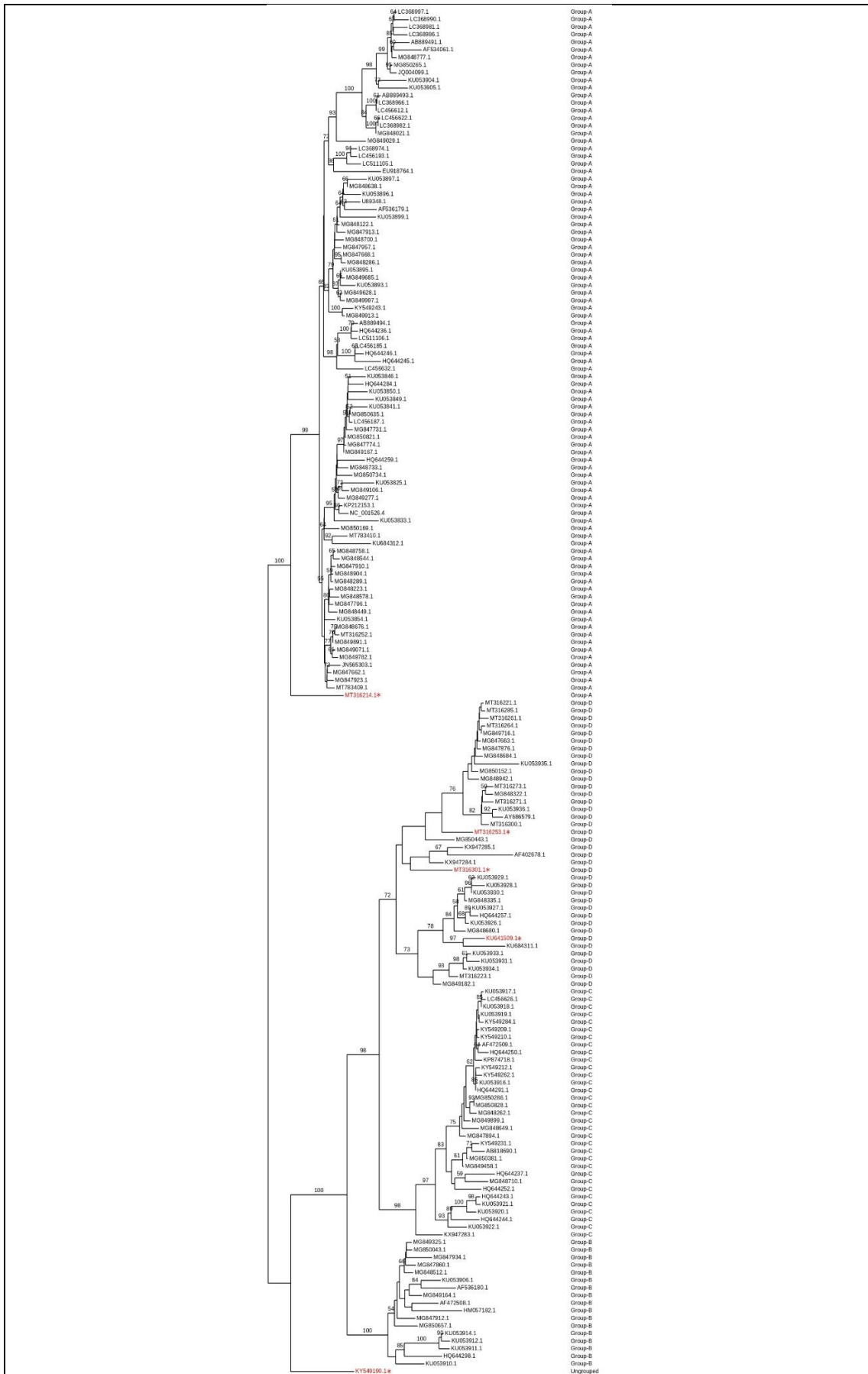
Παρομοίως, η MT316253 είναι η τέταρτη ανασυνδυασμένη αλληλουχία, που λήφθηκε από τους Lou et al., 2020. Είναι μέλος της γενεαλογίας D (υπογενεαλογία D2) που έχει αποκτήσει μέρος του E2 γονιδίου της από την Ευρωπαϊκή γενεαλογία.

Το τελευταίο συμβάν ανασυνδυασμού είναι το KY549190, το οποίο απομονώθηκε στην Ολλανδία και αλληλουχήθηκε με τη μέθοδο αλληλούχισης ολόκληρου γονιδιώματος Sanger (vander Weele et al., 2017). Οι διάφορες αναλύσεις αποκάλυψαν ότι σχεδόν το ήμισυ αυτής της αλληλουχίας (από το δεύτερο μισό του E1 έως το πρώτο μισό L1) ανήκει στην Ευρωπαϊκή γενεαλογία, ενώ το άλλο μέρος ανήκει στη γενεαλογία BCD και πιο συγκεκριμένα στη γενεαλογία C.

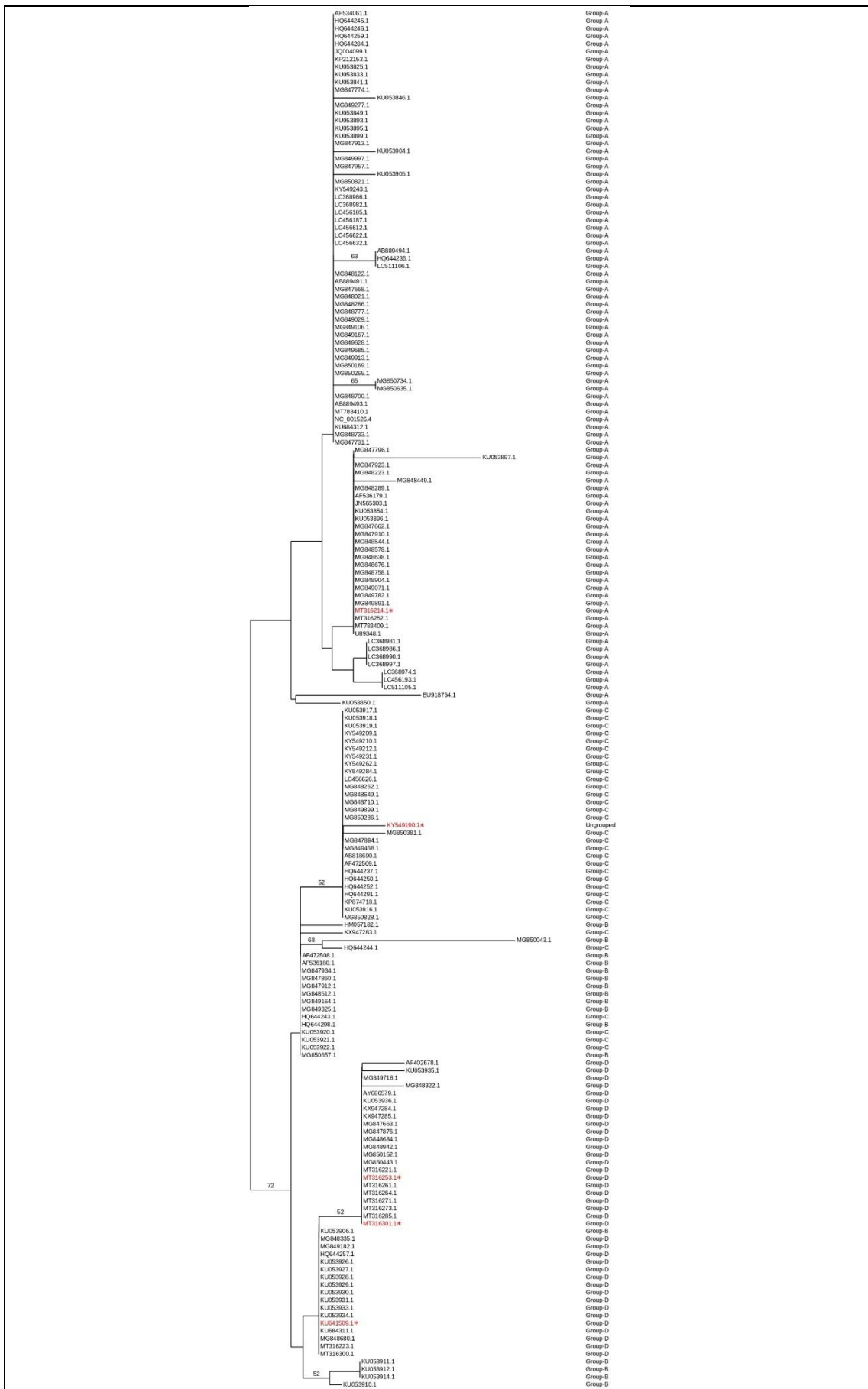
Είναι κατανοητό ότι μερικές από τις αλληλουχίες διαφορετικών γενεαλογιών που έχουν αναμειχθεί στις φυλογενέσεις των E6, E7 και E5 θα μπορούσαν να είναι επιπρόσθετα συμβάντα ανασυνδυασμού. Ωστόσο, αυτές οι γονιδιωματικές περιοχές έχουν μικρά μεγέθη και η συνολική ταύτιση αλληλουχίας μεταξύ των μελών των διαφορετικών σειρών του HPV16 είναι υψηλή.

Οι ταυτοποιημένες ανασυνδυασμένες αλληλουχίες είναι μεμονωμένες ομάδες, έτσι αντιπροσωπεύουν 5 συμβάντα ανασυνδυασμού σε περισσότερες από 1500 αλληλουχίες και δεν φαίνεται να κυκλοφορούν στον γενικό πληθυσμό. Είναι πιθανό ότι αυτοί οι ανασυνδυασμοί είναι εξελικτικά αδιέξοδα υποπροϊόντα χαμηλής συχνότητας του κύκλου μόλυνσης του HPV16.

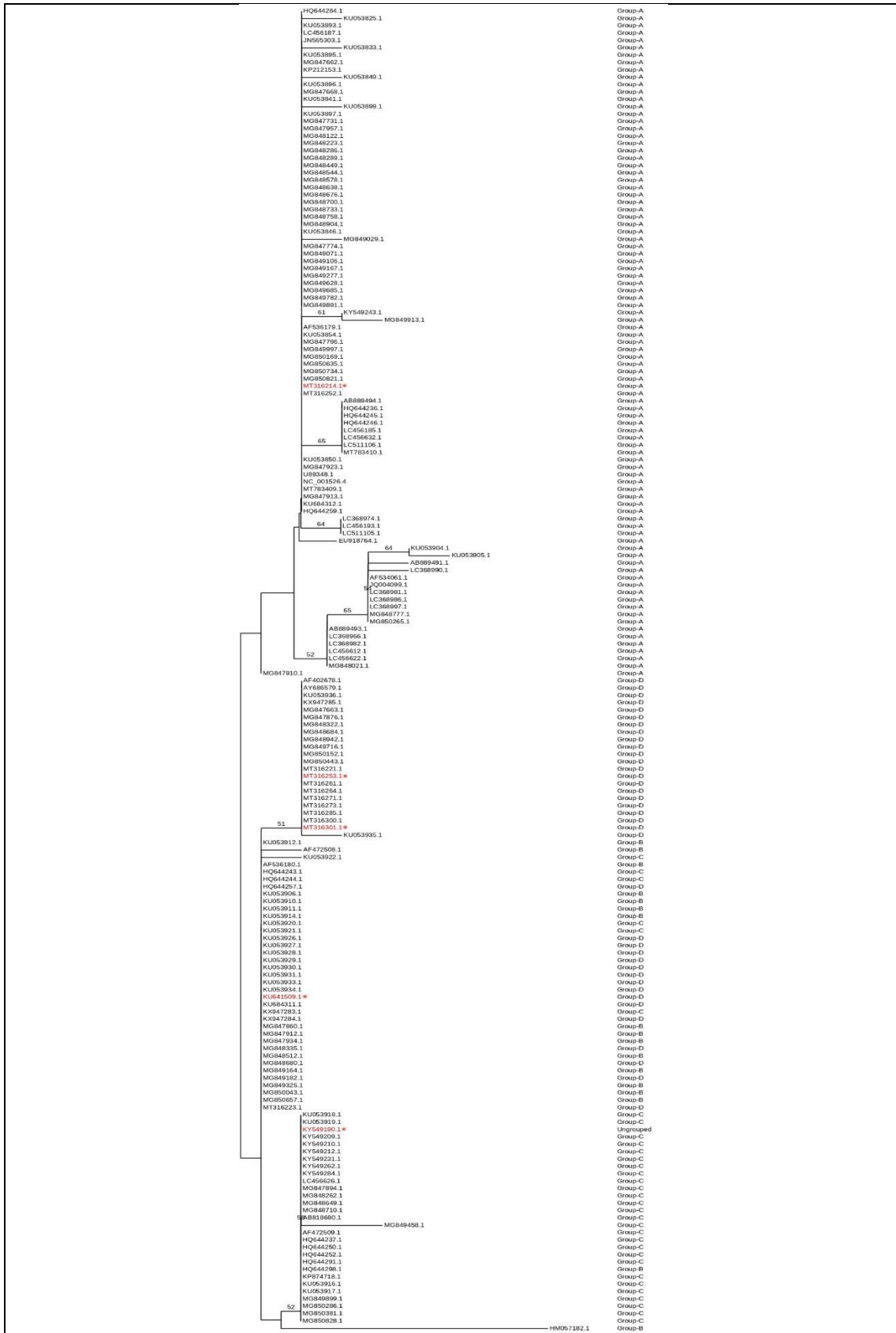
Συνολικά, εντοπίσαμε 4 θέσεις ανασυνδυασμού στην E1, πέντε στην E2 και μία στην L1. Επομένως, τα υπάρχοντα δεδομένα υποστηρίζουν την γενικότερη εικόνα που επικρατεί στην βιβλιογραφία ότι οι HPVs ανασυνδυάζονται με χαμηλή συχνότητα, την οποία υπολογίσαμε στο 0.3%. Ωστόσο, η ανάλυση αυτή βασίσθηκε σε πλήρη γονιδιώματα και σε φυλογενετικές αναλύσεις. Είναι πολύ πιθανόν ότι η ανάλυση περισσοτέρων ακολουθιών με πιο ευαίσθητες μεθόδους θα εντοπίσει πολύ περισσότερα γεγονότα. Επιπλέον, τα υπάρχοντα δεδομένα δεν εντόπισαν σημαντικό αριθμό ανασυνδυασμών στην L1, επομένως, δεν αναμένεται να αποτελέσουν σημαντικό πρόβλημα για τα υπάρχοντα εμβόλια.



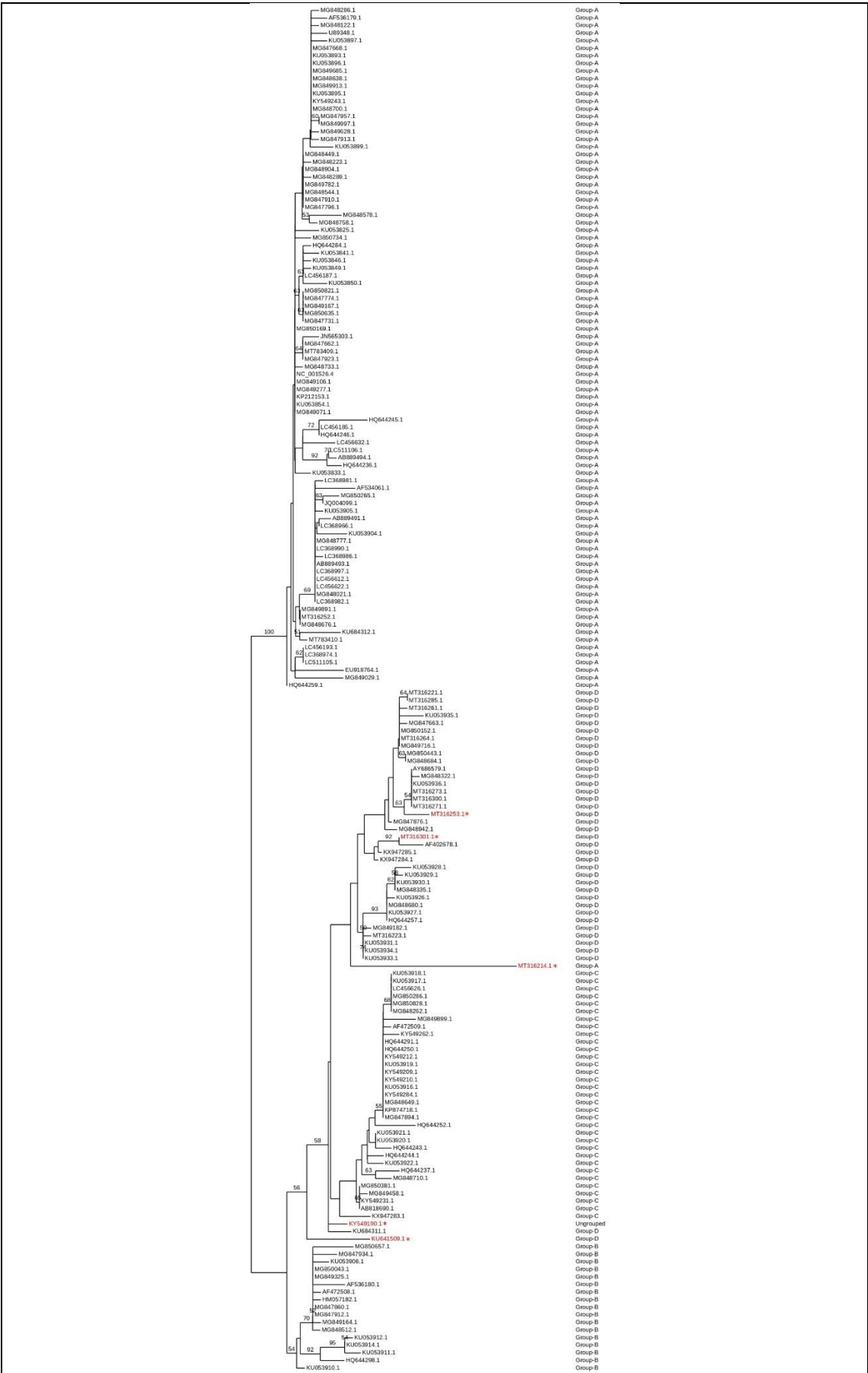
Εικόνα 10. Φυλογενωμικό δέντρο πλήρους γονιδιώματος. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



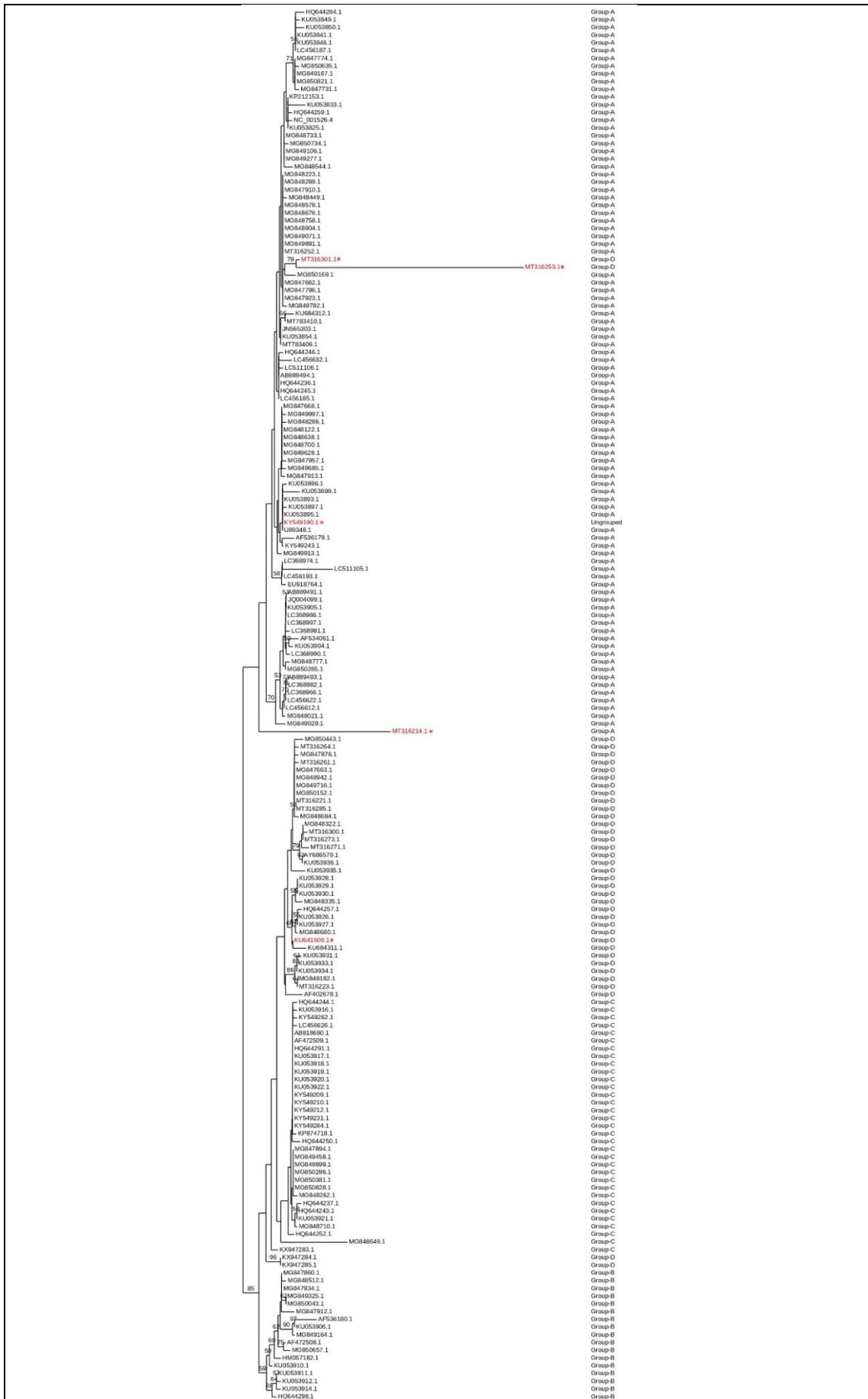
Εικόνα 11. Φυλογενετικό δέντρο γονιδίου E6. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



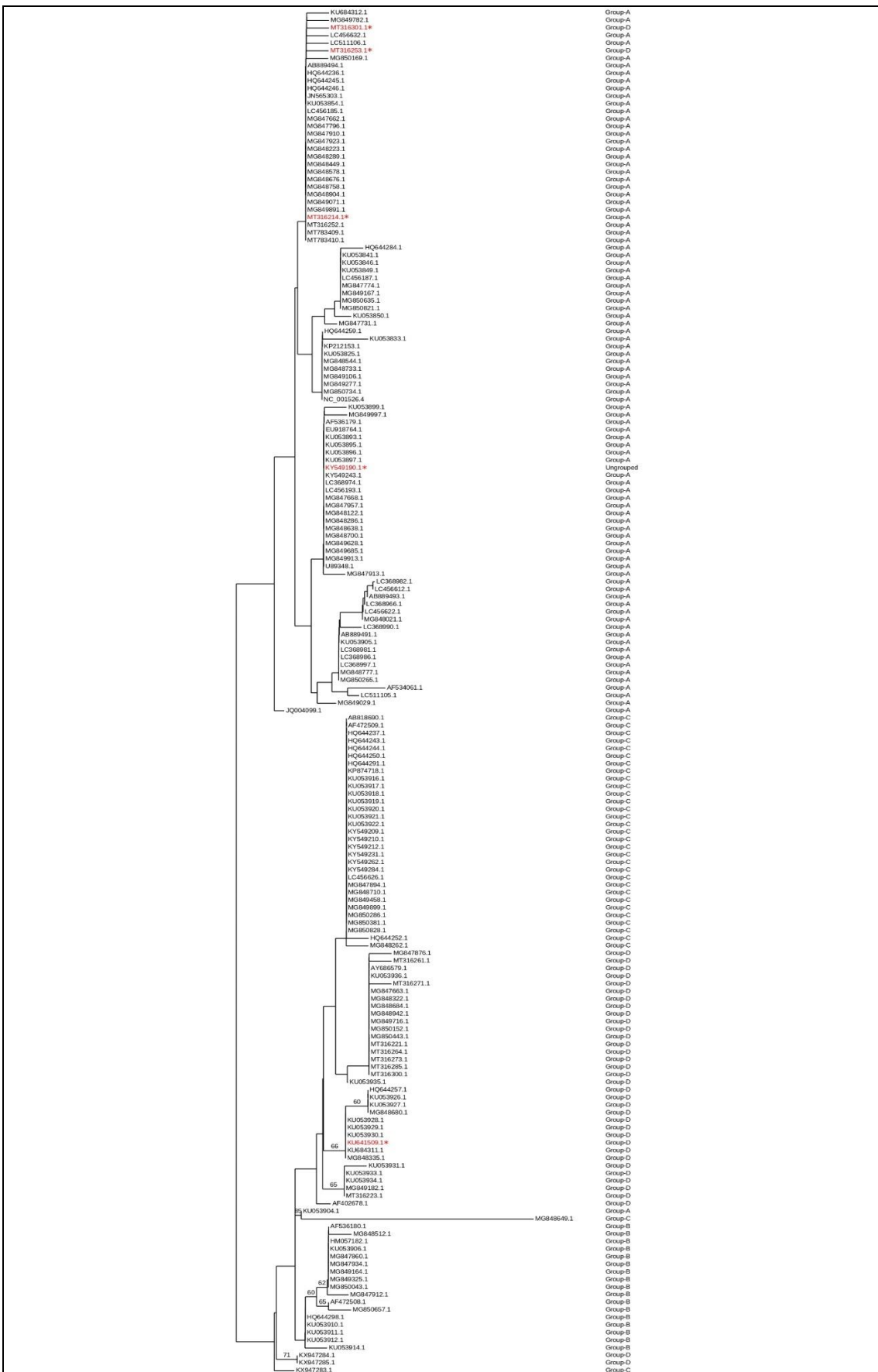
Εικόνα 12. Φυλογενετικό δέντρο γονιδίου E7. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



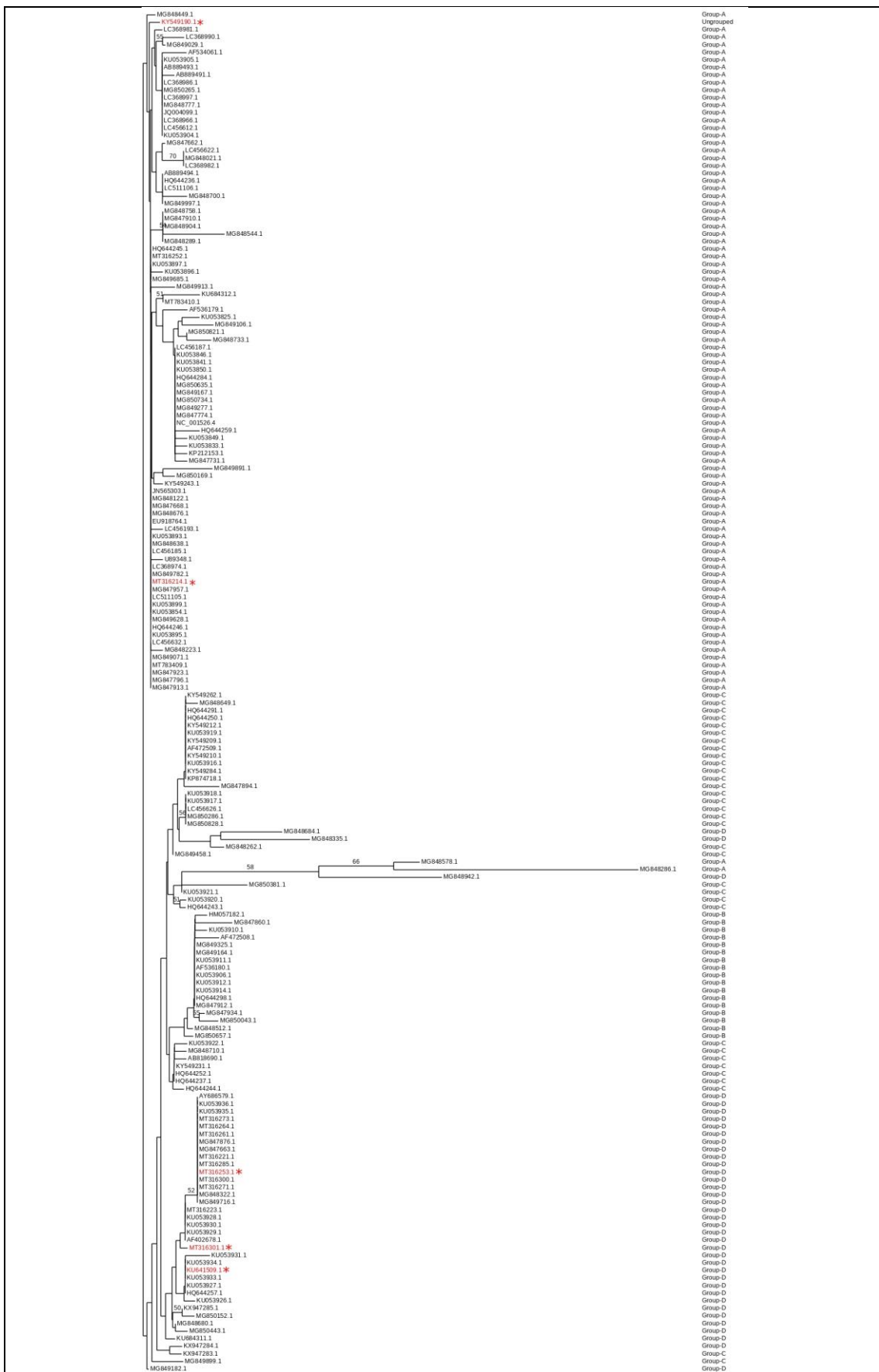
Εικόνα 13. Φυλογενετικό δέντρο γονιδίου E1. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



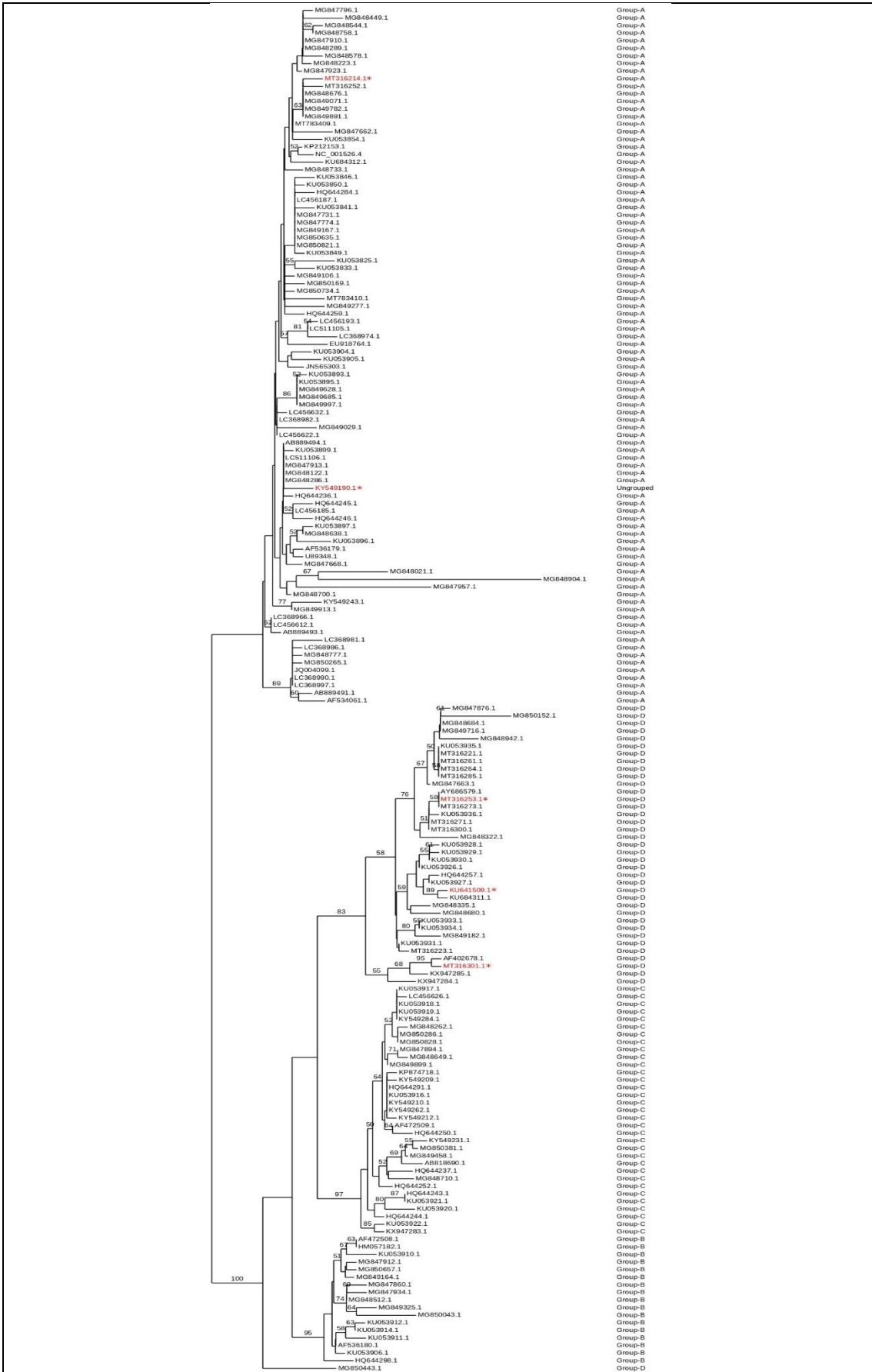
Εικόνα 14. Φυλογενετικό δέντρο γονιδίου E2. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



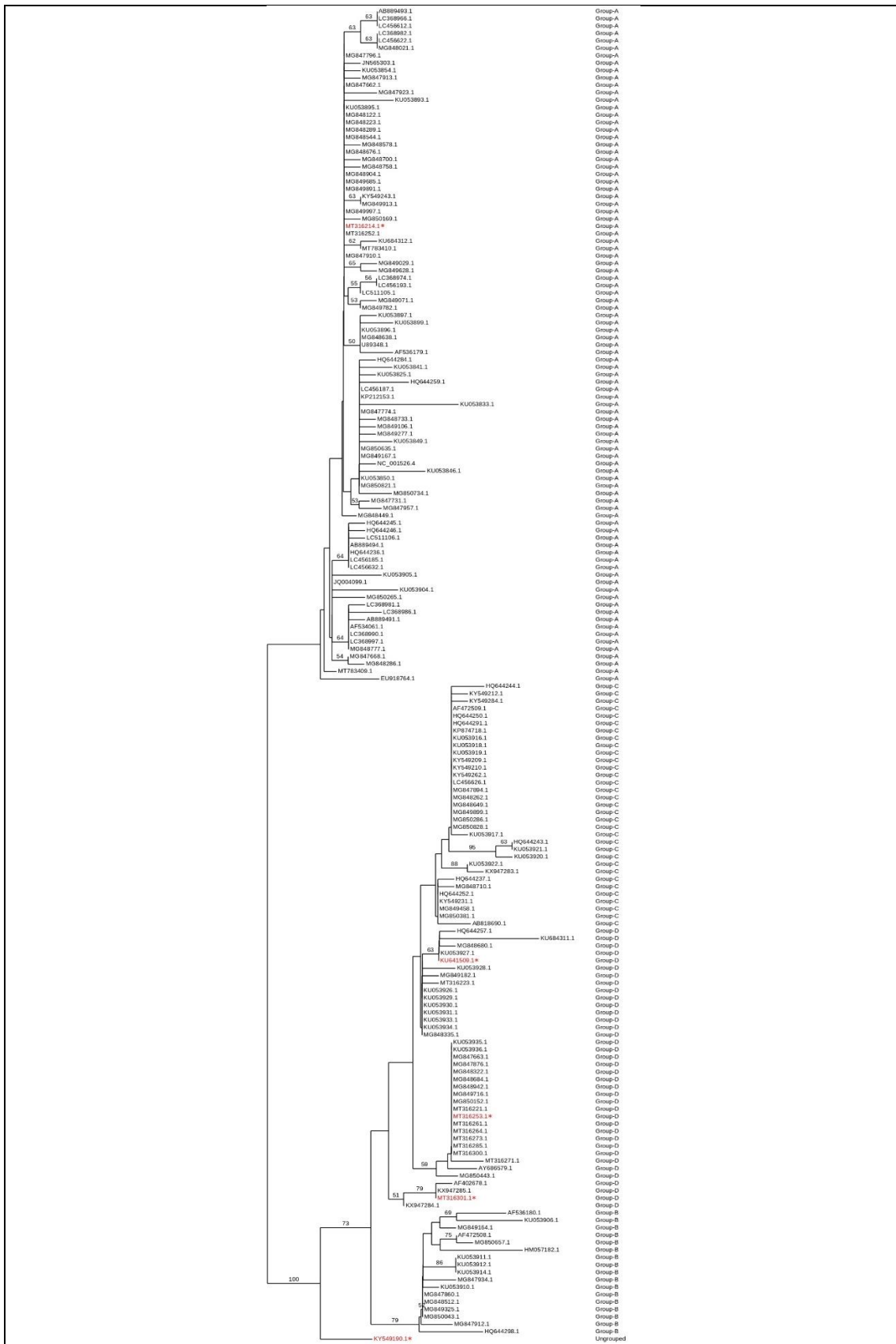
Εικόνα 15. Φυλογενετικό δέντρο γονιδίου E4. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



Εικόνα 16. Φυλογενετικό δέντρο γονιδίου E5. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



Εικόνα 17. Φυλογενετικό δέντρο γονιδίου L2. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.



Εικόνα 18. Φυλογενετικό δέντρο γονιδίου L1. Με κόκκινο χρώμα εμφανίζονται οι 5 αλληλουχίες που θεωρούνται ανασυνδυασμένες.

4. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Chen, Z., Schiffman, M., Herrero, R., DeSalle, R., Anastos, K., Segondy, M., Sahasrabudde, V.V., Gravitt, P.E., Hsing, A.W., Chan, P.K.S., Burk, R.D., 2018. Classification and evolution of human papillomavirus genome variants: Alpha-5 (HPV26, 51, 69, 82), Alpha-6 (HPV30, 53, 56, 66), Alpha-11 (HPV34, 73), Alpha-13 (HPV54) and Alpha-3 (HPV61). *Virology* 516, 86–101. <https://doi.org/10.1016/j.virol.2018.01.002>
- de Oliveira, C.M., 2019. Adaptation of Alpha-Papillomavirus over Millennia. *Acta Cytol* 63, 97–99. <https://doi.org/10.1159/000492658>
- McBride, A.A., 2017. Oncogenic human papillomaviruses. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372, 20160273. <https://doi.org/10.1098/rstb.2016.0273>
- Murahwa, A.T., Nindo, F., Onywera, H., Meiring, T.L., Martin, D.P., Williamson, A.-L., 2019. Evolutionary dynamics of ten novel Gamma-PVs: insights from phylogenetic incongruence, recombination and phylodynamic analyses. *BMC Genomics* 20, 368. <https://doi.org/10.1186/s12864-019-5735-9>
- Van Doorslaer, K., Chen, Z., Bernard, H.-U., Chan, P.K.S., DeSalle, R., Dillner, J., Forslund, O., Haga, T., McBride, A.A., Villa, L.L., Burk, R.D., ICTV Report Consortium, 2018. ICTV Virus Taxonomy Profile: Papillomaviridae. *Journal of General Virology* 99, 989–990. <https://doi.org/10.1099/jgv.0.001105>
- Zhang, J., Zhang, Shaohong, Wang, M., Ding, X., Wen, Q., Chen, Z., Cao, M., Jing, Y., Zhang, Shun, 2018. Genetic variability in E5, E6, E7 and L1 genes of human papillomavirus type 31. *Mol Med Rep* 17, 5498–5507. <https://doi.org/10.3892/mmr.2018.8500>
- Zheng, Z.-M., Baker, C.C., 2006. PAPILOMAVIRUS GENOME STRUCTURE, EXPRESSION, AND POST-TRANSCRIPTIONAL REGULATION. *Front Biosci* 11, 2286–2302.
- Bolatti, E.M., Chouhy, D., Casal, P.E., Pérez, G.R., Stella, E.J., Sanchez, A., Gorosito, M., Bussy, R.F., Giri, A.A., 2016. Characterization of novel human papillomavirus types 157, 158 and 205 from healthy skin and recombination analysis in genus γ -Papillomavirus. *Infect Genet Evol* 42, 20–29. <https://doi.org/10.1016/j.meegid.2016.04.018>
- Borvetó, F., Bravo, I.G., Willemsen, A., 2020. Papillomaviruses infecting cetaceans exhibit signs of genome adaptation following a recombination event. *Virus Evol* 6, veaa038. <https://doi.org/10.1093/ve/veaa038>
- Jiang, M., Xi, L.F., Edelstein, Z.R., Galloway, D.A., Olsem, G.J., Lin, W.C.-C., Kiviat, N.B., 2009. Identification of recombinant human papillomavirus type 16 variants. *Virology* 394, 8–11. <https://doi.org/10.1016/j.virol.2009.08.040>
- Murahwa, A.T., Tshabalala, M., Williamson, A.-L., 2020. Recombination Between High-Risk Human Papillomaviruses and Non-Human Primate Papillomaviruses: Evidence of Ancient Host Switching Among Alphapapillomaviruses. *J Mol Evol* 88, 453–462. <https://doi.org/10.1007/s00239-020-09946-0>
- Pérez-Losada, M., Arenas, M., Galán, J.C., Palero, F., González-Candelas, F., 2015. Recombination in viruses: Mechanisms, methods of study, and

- evolutionary consequences. *Infection, Genetics and Evolution* 30, 296–307. <https://doi.org/10.1016/j.meegid.2014.12.022>
- Rector, A., Stevens, H., Lacave, G., Lemey, P., Mostmans, S., Salbany, A., Vos, M., Van Doorslaer, K., Ghim, S.-J., Rehtanz, M., Bossart, G.D., Jenson, A.B., Van Ranst, M., 2008. Genomic characterization of novel dolphin papillomaviruses provides indications for recombination within the Papillomaviridae. *Virology* 378, 151–161. <https://doi.org/10.1016/j.virol.2008.05.020>
 - Van Doorslaer, K., 2013. Evolution of the papillomaviridae. *Virology* 445, 11–20. <https://doi.org/10.1016/j.virol.2013.05.012>
 - Varsani, A., van der Walt, E., Heath, L., Rybicki, E.P., Williamson, A.L., Martin, D.P., 2006. Evidence of ancient papillomavirus recombination. *J Gen Virol* 87, 2527–2531. <https://doi.org/10.1099/vir.0.81917-0>
 - Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. *Virology* 445, 232–243. <https://doi.org/10.1016/j.virol.2013.07.018>
 - Clifford, G.M., Tenet, V., Georges, D., Alemany, L., Pavón, M.A., Chen, Z., Yeager, M., Cullen, M., Boland, J.F., Bass, S., Steinberg, M., Raine-Bennett, T., Lorey, T., Wentzensen, N., Walker, J., Zuna, R., Schiffman, M., Mirabello, L., 2019. Human papillomavirus 16 sub-lineage dispersal and cervical cancer risk worldwide: Whole viral genome sequences from 7116 HPV16-positive women. *Papillomavirus Res* 7, 67–74. <https://doi.org/10.1016/j.pvr.2019.02.001>
 - Lou, H., Boland, J.F., Torres-Gonzalez, E., Albanez, A., Zhou, W., Steinberg, M.K., Diaw, L., Mitchell, J., Roberson, D., Cullen, M., Garland, L., Bass, S., Burk, R.D., Yeager, M., Wentzensen, N., Schiffman, M., Freites, E.A., Gharzouzi, E., Mirabello, L., Dean, M., 2020. The D2 and D3 Sublineages of Human Papilloma Virus 16-Positive Cervical Cancer in Guatemala Differ in Integration Rate and Age of Diagnosis. *Cancer Res* 80, 3803–3809. <https://doi.org/10.1158/0008-5472.CAN-20-0029>
 - Mirabello, L., Yeager, M., Yu, K., Clifford, G.M., Xiao, Y., Zhu, B., Cullen, M., Boland, J.F., Wentzensen, N., Nelson, C.W., Raine-Bennett, T., Chen, Z., Bass, S., Song, L., Yang, Q., Steinberg, M., Burdett, L., Dean, M., Roberson, D., Mitchell, J., Lorey, T., Franceschi, S., Castle, P.E., Walker, J., Zuna, R., Kreimer, A.R., Beachler, D.C., Hildesheim, A., Gonzalez, P., Porras, C., Burk, R.D., Schiffman, M., 2017. HPV16 E7 Genetic Conservation Is Critical to Carcinogenesis. *Cell* 170, 1164–1174.e6. <https://doi.org/10.1016/j.cell.2017.08.001>
 - Roden, R.B.S., Stern, P.L., 2018. Opportunities and challenges for human papillomavirus vaccination in cancer. *Nat Rev Cancer* 18, 240–254. <https://doi.org/10.1038/nrc.2018.13>
 - Yadav, R., Zhai, L., Tumban, E., 2019. Virus-like Particle-Based L2 Vaccines against HPVs: Where Are We Today? *Viruses* 12. <https://doi.org/10.3390/v12010018>
 - Camacho et al., 2009, n.d.
 - Chevenet, F., Brun, C., Bañuls, A.-L., Jacq, B., Christen, R., 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7, 439. <https://doi.org/10.1186/1471-2105-7-439>

- Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Edgar, Robert C, 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113. <https://doi.org/10.1186/1471-2105-5-113>
- Edgar, Robert C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27, 221–224. <https://doi.org/10.1093/molbev/msp259>
- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059–3066.
- Mandal et al., 2016, n.d.
- Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1, vev003. <https://doi.org/10.1093/ve/vev003>
- Tsimpidis, M., Bachoumis, G., Mimouli, K., Kyriakopoulou, Z., Robertson, D.L., Markoulatos, P., Amoutzias, G.D., 2017. T-RECs: rapid and large-scale detection of recombination events among different evolutionary lineages of viral genomes. *BMC Bioinformatics* 18, 13. <https://doi.org/10.1186/s12859-016-1420-z>
- vander Weele et al., 2017, n.d.