



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ**



**ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ**

**ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ**

**ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΜΑΘΗΜΑΤΙΚΩΝ**

**Διευθυντής ΗΛΙΑΣ ΖΙΝΤΖΑΡΑΣ**

**Διδακτορική Διατριβή**

**«ΓΕΝΩΜΙΚΗ ΣΥΓΚΛΙΣΗ ΣΤΗ ΝΟΣΟ ΤΟΥ PARKINSON»**

υπό

**ΡΙΚΟΥ ΔΗΜΗΤΡΙΟΥ**

Νευρολόγος

Υπεβλήθη για την εκπλήρωση μέρους των

απαιτήσεων για την απόκτηση του

Διδακτορικού Διπλώματος

Λάρισα, 2022

© 2022 ΡΙΚΟΣ ΔΗΜΗΤΡΙΟΣ

Η έγκριση της διδακτορικής διατριβής από το Τμήμα Ιατρικής της Σχολής Επιστημών Υγείας του Πανεπιστημίου Θεσσαλίας δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα (σύμφωνα με τις διατάξεις του άρθρου 202, παράγραφος 2 του Ν.5343/1932).

**1<sup>ος</sup> Εξεταστής Ζιντζαράς Ηλίας**

**(Επιβλέπων)** Καθηγητής, Βιομαθηματικά, Εργαστήριο Βιομαθηματικών, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**2<sup>ος</sup> Εξεταστής Στεφανίδης Ιωάννης**

Καθηγητής, Νεφρολογία, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**3<sup>ος</sup> Εξεταστής Ευθύμιος Δαρδιώτης**

Αν. Καθηγητής, Νευρολογία, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**4<sup>ος</sup> Εξεταστής Κωνσταντίνος Νταφόπουλος**

Καθηγητής, Μαιευτική-Γυναικολογία, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**5<sup>ος</sup> Εξεταστής Δημήτριος Ζαχαρούλης**

Καθηγητής, Χειρουργική, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**6<sup>ος</sup> Εξεταστής Χριστίνα Μεσσήνη**

Επικ. Καθηγήτρια, Μαιευτική - Γυναικολογία, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας

**7<sup>ος</sup> Εξεταστής Ιωάννης Σκουλαρίγκης**

Καθηγητής, Καρδιολογία, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας



## ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να εκφράσω τις ειλικρινείς ευχαριστίες μου στους καθηγητές της τριμελούς επιτροπής για την υποστήριξη κατά την εκπόνηση της παρούσας. Ευχαριστώ τον καθηγητή κ. Ζιντζαρά Ηλία και τον καθηγητή κ. Χατζηγεωργίου Γεώργιο που με εμπιστεύτηκαν και μου έδωσαν την ευκαιρία να εκπονήσω την διδακτορική διατριβή και τον Αναπ. Καθηγητή κ. Δαρδιώτη Ευθύμιο ο οποίος στάθηκε δίπλα μου καθ' όλη την πορεία αυτή σαν δάσκαλος και συνοδοιπόρος.

Θα ήθελα επίσης να εκφράσω τις ευχαριστίες μου στην Αναπ. Καθηγήτρια κα. Ξηρομερήσιου Γεωργία που αφειδώς μοιράστηκε μαζί μου τις γνώσεις και την εμπειρία της. Ιδιαίτερες ευχαριστίες στην Δρ. Τζιαστούδη Μαρία για την πολύτιμη συμβολή της, τον ειδικό στον προγραμματισμό κ. Ανδρέου Άρη και τον αγαπημένο συνάδελφο και φίλο Δρ. Προβατά Αντώνιο για την τεράστια βοήθεια τους.

Τέλος ένα μεγάλο ευχαριστώ οφείλω στη σύζυγό μου και τα παιδιά μου για την υπομονή και την υποστήριξή τους.

Ρίκος Δημήτριος



## ΒΙΟΓΡΑΦΙΚΟ

Όνομα: Δημήτριος Ρίκος

Ημ/νια γέννησης: \*\*\*\*\*

Διεύθυνση: \*\*\*\*\*

Κινητό τηλ.: \*\*\*\*\*

E-mail: \*\*\*\*\*

**Αφαίρεση προσωπικών  
δεδομένων** (Υπηρεσία  
Βιβλιοθήκης & Πληροφόρησης  
Πανεπιστημίου Θεσσαλίας)

### ΕΚΠΑΙΔΕΥΣΗ:

2000-2006 Στρατιωτική Σχολή Αξιωματικών Σωμάτων – Ιατρικό Τμήμα –  
Ιατρική Σχολή ΑΠΘ

2010-2010 Diploma in Aerospace Medicine. Advanced Aerospace  
Medicine for International Medical Officers, USAF School of  
Aerospace Medicine, Brooks city base, San Antonio, Texas

2014-2015 MSc, Research Methodology in Biomedicine, Biostatistics and  
Clinical Bioinformatics at University of Thessaly.

Πτυχιακή εργασία: *Assess the reporting quality of randomized-  
controlled trials in multiple sclerosis from 2000 to 2015, based  
on CONSORT statement*

2016-παρόν Υποψήφιος Διδάκτωρ Παν Θεσσαλίας:

*Γενωμική σύγκλιση στη νόσο του Parkinson*

### ΕΡΓΑΣΙΑΚΗ ΕΜΠΕΙΡΙΑ

**Νοεμ. 2006-Νοεμ 2007:** Κέντρο Εκπαίδευσης Υγειονομικού Πολεμικής Αεροπορίας  
(251 ΓΝΑ). Συμμετοχή ως βοηθός στις κλινικές του 251 ΓΝΑ εκ περιτροπής. Εκτέλεση  
εφημεριών ως βοηθός γενικού εφημερεύοντος στο Τμήμα Επειγόντων Περιστατικών με  
κύρια αρμοδιότητα την αρχική αντιμετώπιση και διαλογή μεγάλης πληθώρας περιστατικών.

**Νοεμ. 2007 – Απρ. 2008:** Πτητική εκπαίδευση στα πλαίσια εξειδίκευσης στην Αεροπορική Ιατρική.

**Απρ 2008 – Σεπ 2013:** Ιατρός Μονάδας σε Πτέρυγες Μάχης της ΠΑ (116ΠΜ, Άραξος – 111ΠΜ, Αλμυρός). Στα καθήκοντά μου περιλαμβάνονταν η αρχική αντιμετώπιση όλου του εύρους περιστατικών σε επίπεδο πρωτοβάθμιας φροντίδας για το προσωπικό της Υπηρεσίας και των οικογενειών τους, ο προληπτικός υγειονομικός έλεγχος του προσωπικού και η Υγειονομική υποστήριξη όλων των δραστηριοτήτων της Μονάδας. Ταυτόχρονα εκτελούσα ενημερωτικές διαλέξεις σε θέματα προληπτικής ιατρικής και εξειδικευμένη εκπαίδευση στο Ιπτάμενο προσωπικό σε θέματα Αεροπορικής Ιατρικής και Φυσιολογίας των πτήσεων.

**Σεπ 2013 – Μάιος 2014:** Ειδικευόμενος Παθολογίας στην 'Α Παθολογική κλινική του 251 ΓΝΑ.

**Μάιος 2014 – Αυγ. 2014:** Ειδικευόμενος Ψυχιατρικής στην Ψυχιατρική κλινική του 251 ΓΝΑ.

**Ιαν 2015 – Νοεμ. 2019:** Ειδικευόμενος Νευρολογίας στην Νευρολογική κλινική του Παν. Γ. Νοσοκομείου Λάρισας. Στα πλαίσια της ειδικότητας παρακολούθησα όλες τις κλινικές και εκπαιδευτικές δραστηριότητες των κλινικών. Συγκεκριμένα στη Νευρολογική κλινική εξέτασα και παρακολούθησα ασθενείς από όλο το φάσμα των Νευρολογικών παθήσεων ως ιατρός της κλινικής, ως εφημερεύων Νευρολόγος στο τμήμα επειγόντων περιστατικών του ΠΓΝΛ και κατά το τελευταίο έτος της ειδικότητας ως ιατρός των τακτικών εξωτερικών ιατρείων της κλινικής.

**Νοεμ. 2019 – Σεπ. 2022:** Ιατρός μονάδας σε Πτέρυγα Μάχης της ΠΑ (111ΠΜ, Αλμυρός)



**Νοεμ. 2019 – παρόν:** Επιστημονικός συνεργάτης της Νευρολογικής Κλινικής του ΠΓΝ Λάρισας.

Επιστημονικός συνεργάτης της Νευρολογικής Κλινικής του 404 Στρατιωτικού Νοσοκομείου Λάρισας.

Ιδιώτης Νευρολόγος.

**Απρ. 2022 – παρόν:** Επιστημονικός συνεργάτης/σύμβουλος Ελληνικής Εταιρείας Νόσου Alzheimer και συναφών διαταραχών Λάρισας

**Σεπ. 2022 – παρόν:** Επιτελής Αρχηγείου Τακτικής Αεροπορίας

### **ΔΙΔΑΚΤΙΚΗ ΕΜΠΕΙΡΙΑ**

Έχω διδάξει επανειλημμένα σε προπτυχιακό και μεταπτυχιακό επίπεδο. Συγκεκριμένα

- Συνεργαζόμενος διδάσκων στο ΠΜΣ «Νευροαποκατάσταση» της Ιατρικής Σχολής του Πανεπιστημίου Θεσσαλίας. (2019-2020, 2020-2021, 2021-2022)
- Συνεργαζόμενος διδάσκων στο ΠΜΣ «Εφαρμοσμένη Κινησιολογία στις Ένοπλες Δυνάμεις» του ΠΘ στο γνωστικό αντικείμενο της αεροπορικής ιατρικής (2014-2015)
- Ιατρική Σχολή ΠΘ – Προπτυχιακό μάθημα Νευρολογίας
- Θεωρητικά μαθήματα Νοσηλευτικών ειδικοτήτων 5<sup>ης</sup> ΥΠΕ στο αντικείμενο της Νευρολογίας.

### **ΛΟΙΠΕΣ ΔΕΞΙΟΤΗΤΕΣ**

Έχω άριστη γνώση της Αγγλικής γλώσσας (C1/2).

Έχω άριστη γνώση χρήσης ηλεκτρονικών υπολογιστών.

### **ΔΗΜΟΣΙΕΥΣΕΙΣ**

- **Dimitrios Rikos**, Efthimios Dardiotis, Georgios Tsivgoulis, Elias Zintzaras, Georgios M. Hadjigeorgiou: *Reporting quality of randomized-controlled trials in multiple sclerosis from 2000 to 2015, based on CONSORT statement*. Multiple Sclerosis and Related Disorders 07/2016; 9., DOI:10.1016/j.msard.2016.07.013
- **Dimitrios Rikos**, Vasileios Siokas, Athina-Maria Aloizou, Zisis Tsouris, Paraskevi Aslanidou, Georgios Koutsis, Maria Anagnostouli, Dimitrios P Bogdanos, Nikolaos Grigoriadis, Georgios M Hadjigeorgiou, Efthimios Dardiotis: *TREM2 R47H (rs75932628) variant is unlikely to contribute to Multiple Sclerosis susceptibility and severity in a large Greek MS cohort*. Multiple Sclerosis and Related Disorders 07/2019; 35., DOI:10.1016/j.msard.2019.07.007
- **Rikos, D.**, Dardiotis, E., Aloizou, A. M., Siokas, V., Zintzaras, E., & Hadjigeorgiou, G. M. (2019). *Reporting quality of randomized controlled trials in Restless Legs Syndrome based on the CONSORT statement*. Tremor and Other Hyperkinetic Movements. DOI: <https://doi.org/10.7916/d8-0f2v-aq62>
- Georgia Xiromerisiou, Katerina Dadouli, Chrysoula Marogianni, Antonios Provatas, Panagiotis Ntellas, **Dimitrios Rikos**, Pantelis Stathis, Despina Georgouli, Gedeon Loules, Maria Zamanakou, Georgios M. Hadjigeorgiou: *A novel homozygous SACS mutation identified by whole exome sequencing-genotype phenotype correlations of all published cases*. Journal of Molecular Neuroscience 11/2019;, DOI:10.1007/s12031-019-01410-z
- Chrysoula Marogianni, **Dimitrios Rikos**, Antonios Provatas, Katerina Dadouli, Panagiotis Ntellas, Panagiota Tsitsi, George Patrinos, Efthimios Dardiotis, George Hadjigeorgiou, Georgia Xiromerisiou: *The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis and the creation of an online Database*. Neurobiology of Aging 04/2019;, DOI:10.1016/j.neurobiolaging.2019.04.012
- Efthimios Dardiotis, Emmanouil Karampinis, Vasileios Siokas, Athina-Maria Aloizou, **Dimitrios Rikos**, Styliani Ralli, Dimitra Papadimitriou, Dimitrios P. Bogdanos, Georgios M. Hadjigeorgiou: *ERCC6L2 rs591486 polymorphism and risk for amyotrophic lateral sclerosis in Greek population*. Neurological Sciences 03/2019; 40(11)., DOI:10.1007/s10072-019-03825-3
- Georgios M. Hadjigeorgiou, Persia-Maria Kountra, Georgios Koutsis, Vana Tsimourtou, Vasileios Siokas, Maria Dardioti, **Dimitrios Rikos**, Chrysoula Marogianni, Athina-Maria Aloizou, Georgia Karadima, Styliani Ralli, Nikolaos Grigoriadis, Dimitrios Bogdanos,

- Marios Panas, Efthimios Dardiotis: *Replication study of GWAS risk loci in Greek multiple sclerosis patients*. *Neurological Sciences* 10/2018;; DOI:10.1007/s10072-018-3617-6
- Efthimios Dardiotis, Athina-Maria Aloizou, Vasileios Siokas, Zisis Tsouris, **Dimitrios Rikos**, Chrysa Marogianni, Michael Aschner, Leda Kovatsi, Dimitrios P. Bogdanos, Aristidis Tsatsakis: *Paraoxonase-1 Genetic Polymorphisms in Organophosphate Metabolism*. *Toxicology* 10/2018; 411., DOI:10.1016/j.tox.2018.10.012
  - Maria Sokratous, Efthimios Dardiotis, Eleni Bellou, Zisis Tsouris, Amalia Michalopoulou, Maria Dardioti, Vasileios Siokas, **Dimitrios Rikos**, Aristidis Tsatsakis, Leda Kovatsi, Dimitrios P. Bogdanos, Georgios M. Hadjigeorgiou: *CpG Island Methylation Patterns in Relapsing-Remitting Multiple Sclerosis*. *Journal of Molecular Neuroscience* 03/2018; 64(3)., DOI:10.1007/s12031-018-1046-x
  - Efthimios Dardiotis, Vasileios Siokas, Eva Pantazi, Maria Dardioti, **Dimitrios Rikos**, Aikaterini Markou, Dimitra Papadimitriou, Matthaios Speletas, Georgios M. Hadjigeorgiou: *A novel mutation in TREM2 gene causing Nasu-Hakola disease and review of the literature*. *Neurobiology of aging* 01/2017; 53., DOI:10.1016/j.neurobiolaging.2017.01.015
  - Vasileios Siokas, Efthimios Dardiotis, Evangelia E. Tsironi, Georgios Tsivgoulis, **Dimitrios Rikos**, Maria Sokratous, Stylianos Koutsias, Konstantinos Paterakis, Georgia Deretzi, Georgios M. Hadjigeorgiou: *The Role of TOR1A Polymorphisms in Dystonia: A Systematic Review and Meta-Analysis*. *PLoS ONE* 01/2017; 12(1):e0169934., DOI:10.1371/journal.pone.0169934
  - Liampas I, Siouras AS, Siokas V, Tsouris Z, **Rikos D**, Brotis A, Aloizou AM, Dastamani M, Dardiotis E. *Migraine in transient global amnesia: a meta-analysis of observational studies*. *J Neurol*. 2021 Jan 2. doi: 10.1007/s00415-020-10363-y. Epub ahead of print. PMID: 33388926.
  - **Rikos D**, Marogianni C, Provatas A, Bourinaris T, Arnaoutoglou M, Stathis P, Patrinos GP, Dardiotis E, Hadjigeorgiou GM, Xiromerisiou G. *Screening for the C9ORF72 expansion in Greek Huntington Disease phenocopies and controls and meta-analysis of current data*. *Tremor Other Hyperkinet Mov (N Y)*. 2020 Jun 12;10:5. doi: 10.5334/tohm.61. PMID: 32775019; PMCID: PMC7394208.
  - Xiromerisiou G, Kalampokini S, **Rikos D**, Provatas A, Tsouris Z, Markou K, Ralli S, Dardiotis E. *Posterior reversible encephalopathy in a GT1a positive oculopharyngeal variant of Guillain-Barré syndrome: A case-report and review of the literature*. *Clin Neurol*

Neurosurg. 2020 Sep;196:106037. doi: 10.1016/j.clineuro.2020.106037. Epub 2020 Jun 22. PMID: 32623212.

- Dardiotis E, **Rikos D**, Siokas V, Aloizou AM, Tsouris Z, Sakalakis E, Brotis AG, Bogdanos DP, Hadjigeorgiou GM. *Assessment of TREM2 rs75932628 variant's association with Parkinson's disease in a Greek population and Meta-analysis of current data.* Int J Neurosci. 2021 Jun;131(6):544-548. doi: 10.1080/00207454.2020.1750388. Epub 2020 Apr 19. PMID: 32250197.
- Marogianni C, Georgouli D, Dadouli K, Ntellas P, **Rikos D**, Hadjigeorgiou GM, Spanaki C, Xiromerisiou G. *Identification of a novel de novo KMT2B variant in a Greek dystonia patient via exome sequencing genotype-phenotype correlations of all published cases.* Mol Biol Rep. 2021 Jan;48(1):371-379. doi: 10.1007/s11033-020-06057-3. Epub 2020 Dec 9. PMID: 33300088.
- Koute V, Michalopoulou A, Siokas V, Aloizou AM, **Rikos D**, Bogdanos DP, Kontopoulos E, Grivea IN, Syrogiannopoulos GA, Papadimitriou A, Hadjigeorgiou GM, Dardiotis E. *Val66Met polymorphism is associated with decreased likelihood for pediatric headache and migraine.* Neurol Res. 2021 Sep;43(9):715-723. doi: 10.1080/01616412.2021.1922181. Epub 2021 May 17. PMID: 34000980.

#### **ΕΡΕΥΝΗΤΙΚΑ ΕΝΔΙΑΦΕΡΟΝΤΑ**

Στα πλαίσια της συνεργασίας μου με την Νευρολογική Κλινική του Πανεπιστημιακού Γενικού Νοσοκομείου Λάρισας συμμετέχω σε ερευνητικές ομάδες με κύριο έργο την διερεύνηση του γενετικού υποβάθρου σύνθετων νευρολογικών νόσων (Gene Association Studies, Genotype-Phenotype correlations, meta-analyses κα).

Συμμετέχω ως ερευνητής σε μελέτη φάσης τέσσερα που αφορά την θεραπεία της αθροιστικής κεφαλαλγίας με μονοκλωνικό αντίσωμα έναντι του CGRP (ερενουμάμπη).

#### **ΚΛΙΝΙΚΑ ΕΝΔΙΑΦΕΡΟΝΤΑ**

Ασχολούμαι με όλο το φάσμα των νευρολογικών παθήσεων με ιδιαίτερο ενδιαφέρον για την ημικρανία και τις κεφαλαλγίες γενικότερα. Έχω συμμετάσχει ενεργά στην οργάνωση και λειτουργία του ειδικού ιατρείου ημικρανίας και παρεμβατικής νευρολογίας της Νευρολογικής κλινικής του ΠΓ Λάρισας και είμαι ενεργό μέλος του κλάδου

Κεφαλαγίας της Ελληνικής Νευρολογικής Εταιρείας. Είμαι επιστημονικός συνεργάτης και λειτουργώ ως σύμβουλος σε κλινικό ενημερωτικό και εκπαιδευτικό επίπεδο με την Ελληνική Εταιρεία νόσου Alzheimer και συναφών διαταραχών Λάρισας (ΕΕΝΑΛ).

#### **ΛΟΙΠΑ ΣΤΟΙΧΕΙΑ**

Είμαι παντρεμένος και έχω τρία παιδιά.



## «ΓΕΝΟΜΙΚΗ ΣΥΓΚΛΙΣΗ ΣΤΗ ΝΟΣΟ ΤΟΥ PARKINSON»

**ΡΙΚΟΣ ΔΗΜΗΤΡΙΟΣ**

Πανεπιστήμιο Θεσσαλίας, Τμήμα Ιατρικής, 2022

### **ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ**

1. **Ζιντζαράς Ηλίας**, Καθηγητής, Βιομαθηματικών, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας **(Επιβλέπων)**
2. **Δαρδιώτης Ευθύμιος**, Αναπ. Καθηγητής, Νευρολογίας, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας
3. **Στεφανίδης Ιωάννης**, Καθηγητής, Νεφρολογίας, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας





# Περιεχόμενα

<b>A ΜΕΡΟΣ: ΓΕΝΙΚΟ</b> .....	<b>21</b>
<b>ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ: ΓΕΝΕΤΙΚΗ ΚΑΙ ΓΕΝΕΤΙΚΗ ΕΠΙΔΗΜΙΟΛΟΓΙΑ</b> .....	<b>23</b>
<b>1. ΕΙΣΑΓΩΓΗ</b> .....	<b>23</b>
<b>2. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΓΕΝΕΤΙΚΗΣ</b> .....	<b>23</b>
2.1 DNA - χρωμοσώματα - γονίδια .....	24
2.2 Αντιγραφή .....	25
2.3 Μεταγραφή - μετάφραση .....	25
2.4 Γονίδια .....	26
2.5 Κληρονομικότητα .....	28
2.6 Μεταλλάξεις .....	29
2.7 Πολυμορφισμοί .....	30
2.8 Μονονουκλεοτιδικοί πολυμορφισμοί .....	30
2.9 Είδη γενετικών παθήσεων .....	30
2.10 Επιγενετική .....	32
<b>3. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΓΕΝΕΤΙΚΗΣ ΕΠΙΔΗΜΙΟΛΟΓΙΑΣ</b> .....	<b>34</b>
3.1 Μοντέλα κληρονόμησης .....	34
3.2 Ανισορροπία σύνδεσης (Linkage Disequilibrium-LD) .....	36
3.3 Κληρονομησιμότητα .....	36
<b>4. ΓΕΝΕΤΙΚΗ ΑΠΟΣΤΑΣΗ ΚΑΙ ΔΕΙΚΤΕΣ</b> .....	<b>37</b>
<b>5. ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΝΔΕΣΗΣ</b> .....	<b>39</b>
<b>6. ΧΑΡΤΟΓΡΑΦΗΣΗ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ</b> .....	<b>41</b>
<b>7. ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ</b> .....	<b>44</b>
<b>8. ΜΕΛΕΤΕΣ ΕΥΡΕΙΑΣ ΓΟΝΙΔΙΩΜΑΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ</b> .....	<b>46</b>
8.1 Τεχνολογία γονοτύπωσης .....	47
8.2 Σχεδιασμός μελέτης .....	48
8.3 Στατιστική ανάλυση .....	48

8.4	Αποτύπωση αποτελεσμάτων .....	49
9.	<b>ΜΕΤΑ ΑΝΑΛΥΣΗ ΓΕΝΕΤΙΚΩΝ ΜΕΛΕΤΩΝ .....</b>	<b>50</b>
9.1	Μετα-ανάλυση ολογενωμικών μελετών σύνδεσης (GWLS) .....	50
9.2	Μετα-ανάλυση ολογενωμικών μελετών συσχέτισης (GWAS) .....	53
10.	<b>ΕΠΙΓΕΝΕΤΙΚΗ .....</b>	<b>55</b>
	<b>ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ: Η ΝΟΣΟΣ PARKINSON .....</b>	<b>57</b>
1.	Εισαγωγή .....	57
2.	Διάγνωση – Κλινική εικόνα .....	57
3.	Επιδημιολογία .....	59
4.	Παθοφυσιολογία .....	59
5.	Διαφορική διάγνωση .....	60
6.	Διάγνωση .....	61
7.	Γενετική της νόσου Parkinson .....	61
8.	Θεραπεία .....	63
	<b>Β ΜΕΡΟΣ: ΕΙΔΙΚΟ .....</b>	<b>65</b>
	<b>ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ: ΜΕΤΑ ΑΝΑΛΥΣΗ ΟΛΟΓΕΝΩΜΙΚΩΝ ΜΕΛΕΤΩΝ ΣΥΝΔΕΣΗΣ</b>	
	<b>(GWLS) .....</b>	<b>66</b>
1.	<b>ΕΙΣΑΓΩΓΗ .....</b>	<b>66</b>
2.	<b>ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΣ .....</b>	<b>66</b>
2.1	Συστηματική αναζήτηση μελετών .....	66
2.2	Εφαρμογή κριτηρίων ένταξης και αποκλεισμού .....	66
2.3	Εξαγωγή δεδομένων .....	67
2.4	Ανάλυση δεδομένων .....	71
3.	<b>ΑΠΟΤΕΛΕΣΜΑΤΑ .....</b>	<b>71</b>
	<b>ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ: ΜΕΤΑ ΑΝΑΛΥΣΗ ΟΛΟΓΕΝΩΜΙΚΩΝ ΜΕΛΕΤΩΝ ΣΥΧΕΤΙΣΗΣ</b>	
	<b>(GWAS) .....</b>	<b>82</b>
1.	<b>ΕΙΣΑΓΩΓΗ .....</b>	<b>82</b>
2.	<b>ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΣ .....</b>	<b>82</b>

2.1	Βιβλιογραφική αναζήτηση.....	82
2.2	Κριτήρια ένταξης.....	83
2.3	Εξαγωγή δεδομένων.....	83
2.4	Διαχωρισμός των bins .....	83
2.5	Ανεύρεση σημαντικών SNP.....	83
2.6	Μετα ανάλυση μελετών.....	84
3.	<b>ΑΠΟΤΕΛΕΣΜΑΤΑ</b> .....	85
	<b>ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ: ΓΕΝΩΜΙΚΗ ΣΥΓΚΛΕΙΣΗ</b> .....	90
	<b>ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ: ΣΥΖΗΤΗΣΗ</b> .....	91
	<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b> .....	96
	<b>Γ ΜΕΡΟΣ: ΠΑΡΑΡΤΗΜΑ</b> .....	101



## **A ΜΕΡΟΣ: ΓΕΝΙΚΟ**



# ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ: ΓΕΝΕΤΙΚΗ ΚΑΙ ΓΕΝΕΤΙΚΗ ΕΠΙΔΗΜΙΟΛΟΓΙΑ

## 1. ΕΙΣΑΓΩΓΗ

Η μελέτη των παθήσεων και αποκάλυψη της συνεισφοράς της γενετικής ποικιλομορφίας στην εμφάνιση και εξέλιξη αυτών έχει ως απώτερο σκοπό την αναγνώριση και καταγραφή των πιθανών γενετικών τόπων που διαδραματίζουν κάποιο ρόλο στις παθήσεις αυτές. Η επίδραση αυτή μπορεί να είναι καθοριστική ή να προσθέτει απλά κάποιο ρίσκο για την εμφάνιση της πάθησης ανάλογα με τον τρόπο κληρονομίσης. Για να μελετηθεί το πολυσύνθετο αυτό θέμα γίνεται χρήση των γνώσεων μας για την γενετική και διαχείριση αυτών με όρους επιδημιολογίας και στατιστικής στα πλαίσια ενός σχετικά νέου πεδίου που ονομάστηκε γενετική επιδημιολογία που έχει ως βασικό αντικείμενο μελέτης τους πληθυσμούς και όχι τα άτομα. Στο εισαγωγικό κεφάλαιο της παρούσας θα γίνει αναφορά στις βασικές αρχές της γενετικής και στη συνέχεια της γενετικής επιδημιολογίας με έμφαση στις μελέτες γενετικής συσχέτισης, μελέτες γενετικής σύνδεσης ευρέως γονιδιώματος και μελέτες γενετικής συσχέτισης ευρέως γονιδιώματος.

## 2. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΓΕΝΕΤΙΚΗΣ

Το γενετικό υλικό (DNA) δύο οποιονδήποτε ανθρώπων είναι όμοιο κατά περίπου 99.5%. Αυτή η μικρή διαφορά είναι που προκαλεί την μεγάλη γενετικά καθορισμένη ετερογένεια και ποικιλομορφία μεταξύ των ανθρώπων και επηρεάζει από τα εξωτερικά χαρακτηριστικά μέχρι την προσωπικότητα και την επιρρέπεια σε ασθένειες. Από την πρώτη αναφορά σε κάποιο πιθανό κληρονομήσιμο στοιχείο τους ιπποκρατικούς χρόνους μέχρι σήμερα έχουν γίνει άλματα στην κατανόηση μας για τον τρόπο που μεταβιβάζεται η γενετική

πληροφορία. Στο κεφάλαιο αυτό θα γίνει μια σύντομη αναφορά στις βασικές αρχές της γενετικής επιστήμης και τις αρχές που την διέπουν.

## 2.1 DNA – χρωμοσώματα – γονίδια

Η βασική δομική και λειτουργική μονάδα του οργανισμού είναι τα κύτταρα. Σε κάθε οργανισμό υπάρχουν πολλά είδη κυττάρων τα οποία διαφοροποιούνται με σκοπό να επιτελέσουν συγκεκριμένες λειτουργίες. Η βασική μονάδα που περιέχει, και μεταφέρει την γενετική πληροφορία από το ένα κύτταρο στα παράγωγά του είναι μια χημική ένωση που ονομάζεται δεοξυ-ρυβονουκλεϊνικό οξύ ή DNA. Το DNA αποτελείται 4 βάσεις αμινοξέων (αδενίνη [A], θυμίνη [T], κυτοσίνη [C] και γουανίνη [G]) οι οποίες ενώνονται με δεσμούς ανά ζεύγη (A-T, C-G) και σχηματίζουν συμπληρωματικές αλυσίδες δίνοντας στο μόριο διπλή ελικοειδή δομή. Την αλληλουχία του DNA την διαβάζουμε από το 5' προς το 3' άκρο της. Το σύνολο του γενετικού υλικού ενός ανθρώπου ονομάζεται γονιδίωμα και περιέχει περίπου 3 δισεκατομμύρια ζεύγη βάσεων. Το μήκος του συνολικού DNA ενός κυττάρου σε πλήρη έκταση θα ήταν περίπου 2 μέτρα και ένα άτομο διαθέτει περίπου 100 δισεκατομμύρια μέτρα γενετικού υλικού. Το DNA με σκοπό την «συσκευασία» αλλά και την ρύθμιση της γονιδιακής έκφρασης οργανώνεται γύρω από πρωτεϊνικά μόρια που ονομάζονται ιστόνες και σχηματίζει τα νουκλεοσώματα τα οποία με την σειρά τους οργανώνονται σε χρωματοσώματα και τελικά σχηματίζουν τα χρωμοσώματα. Ο ανθρώπινος οργανισμός διαθέτει 23 ζεύγη χρωμοσωμάτων. Τα πρώτα 22 ονομάζονται αυτοσωματικά και αριθμούνται με τους αριθμούς 1-22 και το τελευταίο ζεύγος αποτελείται από δύο φυλετικά χρωμοσώματα που κωδικοποιούνται ως X και Y. Στην αλληλουχία των βάσεων του DNA εδράζεται η γενετική πληροφορία. Από τον τεράστιο όμως αριθμό βάσεων ένα μικρό μόνο μέρος κωδικοποιεί κάποια πληροφορία. Αυτή βρίσκεται σε τμήματα του DNA που ονομάζονται γονίδια και αποτελούν μια διακριτή μονάδα κληρονομικών πληροφοριών. Η πληροφορία αυτή με την διαδικασία της μεταγραφής και της μετάφρασης να οδηγήσει στην



παραγωγή πρωτεϊνικών μορίων που με την σειρά τους θα καθορίσουν και τον φαινότυπο του οργανισμού δηλαδή τα φυσικά εκφρασμένα χαρακτηριστικά του.

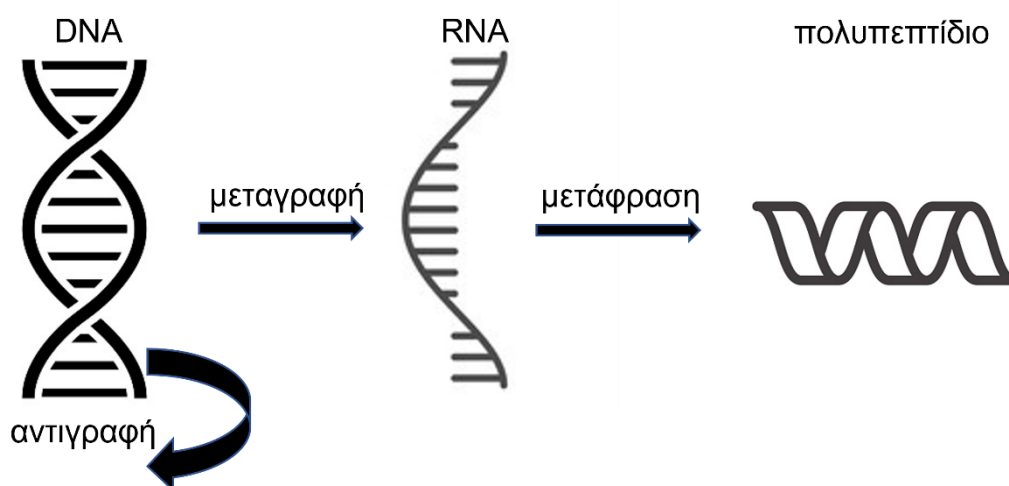
## 2.2 Αντιγραφή

Για την ανάπτυξη και διατήρηση του οργανισμού απαιτείται η διαίρεση του κυττάρου. Για να γίνει κάτι τέτοιο πρώτα θα πρέπει να μπορεί να φτιάξει ένα πλήρες αντίγραφο του γενετικού του κώδικα. Σε αυτήν την διαδικασία (replication) μια διπλή έλικα DNA αντιγράφεται ώστε να παράξει δύο πανομοιότυπες θυγατρικές έλικες που ονομάζονται αδερφές χρωματίδες. Η αντιγραφή του DNA είναι μια διαδικασία ημι-συντηριτική εφόσον κάθε αδερφή χρωματίδα περιέχει μια γνήσια και ένα αντίγραφο.

## 2.3 Μεταγραφή – μετάφραση

Εκτός από το DNA στον πυρήνα του κυττάρου υπάρχει και το ριβονουκλεϊνικό οξύ (RNA). Το RNA είναι ουσιαστικά ένα μονόκλωνο αντίγραφο του DNA και επιτρέπει την μεταφορά μικρών κομματιών πληροφορίας σε άλλα σημεία του κυττάρου και αποτελούν το πρότυπο (μήτρα) για την σύνθεση πρωτεϊνών. Η διαδικασία αυτή ξεκινάει με την μεταγραφή του DNA σε RNA και ακολουθείται από την μετάφραση του RNA σε αλληλουχία αμινοξέων με την δράση ενζύμων που ονομάζονται ριβοσώματα. Για να ξεκινήσει η μεταγραφή του DNA σε RNA θα πρέπει να υπάρξει ένας υποκινητής δηλαδή μια μικρή συγκεκριμένη αλληλουχία DNA στην οποία θα προσδεθεί το κατάλληλο ένζυμο (πολυμεράση). Οι υποκινητές βρίσκονται μέχρι και 50 χιλιάδες βάσεις (50kb) πριν από την έναρξη του γονιδίου. Το RNA που σχηματίζεται ονομάζεται μεταφορικό RNA (mRNA) και είναι αυτό που συνδέεται με το πρωτεϊνικό σύμπλοκο του ριβοσώματος το οποίο αντιστοιχίζει μία τριπλέτα βάσεων mRNA με ένα συγκεκριμένο αμινοξύ σε σειρά σχηματίζοντας τελικά το πρωτεϊνικό παράγωγο

της γενετικής πληροφορίας ([Εικόνα 1](#)). Η αντιστοίχιση αυτή γίνεται με βάση των γενετικό κώδικα. Πρόκειται για έναν κώδικα αντιστοίχισης τριπλέτας νουκλεοτιδίων (κωδικόνιο) με αμινοξέα. Υπάρχουν είκοσι διαθέσιμα αμινοξέα και 64 διαφορετικά κωδικόνια το οποίο σημαίνει ότι περισσότερα από ένα κωδικόνια αντιστοιχούν σε κάποιο αμινοξύ ([Πίνακας 1](#)). (3)



Εικόνα 1. Το βασικό δόγμα της βιολογίας

## 2.4 Γονίδια

Το μεταγραφικά ενεργό μέρος ενός χρωμοσώματος ονομάζεται γονίδιο. Τα γονίδια οργανώνονται σε εξώνια και ενδόνια. Τα εξώνια είναι οι περιοχές που (από)κωδικοποιούνται και αντιστοιχούν σε πρωτεϊνικό παράγωγο ενώ τα ενδόνια δεν μεταγράφονται. Υπάρχουν επιπλέον μη μεταγραφικά ενεργές περιοχές στην αρχή και το τέλος ενός γονιδίου. Η έναρξη του γονιδίου σηματοδοτείται από το κωδικόνιο έναρξης (ATG) στο 5' άκρο του. Άνωθεν αυτού υπάρχουν μη κωδικοποιημένες περιοχές που αποτελούν τους εκκινητές του γονιδίου και ακόμη πιο κεντρικά ρυθμιστικές περιοχές με συγκεκριμένα μοτίβα που διαδραματίζουν ρυθμιστικό ρόλο έκφρασης των γονιδίων και ανταποκρίνονται σε συγκεκριμένες πρωτεΐνες με τρόπο ειδικό ανάλογα με τον ιστό και προωθούν την μεταγραφή.

Στο 3' άκρο του γονιδίου υπάρχει το κωδικόνιο λήξης και στη συνέχεια μια ουρά πολλαπλών νουκλεοτιδίων αδενίνης. Σύμφωνα με το Human Genome Project [ΡΕΦ] υπάρχουν περίπου 20.000 με 25.000 γονίδια σε έναν άνθρωπο. Κάθε άτομο έχει δύο αντίγραφα κάθε γονιδίου, ένα από κάθε γονέα. Τα περισσότερα γονίδια είναι ίδια σε όλους τους ανθρώπους όμως ένα πολύ μικρό ποσοστό είναι διαφορετικό. Οι μορφές ενός γονιδίου που έχουν μικρές διαφορές στην αλληλουχία τους ονομάζονται αλλήλια.

		2 <sup>η</sup> βάση κωδικονίου					
		T	C	A	G		
1 <sup>η</sup> βάση κωδικονίου	T	Phe	Ser	Tyr	Cys	U	3 <sup>η</sup> βάση κωδικονίου
		Phe	Ser	Tyr	Cys	C	
		Leu	Ser	STOP	STOP	A	
		Leu	Ser	STOP	Trp	G	
	C	Leu	Pro	His	Arg	U	
		Leu	Pro	His	Arg	C	
		Leu	Pro	Gln	Arg	A	
		Leu	Pro	Gln	Arg	G	
	A	Ile	Thr	Asn	Ser	U	
		Ile	Thr	Asn	Ser	C	
		Ile	Thr	Lys	Arg	A	
		Met	Thr	Lys	Arg	G	
	G	Val	Ala	Asp	Gly	U	
		Val	Ala	Asp	Gly	C	
		Val	Ala	Glu	Gly	A	
		Val	Ala	Glu	Gly	G	
Αμινοξέα							
Ala: Αλανίνη Arg: Αργινίνη Asn: Ασπαραγίνη Asp: Ασπαρτικό οξύ							
Cys: Κυστεΐνη Gln: Γλουταμίνη Glu: Γλουταμικό οξύ							
Gly: Γλυκίνη His: Ιστιδίνη Ile: Ισολευκίνη Leu: Λευκίνη							
Lys: Λυσίνη Met: Μεθειονίνη Phe: Φενυλαλανίνη Pro: Προλίνη							
Ser: Σερίνη Thr: Θρεονίνη Trp: Τρυπτοφάνη Tyr: Τυροσίνη							
Val: Βαλίνη							

## 2.5 Κληρονομικότητα

Η διαδικασία της μεταφοράς της γενετικής πληροφορίας από τους γονείς στους απογόνους γίνεται με την μεταφορά των γονιδίων από τα αναπαραγωγικά κύτταρα, τους γαμέτες, στην επόμενη γενιά. Οι γαμέτες είναι απλοειδή κύτταρα δηλαδή περιέχουν ένα σετ χρωμοσωμάτων (23) σε αντίθεση με τα σωματικά κύτταρα τα οποία είναι διπλοειδή και περιέχουν 2 σετ χρωμοσωμάτων (46), ένα από κάθε γονέα. Κατά την γονιμοποίηση οι άρρενες και θήλεις γαμέτες (σπερματοζώαριο και ωάριο) ενώνονται και μεταβιβάζουν γονίδια στους απογόνους. Κάθε γαμέτης προσφέρει το ένα σετ χρωμοσωμάτων και το κύτταρο που προκύπτει (το ζυγωτό) είναι διπλοειδές. Η συνεχής αναπαραγωγή του ζυγωτού με την διαδικασία της μίτωσης παράγει όλα τα σωματικά κύτταρα του οργανισμού που περιέχουν ακριβή αντίγραφα του συνόλου των χρωμοσωμάτων. Εξαιρέση αποτελούν οι γαμέτες του νέου ατόμου οι οποίοι είναι απλοειδείς και προκύπτουν από την διαδικασία της μείωσης. Κατά την διαδικασία αυτή τα διπλά αντίγραφα χρωμοσωμάτων υπόκεινται σε διασταύρωση (crossing) ανταλλάσσοντας τμήματα τους σε περιοχές που ονομάζονται χιάσματα, και διαιρούνται σε τέσσερις χρωματίδες σε αντίστοιχα απλοειδή κύτταρα τα οποία διαφέρουν ως προς την γενετική πληροφορία μεταξύ τους αλλά και από τα προγονικά κύτταρα (segregation). Η διαδικασία αυτή έχει σαν αποτέλεσμα την γενετική ποικιλομορφία. Αυτό συμβαίνει γιατί τα χρωμοσώματα κατά την διάρκεια της μείωσης τοποθετούνται ανεξάρτητα (τυχαία) πιο κοντά στον έναν ή τον άλλο πόλο του κυττάρου με αποτέλεσμα η πιθανότητα το απλοειδές κύτταρο να περιέχει το μητρικό ή το πατρικό ομόλογο να είναι 50%-50% (independent assortment). Επιπλέον κατά την μείωση επιτελείται το crossing με αποτέλεσμα τα παράγωγα χρωμοσώματα να μην είναι αποκλειστικά πατρικά ή μητρικά αλλά ανασυνδυασμένα. Σε αυτά τα αίτια γενετικής ποικιλομορφίας θα πρέπει να προστεθεί και η τυχαία γονιμοποίηση, δηλαδή η τυχαία ένωση ενός σπερματοζωαρίου και ενός ωαρίου. Οι ανωτέρω λόγοι καθιστούν τις δυνατότητες πιθανών συνδυασμών αστρονομικές. Το αποτέλεσμα αυτών των διεργασιών δηλαδή τα γενικά και ειδικά παρατηρούμενα χαρακτηριστικά του οργανισμού

που προκύπτει, μετά και από την επίδραση του περιβάλλοντος ονομάζεται φαινότυπος.

## 2.6 Μεταλλάξεις

Εάν και όλα τα κύτταρα προέρχονται από αντιγραφή του ίδιου ζυγωτού, οι πολλοί κύκλοι αντιγραφής οδηγούν αναπόφευκτα σε αλλαγές στην αλληλουχία του DNA που παρά τις πολλαπλές δικλείδες ελέγχου και επιδιόρθωσης παραμένουν μη ανιχνεύσιμες με αποτέλεσμα να αναπαράγονται. Τις αλλαγές αυτές τις ονομάζουμε μεταλλάξεις και μπορεί να έχουν ή να μην έχουν κάποια επίδραση στην φυσιολογική λειτουργία του κυττάρου και του οργανισμού. Εάν η μετάλλαξη συμβαίνει σε κάποιον γαμέτη θα περάσει και στους απογόνους ενώ στην αντίθετη περίπτωση της παρουσίας μετάλλαξης σε σωματικό κύτταρο αυτή δεν θα μεταβιβασθεί στους απογόνους του οργανισμού. Η απλούστερη μορφή μετάλλαξης είναι η σημειακή μετάλλαξη όπου ένα νουκλεοτίδιο αλλάζει με ένα άλλο. Μια τέτοια μετάλλαξη μπορεί να προκαλέσει αλλαγή στο κωδικώνιο με αποτέλεσμα και την αλλαγή του αμινοξέος που αντιστοιχεί ή τον πρόωρο τερματισμό της μετάφρασης. Μια τέτοια μετάλλαξη ονομάζεται μη συνώνυμη ενώ στην αντίθετη περίπτωση ονομάζεται συνώνυμη. Ένα άλλο είδος μεταλλάξεων είναι οι εισαγωγές (insertions) και οι διαγραφές (deletions) που μπορεί να αφορούν σε ζεύγη βάσεων ή και σε μεγαλύτερα κομμάτια γενετικού υλικού. Η επίδραση που μπορεί να έχουν τέτοιες μεταλλάξεις εξαρτάται εκτός από το μέγεθος τους και από το σημείο που λαμβάνουν χώρα. Μια διαγραφή ή προσθήκη ενός ζεύγους βάσεων στο μεταγραφικό μέρος ενός γονιδίου μπορεί να προκαλέσει μετατόπιση του πλαισίου ανάγνωσης του με καταστροφικά αποτελέσματα. Σε ένα ανώτερο επίπεδο μια τέτοια μετάλλαξη μπορεί να επηρεάσει ένα γονίδιο ή μια ομάδα γονιδίων ή σε κάποιες περιπτώσεις και ολόκληρα χρωμοσώματα. Όταν ένα κομμάτι γενετικού υλικού αποκόπτεται από ένα σημείο και επικολλάται σε ένα άλλο έχουμε το φαινόμενο της μετάθεσης (translocation) ενώ όταν ένα κομμάτι αποκόπτεται αλλάζει φορά αλλά επικολλάται στο ίδιο σημείο το φαινόμενο της αναστροφής (inversion).

## 2.7 Πολυμορφισμοί

Εάν μια μετάλλαξη έχει συμβεί και έχει την δυνατότητα να μεταβιβασθεί στους απογόνους τότε λέμε ότι υπάρχει μια γενετική παραλλαγή (variant). Παραλλαγές που έχουν συχνότητα στον γενικό πληθυσμό μεγαλύτερο από 1% ονομάζονται πολυμορφισμοί ενώ κάτω από αυτό το όριο ονομάζονται σπάνιες παραλλαγές.

## 2.8 Μονονουκλεοτιδικοί πολυμορφισμοί

Ενώ υπάρχουν διάφορες μορφές παραλλαγών οι μονονουκλεοτιδικοί πολυμορφισμοί (Single Nucleotide Polymorphisms – SNPs) αποτελούν τους πιο συχνούς. Τα SNPs αφορούν παραλλαγές σε μια μοναδική βάση του DNA. Αυτό σημαίνει ότι σε ένα συγκεκριμένο σημείο του DNA (locus) σε ένα ομόλογο ζευγάρι χρωμοσωμάτων το άτομο μπορεί να έχει την ίδια νουκλεοτιδική βάση και στα δυο μέλη (πχ AA) ή να έχει διαφορετικές (πχ AT). Στο ίδιο ακριβώς σημείο ένα άλλο άτομο μπορεί να εμφανίζει κάποιον άλλον συνδυασμό των βάσεων αυτών (πχ TT). Εξαιτίας αυτού του γεγονότος λέμε ότι το σημείο αυτό είναι πολυμορφικό και οι νουκλεοτιδικές βάσεις που μπορεί να εμφανισθούν ονομάζονται αλληλία. Τα άτομα που διαθέτουν ένα ζευγάρι από ίδιες βάσεις (πχ AA ή TT) ονομάζονται ομοζυγώτες για το αντίστοιχο αλληλίο ενώ αυτά που διαθέτουν διαφορετικές βάσεις ετεροζυγώτες. Επιπλέον το αλληλίο που εμφανίζεται πιο σπάνια στον πληθυσμό το ονομάζουμε minor allele ενώ το πιο συχνό major allele (4).

## 2.9 Είδη γενετικών παθήσεων

Εάν και τα γενετικά νοσήματα θεωρούνται σπάνια αποτελούν το 80% των σπάνιων παθήσεων που συνολικά αφορούν μερικές χιλιάδες παθήσεις και αθροιστικά αφορούν περίπου 1 ανά 17 άτομα πληθυσμού. Επιπλέον ο γενετικός μας κώδικας διαδραματίζει κάποιον ρόλο σε μικρότερο ή μεγαλύτερο βαθμό σε όλες τις παθολογικές διεργασίες συμπεριλαμβανομένων και των κοινών παθήσεων. Κάποιο μέρος της γενετικής ποικιλομορφίας μας μπορεί να καθιστά ένα άτομο πιο επιρρεπές στην εμφάνιση μιας

πάθησης ή στην αντίθετη περίπτωση να τον προστατεύει από αυτήν. Το περιβάλλον μπορεί να επιδρά σημαντικά στο άτομο αλλά ακόμη και σε αυτή τη περίπτωση υπάρχει η πιθανότητα η επίδραση αυτή να ρυθμίζεται από το γενετικό υπόβαθρο του ατόμου. Μπορούμε να περιγράψουμε τις παθήσεις που σχετίζονται με το γενετικό μας υλικό σε τρεις βασικές κατηγορίες. Τις χρωμοσωμικές παθήσεις, τις μονογονιδιακές παθήσεις και τις σύνθετες.

Όπως περιεγράφηκε στο κεφάλαιο 2 κάθε κύτταρο διαθέτει ένα πλήρες αντίγραφο του γονιδιώματος το οποίο αποτελείται από 22 ζεύγη ομόλογων χρωμοσωμάτων και τα φυλετικά χρωμοσώματα ΧΧ στα θήλεα άτομα και ΧΥ στους άρρενες. Οι γαμέτες διαθέτουν 23 χρωμοσώματα. Παρόλη την παρουσία πολλαπλών μηχανισμών που έχουν σκοπό την ακριβή μεταφορά ενός πλήρους αντιγράφου του γονιδιώματος κατά την κυτταρική διαίρεση κάποιες φορές μπορεί να υπάρξουν λάθη. Αυτό μπορεί να οδηγήσει σε κύτταρα με χρωμοσωμικές ανωμαλίες που μπορούν να κατηγοριοποιηθούν σε αριθμητικές ή δομικές ανωμαλίες.

Εάν ένα κύτταρο περιέχει περισσότερα από δύο σετ απλοειδούς γονιδιώματος λέμε ότι εμφανίζει πολυπλοειδία. Στην περίπτωση που ένας γαμέτης περιέχει περισσότερα ή λιγότερα χρωμοσώματα από το φυσιολογικό οδηγεί στην δημιουργία ενός ζυγωτού που θα έχει έλλειψη ενός χρωμοσώματος ή μεγαλύτερο αριθμό χρωμοσωμάτων από το φυσιολογικό και ονομάζεται ανευπλοειδικό. Συνήθως αυτή η κατάσταση δεν είναι συμβατή με την ζωή όμως υπάρχουν και περιπτώσεις όπου το έμβρυο επιβιώνει μέχρι την γέννηση όπως η τρισωμία 13 και 18 ή επιβιώνει μέχρι μεγάλη ηλικία όπως η τρισωμία 21.

Δομική βλάβη του DNA μπορεί να οδηγήσει σε θραύση αυτού με αποτέλεσμα οι μηχανισμοί επιδιόρθωσης να επισκευάσουν λανθασμένα το DNA οδηγώντας σε απώλεια υλικού (deletion) αναστροφή ή δημιουργία ενός κυκλικού μορίου που εφόσον διαθέτει ένα κεντρομερίδιο να μπορεί να μεταβιβασθεί στους απογόνους.

Η εμφάνιση πολλών παθήσεων εξαρτάται από ένα σημείο στο γονιδίωμα ή ένα γονίδιο και τις παθήσεις αυτές τις ονομάζουμε μονογονιδιακές. Εκτός από κάποιες εξαιρέσεις ακολουθούν τους νόμους του Mendel και κληρονομούνται με έναν από τους πέντε τρόπους κληρονόμησης (αυτοσωματικός επικρατητικός, αυτοσωματικός υπολειπόμενος, Χ-φυλοσύνδετος επικρατητικός, Χ-φυλοσύνδετος υπολειπόμενος και Υ-φυλοσύνδετος). Οι πολυμορφισμοί ή μεταλλάξεις που οδηγούν σε τέτοιες παθήσεις μπορούν να είναι των μορφών που περιγράφηκαν στην παράγραφο 2.6.

Ιδιαίτερη περίπτωση αποτελούν οι μιτοχονδριακές παθήσεις. Τα μιτοχόνδρια αποτελούν οργανίλια του κυττάρου που διαθέτουν ξεχωριστό γενετικό υλικό από αυτό του πυρήνα στην μορφή ενός κυκλικού μορίου DNA που κωδικοποιεί 37 γονίδια. Τα μιτοχόνδρια μεταβιβάζονται στους απογόνους αποκλειστικά από τη μητέρα και επιπλέον μπορεί να εμφανίζουν ετεροπλασμία δηλαδή κάποια από αυτά να περιέχουν μια μετάλλαξη ενώ άλλα όχι στον ίδιο οργανισμό. Έτσι η κληρονομηση μιας μετάλλαξης και έκφραση του φαινοτύπου εξαρτάται από το μιτοχονδριακό γενετικό υλικό της μητέρας αλλά και από το μεταλλαγμένο φορτίο που θα διαθέτει ο γαμέτης που θα γονιμοποιηθεί.

Εκτός από τις ανωτέρω παθήσεις μεγάλη πρόκληση αποτελούν οι σύνθετες παθήσεις στις οποίες ο τελικός φαινότυπος επηρεάζεται σε διάφορους βαθμούς από σύνθετους γενετικούς και περιβαλλοντικούς παράγοντες και την αλληλεπίδρασή τους. Η μελέτη αυτών των σύνθετων νοσημάτων ανάμεσά τους και η νόσος Parkinson θα απασχολήσει κυρίως την παρούσα διατριβή.

## 2.10 Επιγενετική

Περιγράψαμε ότι η νουκλεοτιδική αλληλουχία περιέχει την γενετική πληροφορία και αυτή μεταβιβάζεται στους απογόνους και ότι αλλαγές σε αυτήν την αλληλουχία σε διάφορα επίπεδα επηρεάζει σε άλλοτε άλλο βαθμό την επιρρέπεια ενός οργανισμού να εμφανίσει κάποιο χαρακτηριστικό ή κάποια ασθένεια. Πέρα όμως από το επίπεδο της νουκλεοτιδικής αλληλουχίας υπάρχει και ένα ακόμη επίπεδο πληροφορίας που επίσης μπορεί να μεταβιβάζεται σε παράγωγες γενετικές σειρές και σε αυτήν την περίπτωση ρυθμίζει την έκφραση των γονιδίων. Η μελέτη αυτού του επιπέδου ρύθμισης ονομάζεται επιγενετική. Η επιγενετική πληροφορία παίρνει την μορφή χημικών ομάδων (και όχι νουκλεοτιδίων) οι οποίες αλλάζουν την συμπεριφορά των γονιδίων στα οποία βρίσκονται. Στα ανθρώπινα κύτταρα ένα δινουκλεοτίδιο κυτοσίνης – γουανίνης (CpG) σε σειρά με έναν φωσφοδιεστερικό δεσμό ανάμεσα τους (και όχι δυο συμπληρωματικά νουκλεοτίδια) μπορούν να δεχθούν μια ένωση μεθυλίου ( $\text{CH}_3$ ) με την διαδικασία που ονομάζεται μεθυλίωση με την δράση ενζύμων που ονομάζονται μεθυλτρανσφεράσες (DNMTs). Κατά την αντιγραφή αυτών των αλληλουχιών παράγονται δυο ημιμεθυλιωμένες άλυστοι οι οποίες αναγνωρίζονται και οι μη μεθυλιωμένες συμπληρωματικές άλυστοι μεθυλιώνονται. Έτσι η πληροφορία αυτή της μεθυλίωσης σε κάποια σημεία μεταβιβάζεται στα παράγωγα



κύτταρα. Αυτά τα σημεία αναγνωρίζονται από συγκεκριμένες ομάδες πρωτεϊνών που συμμετέχουν στην οργάνωση του DNA σε χρωματίδες μέσω της αναδίπλωσης τους σε οκταμερείς ιστόνες. Αντίστοιχα οι μεθυλιωμένες περιοχές μπορούν να απομεθυλιώνονται. Η διαδικασία αυτή είναι δυναμική και μπορεί να ανταποκρίνεται σε ερεθίσματα ή σε συγκεκριμένες φάσεις της ανάπτυξης. Με την μεθυλίωση ή απομεθυλίωση των περιοχών αυτών υπάρχει η δυνατότητα της ρύθμισης της έκφρασης ενός γονιδίου θέτοντας το σε κατάσταση ενεργή ή ανενεργή (5). Την ίδια ρυθμιστική λειτουργία μπορεί να έχει η ακετυλίωση των ίδιων των ιστονών.

Αντιγραφή	Η παραγωγή δύο πανομοιότυπων αδελφών χρωματίδων από μια διπλή έλικα DNA
Μεταγραφή	Η δημιουργία μια συμπληρωματικής έλικας RNA από μια έλικα DNA
Μετάφραση	Η παραγωγή ενός πολυπεπτιδίου από μια έλικα RNA
Γενετικός κώδικας	Κώδικας αντιστοίχισης μιας τριπλέτας νουκλεοτιδίων (κωδικόνιο) με ένα αμινοξύ
Γονίδιο	Το μεταγραφικά ενεργό μέρος ενός χρωμοσώματος. Η βασική μονάδα κληρονομικότητας
Μείωση	Η διαδικασία κατά την οποία ένα κύτταρο με δύο αντίγραφα κάθε χρωμοσώματος (διπλοειδές) παράγει δύο κύτταρα με ένα αντίγραφο χρωμοσωμάτων (απλοειδές)

Μετάλλαξη (mutation)	Αλλαγή στην αλληλουχία του γενετικού υλικού
Παραλλαγή (variant)	Μετάλλαξη που είναι δυνητικά κληρονομήσιμη στους απογόνους
Πολυμορφισμός	Παραλλαγή με συχνότητα στον πληθυσμό >1%
Μονονουκλεοτιδικός πολυμορφισμός (SNP)	Πολυμορφισμός σε μια μοναδική βάση DNA
Αλληλίο	Μία ή περισσότερες διαφορετικές εκδοχές μιας αλληλουχίας DNA σε ένα συγκεκριμένο σημείο αυτής.
Απλότυπος	Ο συνδυασμός των αλληλίων (ή πολυμορφισμών) που βρίσκονται στο ίδιο χρωμόσωμα
Φαινότυπος	Το σύνολο των παρατηρούμενων χαρακτηριστικών ενός οργανισμού

Πίνακας 2. Επεξήγηση βασικών όρων γενετικής

### 3. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΓΕΝΕΤΙΚΗΣ ΕΠΙΔΗΜΙΟΛΟΓΙΑΣ

#### 3.1 Μοντέλα κληρονομής

Η ανθρώπινη γενετική επιδημιολογία έχει ως στόχο την μελέτη των κληρονομήσιμων στοιχείων και την επίδρασή τους στην εμφάνιση χαρακτηριστικών και περισσότερο των ασθενειών στον άνθρωπο με την εκμετάλλευση των γνώσεων της γενικής και αναλυτικής επιδημιολογίας, της γενετικής, της μοριακής βιολογίας, της βιοστατιστικής και της βιοπληροφορικής.

Όλη αυτή η διεργασία γίνεται σε επίπεδο πληθυσμών με σκοπό την αποκάλυψη της γενετικής αρχιτεκτονικής των παθήσεων που όπως φάνηκε μετά τα πρώτα βήματα της γενετικής είναι ιδιαίτερα πολύπλοκη.

Ο Μέντελ ήταν ο πρώτος που πρότεινε ένα σαφές μοντέλο για να εξηγήσει την κληρονομία κάποιου παράγοντα και την επίδρασή του στον φαινότυπο. Περιέγραψε του νόμους του διαχωρισμού (segregation) και της ανεξάρτητης ταξινόμησης (independent assortment) καταγράφοντας ότι όταν ένας οργανισμός παράγει γαμέτες τα αλληλία ενός γονιδίου διαχωρίζονται ώστε κάθε γαμέτης να λάβει από ένα και ότι αλληλία από διαφορετικά σημεία (loci) κληρονομούνται ξεχωριστά. Το μοντέλο του περιλάμβανε επιπλέον την έννοια του επικρατούς και υπολειπόμενου αλληλίου με την έννοια ότι ο φαινότυπος ενός υπολειπόμενου αλληλίου για να εκφρασθεί θα πρέπει ο οργανισμός να είναι ομόζυγος για αυτό το αλληλίο.

Εάν σε ένα locus τα αλληλία είναι A και B και η συχνότητα του αλληλίου A είναι p τότε η συχνότητα του B είναι q=1-p. Σε μεγάλους πληθυσμούς όταν τα αλληλία κληρονομούνται ανεξάρτητα οι αναμενόμενες συχνότητες των αλληλίων θα είναι συνάρτηση των συχνοτήτων τους και δεν θα αλλάζουν από γενιά σε γενιά. Όταν συμβαίνει αυτό λέμε ότι το Locus αυτό είναι σε ισορροπία Hardy-Weinberg (HWE). Εφόσον η συχνότητα των αλληλίων αναφέρεται στα χρωμοσώματα και όχι στους διπλοειδείς οργανισμούς και ο όρος γονότυπος αναφέρεται στα ζεύγη αλληλίων τότε η γονοτυπική συχνότητα των πιθανών συνδυασμών των αλληλίων A και B δηλαδή AA, AB και BB θα είναι  $p^2$ ,  $2p(i-p)$  και  $(1-p)^2$ . Η εξίσωση λοιπόν HW Μπορεί να εκφρασθεί:

$$p^2+2pq+q^2=1$$

Όπου p και q οι συχνότητες των αλληλίων A και B και ισχύει  $p+q=1$ . Για να ισχύει αυτή η σχέση θα πρέπει να υπάρχει HWE δηλαδή το ταίριασμα των συντρόφων στον πληθυσμό να γίνεται τυχαία, να μην υπάρχει επιλογή και μεταλλάξεις καθώς και να μην υπάρχει σχέση του locus με την ικανότητα του ατόμου να επιβιώσει τουλάχιστον μέχρι την αναπαραγωγική ηλικία. Μια ακόμη προϋπόθεση είναι ο πληθυσμός να είναι αρκετά μεγάλος ώστε να αποφεύγεται η επίδραση της ποικιλομορφίας του ιδρυτικού πληθυσμού στον τελικό πληθυσμό (founder effect).

### 3.2 Ανισορροπία σύνδεσης (Linkage Disequilibrium-LD)

Όταν δυο loci βρίσκονται σε κοντινή φυσική απόσταση τείνουν να κληρονομούνται μαζί. Ο όρος ανισορροπία σύνδεσης (LD) αναφέρεται σε μια παρόμοια περίπτωση όμως στο επίπεδο του πληθυσμού. Γενικά περιγράφει την κατάσταση κατά την οποία υπάρχει «σχέση» και μάλλον μη τυχαία μεταξύ δύο αλληλίων. Έτσι η συχνότητα του απλότυπου αυτών των αλληλίων είναι διαφορετική από την αναμενόμενη με βάση τις συχνότητες των αλληλίων.

### 3.3 Κληρονομησιμότητα

Ως κληρονομησιμότητα (Heredability) αναφέρουμε το ποσοστό της ποικιλότητας ενός χαρακτηριστικού που οφείλεται στη γενετική ποικιλομορφία. Σαν μαθηματικό μέγεθος θα λέγαμε ότι αποτελεί τον λόγο της γενετικής διακύμανσης προς την συνολική διακύμανση. Περιλαμβάνει το επικρατές (dominant) το αθροιστικό (additive) και το επιστατικό (epistatic) μοντέλο επιρροής αλλά με την στενή έννοια μπορούμε να πούμε ότι η προσοχή περιορίζεται στο αθροιστικό μοντέλο. Έτσι η κληρονομησιμότητα  $H$  μπορεί να πάρει τιμές από 0 έως 1 με τις τιμές κοντά στο 1 να υπονοούν ότι η γενετική συμμετοχή στην εμφάνιση ενός χαρακτηριστικού είναι μεγάλη ενώ τιμές κοντά στο 0 ότι μόνο ασθενώς συμμετέχει η γενετική ποικιλομορφία στην εμφάνιση του χαρακτηριστικού αυτού.

Για την ποσοτικοποίηση των ανωτέρω μπορούμε να θεωρήσουμε ότι η γενετική ποικιλομορφία συμμετέχει στην διαμόρφωση της φαινοτυπικής ποικιλομορφίας και εφόσον μπορούμε να περιγράψουμε ένα κανονικά κατανομημένο χαρακτηριστικό με βάση τον μέσο όρο και την διακύμανση του θα λέγαμε ότι η γενετική διακύμανση καθορίζει την φαινοτυπική διακύμανση. Έτσι η φαινοτυπική διακύμανση θα μπορούσε να περιγραφεί ως μία γραμμική σχέση της μορφής:  $Y_{gc} = \mu + G_g + E_c$ , όπου η φαινοτυπική τιμή  $Y$  σε ένα άτομο με γονότυπο  $g$  και περιβάλλον  $c$  ισούται με μία απόκλιση από έναν μέσο όρο του πληθυσμού  $\mu$  κατά την γενετική ( $G$ ) και περιβαλλοντική ( $E$ ) επιρροή.

Ένας τρόπος που έχει χρησιμοποιηθεί για την μελέτη της ισχύος του γενετικού στοιχείου σε έναν πληθυσμό είναι οι μελέτες διδύμων. Οι μονοζυγωτικοί δίδυμοι είναι γενετικά ταυτόσημοι ενώ οι διζυγωτικοί δίδυμοι μπορούν να θεωρηθούν σαν ζευγάρια ίδιας ηλικίας με κοινή περιβαλλοντική επιρροή. Έτσι εάν η συσχέτιση μονοζυγωτικών και διζυγωτικών είναι παρόμοια μπορούμε να θεωρήσουμε ότι υπάρχει μικρή γενετική συμμετοχή ενώ στην αντίθετη περίπτωση της μεγάλης γενετικής επιρροής η συσχέτιση των μονοζυγωτικών θα είναι μεγαλύτερη από τους διζυγωτικούς δίδυμους.

#### 4. ΓΕΝΕΤΙΚΗ ΑΠΟΣΤΑΣΗ ΚΑΙ ΔΕΙΚΤΕΣ

Όπως αναφέρθηκε και νωρίτερα η διαδικασία της παραγωγής γαμετών δηλαδή απλοειδών κυττάρων ονομάζεται μείωση και κατά αυτή τα διπλοειδή κύτταρα διαχωρίζουν τα δύο σετ χρωμοσωμάτων τους δημιουργώντας 4 αδερφές χρωματίδες και στη συνέχεια αντίστοιχους γαμέτες. Κατά την μείωση τα χρωμοσώματα συστοίχονται σε μια γραμμή και κάποιες φορές ανταλλάσσουν υλικό με την διαδικασία της διασταύρωσης (crossover) σχηματίζοντας χιάσματα. Αποτέλεσμα αυτής της διαδικασίας είναι η παραγωγή νέων συνδυασμών αλληλίων και αυτό ονομάζεται ανασυνδυασμός (recombination). Τέτοιος ανασυνδυασμός είναι πιο εύκολο να συμβεί μεταξύ περιοχών που βρίσκονται πιο μακριά σε αντίθεση με περιοχές που είναι πιο κοντά και μπορεί να έχουν κάποια σύνδεση με αποτέλεσμα να κληρονομούνται συνήθως μαζί.

Ως γενετική απόσταση μεταξύ δύο σημείων (loci) που εδράζονται στο ίδιο χρωμόσωμα ορίζεται ως ο μέσος όρος των γεγονότων διασταύρωσης μεταξύ αυτών των δύο σημείων σε όλες τις μειώσεις και από πολλαπλές οικογένειες. Η γενετική απόσταση μετριέται σε μονάδες Morgan ή συχνότερα στην εκατοστιαία υποδιαίρεση centimorgan (cM). Δύο σημεία ή γενετικοί δείκτες έχουν γενετική απόσταση 1 cM όταν συμβαίνει μια διασταύρωση μεταξύ τους κατά την μείωση μία στις 100 φορές (0,01) δηλαδή με συχνότητα 1%. Μία σχετική ποσότητα με τον ρυθμό ανασυνδυασμού (recombination rate) που καθορίζει την γενετική απόσταση είναι ο λόγος ανασυνδυασμού (recombination fraction -  $\theta$ ) που περιγράφει την πιθανότητα να συμβεί ένας ανασυνδυασμός μεταξύ δύο σημείων

και λαμβάνει τιμές από 0 έως 0.5. Εάν  $\theta=0.5$  υποδηλώνει υπάρχει 50%-50% πιθανότητα να συμβεί ανασυνδυασμός δηλαδή οι δείκτες κληρονομούνται ανεξάρτητα. Εάν  $\theta<0.5$  υποδηλώνεται ότι υπάρχει κάποια σύνδεση μεταξύ των σημείων. Εφόσον ο ρυθμός ανασυνδυασμού ισούται με το άθροισμα της πιθανότητας να συμβεί ένας ανασυνδυασμός και της πιθανότητας να συμβούν δύο ανασυνδυασμοί κοκ μπορούμε να πούμε ότι για κοντινές αποστάσεις όπου η πιθανότητα για περισσότερους από ένα ανασυνδυασμούς είναι μηδέν τότε ο ρυθμός ανασυνδυασμού δηλαδή η γενετική απόσταση ισούται με τον λόγο ανασυνδυασμού recombination fraction  $\theta$ . Εκτός από αυτήν την σχέση που συνδέει τις δύο ποσότητες υπάρχουν και άλλες με λιγότερους περιορισμούς και πιο πολύπλοκες. Το ενδιαφέρον μας για την σχέση των δύο ποσοτήτων έγκειται στο γεγονός ότι ενώ η γενετική απόσταση μας εξυπηρετεί στην δημιουργία γενετικών χαρτών εφόσον μπορεί μια γενετική απόσταση να αθροισθεί με μια άλλη, το recombination fraction είναι πιο εύκολο να υπολογισθεί.

Με βάση τα ανωτέρω δημιουργήθηκαν γενετικοί χάρτες χρησιμοποιώντας οικογένειες αναφοράς με πρώτες τις οικογένειες CEPH (Centre d'Etudes du Polymorphisme Humain) από τη Γαλλία και την Utah και υπολογίζοντας τον αριθμό των ανασυνδυασμών μεταξύ πολλαπλών γενετικών δεικτών (6). Σταδιακά προστέθηκαν και άλλοι δείκτες πιο πληροφοριακοί προσθέτοντας περισσότερη ακρίβεια στους γενετικούς χάρτες. Από τους πιο γνωστούς ολογωνιδωματικούς γενετικούς χάρτες είναι ο χάρτης Marshfield στον οποίο χρησιμοποιήθηκαν οκτώ οικογένειες CEPH και γονοτυπικά δεδομένα από πολυμορφικούς μικροδορυφορικούς δείκτες (7). Η επιλογή των δεικτών (markers) για γονοτύπωση και δημιουργία γενετικών χαρτών είναι σημαντική και με τον ίδιο τρόπο σημαντική είναι η επιλογή δεικτών για μελέτες σύνδεσης. Οι δείκτες αυτοί θα προσφέρουν μεγαλύτερη πληροφορία εάν είναι ετερόζυγοι στους γονείς εφόσον στην αντίθετη περίπτωση δεν θα είναι ορατό εάν υπήρξε ανασυνδυασμός ή όχι στους απογόνους. Επιπλέον θα πρέπει αυτοί οι δείκτες να έχουν ικανή συχνότητα στον πληθυσμό. Οι μικροδορυφορικοί πολυμορφισμοί πληρούν αυτά τα χαρακτηριστικά και επειδή είναι συχνοί στον πληθυσμό και επιπλέον εμφανίζουν μεγάλο βαθμό ετερογένειας με πολλαπλά αλληλία τα οποία καθορίζονται από τις επαναλήψεις των βάσεων τους. Πέρα από τους μικροδορυφορικούς πολυμορφισμούς ως δείκτες μπορούν να χρησιμοποιηθούν

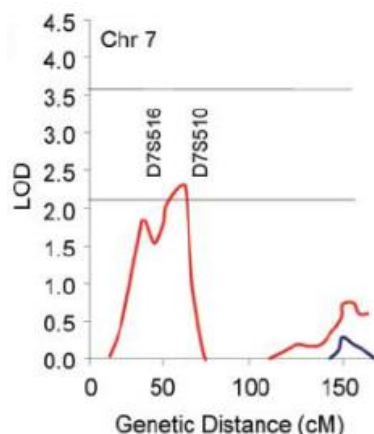
και τα SNPs προκαλώντας όμως μερικές φορές στρεβλώσεις στα αποτελέσματα των μελετών σύνδεσης εάν δεν γίνει έλεγχος των πιθανών περιπτώσεων ύπαρξης ανισορροπίας σύνδεσης.

## 5. ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΝΔΕΣΗΣ

Οι μελέτες γενετικής σύνδεσης χρησιμοποιούνται για να αναγνωρίσουν περιοχές του γονιδιώματος που προδιαθέτουν σε παθήσεις. Η ανάλυση σύνδεσης βασίζεται στο γεγονός ότι κάποιες περιοχές του γονιδιώματος βρίσκονται σε σύνδεση δηλαδή έχουν την τάση να συμμεταβιβάζονται μη τυχαία το οποίο περιγράφεται από την ποσότητα  $\theta$  η οποία θα έχει τιμή  $< 0.5$ . Η ανάλυση σύνδεσης απαιτεί την γονοτύπωση αρκετών μελών μιας οικογένειας η οποία παρουσιάζει σε μέλη της τον μελετώμενο φαινότυπο και την παρατήρηση της συμμεταβίβασης του φαινοτύπου με κάποιον γενετικό τόπο ο οποίος μας προσφέρεται από τους καθορισμένους γενετικούς δείκτες (markers). Προηγείται η δημιουργία του γενεαλογικού δένδρου της οικογένειας και οι εκτιμήσεις των συχνοτήτων του ανασυνδυασμού. Έτσι επιτυγχάνεται η απόδειξη ότι ο συγκεκριμένος γενετικός τόπος ή πιο πιθανά κάποια άλλη κοντινή και συνδεδεμένη σε αυτόν μεταγραφικά ενεργή περιοχή συμμετέχει στην εμφάνιση του φαινοτύπου (4). Οι μελέτες αυτές ονομάζονται μελέτες γενετικής σύνδεσης και εάν αφορούν όλο το εύρος του γονιδιώματος ολογενωμικές μελέτες γενετικής σύνδεσης (Genome Wide Linkage Studies – GWLS ή εν συντομία Genome Scans).

Απόδειξη για τη συσχέτιση αυτή παρουσιάζεται τυπικά με το στατιστικό logarithm of the odds (LOD) score. Το LOD score χρησιμοποιεί την μεθοδολογία του λόγου πιθανοφάνειας για να συγκρίνει την πιθανότητα δύο περιοχές και συγκριμένα το γονίδιο που αναζητάμε και ένας (στην περίπτωση του two-point) ή περισσότεροι (στην περίπτωση του multi-point) γενετικοί δείκτες βρίσκονται σε σύνδεση δηλαδή έχουν  $\theta < 0.5$  ή αυτό το αποτέλεσμα να οφείλεται στην τύχη. Μάλιστα χρησιμοποιείται η μεθοδολογία της μέγιστης πιθανοφάνειας (maximum likelihood) όπου συγκρίνονται οι πιθανοφάνειες του εναλλακτικού ( $\theta < 0.5$ ) και του μηδενικού μοντέλου ( $\theta = 0.5$ ) και ο δεκαδικός λογάριθμος κατανέμεται σε κατανομή της οποίας οι μέγιστες τιμές αποτελούν απόδειξη εναντίον της μηδενικής υπόθεσης. Τιμές μεγαλύτερες από 3 αποτελούν ισχυρή ένδειξη υπέρ της σύνδεσης. Οι τιμές του LOD score στις μελέτες σύνδεσης συνήθως

κατανέμονται ανά χρωμόσωμα σε διάγραμμα όπου στον έναν άξονα (x) βρίσκεται η γενετική απόσταση σε cM και στον άλλον (y) το στατιστικό που χρησιμοποιήθηκε (LOD score) (Εικόνα 2) (8).



Εικόνα 2. Αποτύπωση

LOD score

Σε μια μελέτη γενετικής σύνδεσης θα πρέπει να καθορισθεί εκ των προτέρων το γενετικό μοντέλο συνήθως από μελέτες της ασθένειας σε πληθυσμούς. Πέραν αυτού υπάρχει η πιθανότητα η γενετική αιτιολογία να παρουσιάζει ετερογένεια και διαφορετικοί γενετικοί τόποι να προκαλούν παρόμοιο φαινότυπο. Ένα τέτοιο φαινόμενο προκαλεί μείωση της δύναμης της παραμετρικής ανάλυσης σύνδεσης. Μία λύση σε αυτό το πρόβλημα αποτελεί ο υπολογισμός του LOD score ετερογένειας όπου το LOD score μεγιστοποιείται σε συνάρτηση με μια ακόμη παράμετρο, τον αριθμό των οικογενειών που συνδέονται με αυτόν το γενετικό τόπο.

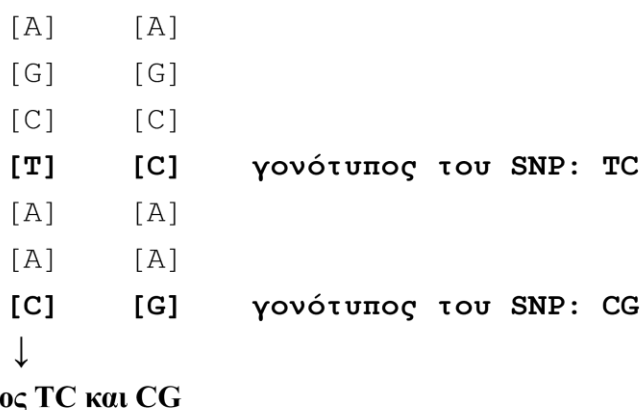
Για σύνθετες παθήσεις όπου πολλαπλά γονίδια μπορεί να συνεισφέρουν στο ρίσκο για την εμφάνιση της υπάρχουν μη παραμετρικές μέθοδοι ανάλυσης που δεν απαιτούν τον καθορισμό του μοντέλου κληρονομής και είναι ελεύθερες υπόθεσης (Model-free). Μία προσέγγιση σε τέτοιου είδους μελέτη είναι με την χρήση ζευγών αδερφών (sibling pairs). Σε ένα τέτοιο ζευγάρι γνωρίζουμε ότι για έναν πολυμορφισμό θα είναι πανομοιότυποι με κάθε αλληλίο



να προέρχεται από τον ίδιο γονέα (identical by descent - IBD) και όχι πανομοιότυποι αλλά από διαφορετικό γονέα (identical by state - IBS) με συγκεκριμένο τρόπο. Δηλαδή η πιθανότητα το ζεύγος να μοιράζεται κανένα, ένα ή δύο IBD αλληλία είναι αντίστοιχα 0.5, 0.25 και 0.25. Αυτές οι προβλέψεις μπορούν να συγκριθούν με τις παρατηρούμενες συχνότητες κοινών IBD δεικτών (9).

## 6. ΧΑΡΤΟΓΡΑΦΗΣΗ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ

Η αλληλουχία των αλληλίων και στα δύο χρωμοσώματα ενός διπλοειδούς κυττάρου ονομάζεται γονότυπος. Η σειρά των βάσεων σε ένα χρωμόσωμα ονομάζεται απλότυπος. Πολλές φορές ο απλότυπος περιέχει για συντομία μόνο τους τόπους που είναι πολυμορφικοί στον πληθυσμό (Εικόνα 3).



Εικόνα 3. Γονότυπος και απλότυπος

Όπως έχει αναφερθεί σε διάφορα σημεία του γονιδιώματος κατά την διάρκεια της μείωσης γίνεται ανασυνδυασμός, σε άλλες περιοχές περισσότερο και σε άλλες λιγότερο. Οι γενετικοί τόποι που αναδυνδύαζονται με τρόπο τυχαίο βρίσκονται σε ισορροπία σύνδεσης ενώ όταν μεταβιβάζονται μαζί με τρόπο μη τυχαίο βρίσκονται σε ανισορροπία σύνδεσης (LD). Η LD λοιπόν είναι η σχέση που έχουν δύο αλληλία σε δύο ή περισσότερους γενετικούς τόπους σε έναν πληθυσμό και περιγράφει την πιθανότητα να υπάρχει ένας συγκεκριμένος απλότυπος πιο

συχνά από όσο θα αναμέναμε από τύχη. Η ποσοτικοποίηση της LD μπορεί να γίνει με διάφορους τρόπους. Μπορούμε να πούμε ότι η πιθανότητα να υπάρχει ένας απλότυπος πχ A-B ισούται με το γινόμενο των πιθανοτήτων να υπάρχει το κάθε αλληλίο ξεχωριστά:

$$P(\text{απλοτυπος A-B}) = P(\text{αλληλίο A}) \times P(\text{αλληλίο B})$$

Εάν τα δύο αλληλία βρίσκονται σε LD είναι δηλαδή συνδεδεμένα τότε:

$$P(\text{απλοτυπος A-B}) = P(\text{αλληλίο A}) \times P(\text{αλληλίο B}) + D$$

Όπου D ο συμπαράγοντας ανισορροπίας (disequilibrium coefficient). Εάν λύσουμε την εξίσωση για τον D τότε αυτός ισούται με την πιθανότητα να υπάρχει ο συγκεκριμένος απλότυπος μείον το γινόμενο των πιθανοτήτων να υπάρχει το κάθε αλληλίο ξεχωριστά:

$$D = p_{AB} - p_A p_B$$

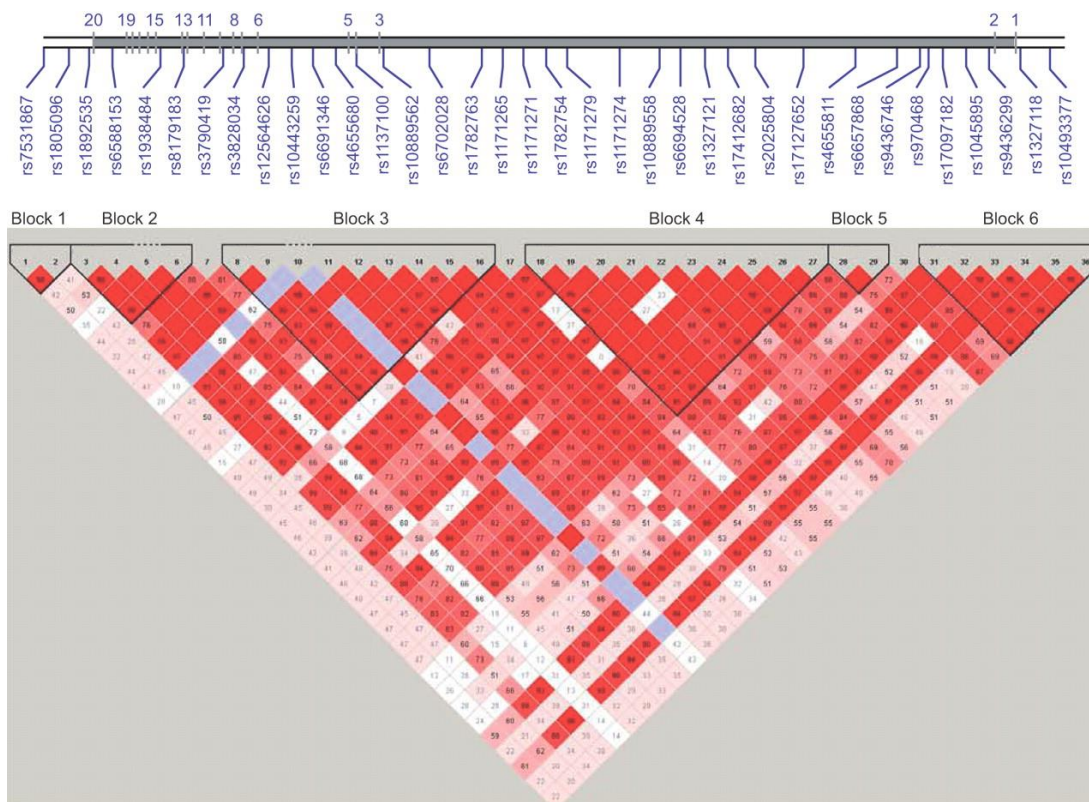
Εάν δύο αλληλία βρίσκονται σε ισορροπία σύνδεσης τότε ο D είναι ίσος με μηδέν. Επειδή αυτό το μέτρο LD επηρεάζεται σημαντικά από την συχνότητα των αλληλίων στον πληθυσμό για την περιγραφή της LD υπάρχει και το Lewontin's D' το οποίο κανονικοποιεί το D παίρνοντας τον λόγο αυτού προς την μέγιστη τιμή που μπορεί να λάβει και έτσι λαμβάνει τιμές από -1 έως ένα. Άλλο μετρικό και ίσως το πιο δημοφιλές είναι ο Pearson correlation coefficient r και συγκεκριμένα το τετράγωνο αυτού  $r^2$  το οποίο ισούται με:

$$r^2 = D^2 / p(1-p)q(1-q)$$

όπου p και q είναι οι συχνότητες των αλληλίων σε κάθε τόπο.

Με την πάροδο του χρόνου και τις πολλαπλές μειώσεις η ανισορροπία σύνδεσης που μπορεί να υπάρχει μεταξύ γενετικών τόπων εκφυλίζεται. Εάν οι τόποι αυτοί είναι αρκετά κοντά η LD μπορεί να παραμένει εάν όχι τότε σταδιακά μπορεί να επέλθει ισορροπία σύνδεσης. Έχει φανεί ότι δεν υπάρχει ομοιομορφία στις τιμές LD μεταξύ ζευγών SNPs σε διάφορες περιοχές. Έτσι προτάθηκε και στη συνέχεια αποδείχθηκε ότι υπάρχουν μετρήσιμα τμήματα γονιδιώματος τα οποία έχουν μικρές μόνο ενδείξεις ότι υπήρξε ανασυνδυασμός ιστορικά και ότι μόνο λίγοι απλότυποι παρατηρούνται σε αυτές. Τα τμήματα αυτά ονομάστηκαν haplotype blocks (Εικόνα 3). Αντίθετα τα σημεία ανάμεσα σε τέτοια τμήμα φαίνεται ότι είναι σημεία με μεγάλη πιθανότητα να συμβεί ανασυνδυασμός και ονομάστηκαν hotspots (10). Τα haplotype blocks μπορεί να έχουν μήκος από λίγες χιλιάδες βάσεις(kb) έως και 300kb. Με σκοπό την οπτικοποίηση των τμημάτων αυτών και την πιο αποδοτική μελέτη τους δεδομένου ότι η αριθμητική πληροφορία της ανά ζεύγη μετρήσιμης LD είναι πολύ μεγάλη χρησιμοποιείται συνήθως το λογισμικό haplview (1) το οποίο υπολογίζει το μέγεθος της ανισορροπίας σύνδεσης και την αναπαριστά με τρόπο εύκολα κατανοητό και διαδραστικό (Εικόνα 4). Η σημαντικότητα της παρουσίας των haplotype blocks δηλαδή η αυξημένη LD μεταξύ SNP (υψηλό  $r^2$ ) βρίσκεται στην επιλογή κατάλληλων πολυμορφισμών-δεικτών. Εάν δύο πολυμορφισμοί έχουν αυξημένη LD τότε μας προσφέρουν σχεδόν την ίδια πληροφορία. Έτσι σκοπός είναι να βρεθούν οι κατάλληλοι δείκτες που ανήκουν ουσιαστικά σε περιοχές με μεγάλη πιθανότητα ανασυνδυασμού δηλαδή σε διαφορετικά haplotype blocks.

Τέτοιοι πολυμορφισμοί ονομάζονται SNPs σηματοδότες (tagging-SNPs – tSNPs) γιατί σηματοδοτούν όλους τους άλλους πολυμορφισμούς της περιοχής. Έτσι για λόγους οικονομίας και υπολογιστικής αποδοτικότητας επιλέγουμε τον μικρότερο και αντιπροσωπευτικότερο αριθμό tSNPs για την χαρτογράφηση μιας περιοχής που θα χρησιμοποιηθεί στις μελέτες γενετικής συσχέτισης.



Εικόνα 4. Παράδειγμα Harplotype view (11). Απεικονίζονται τετράγωνα με την ανα ζεύγη μέτρηση της LD. Τα κόκκινα τετράγωνα υποδεικνύουν υψηλή LD, τα μπλε μη σημαντική και τα άσπρα χαμηλή ή μηδενική. Τα harplotype blocks επισημαίνονται με τρίγωνα μαύρης γραμμής

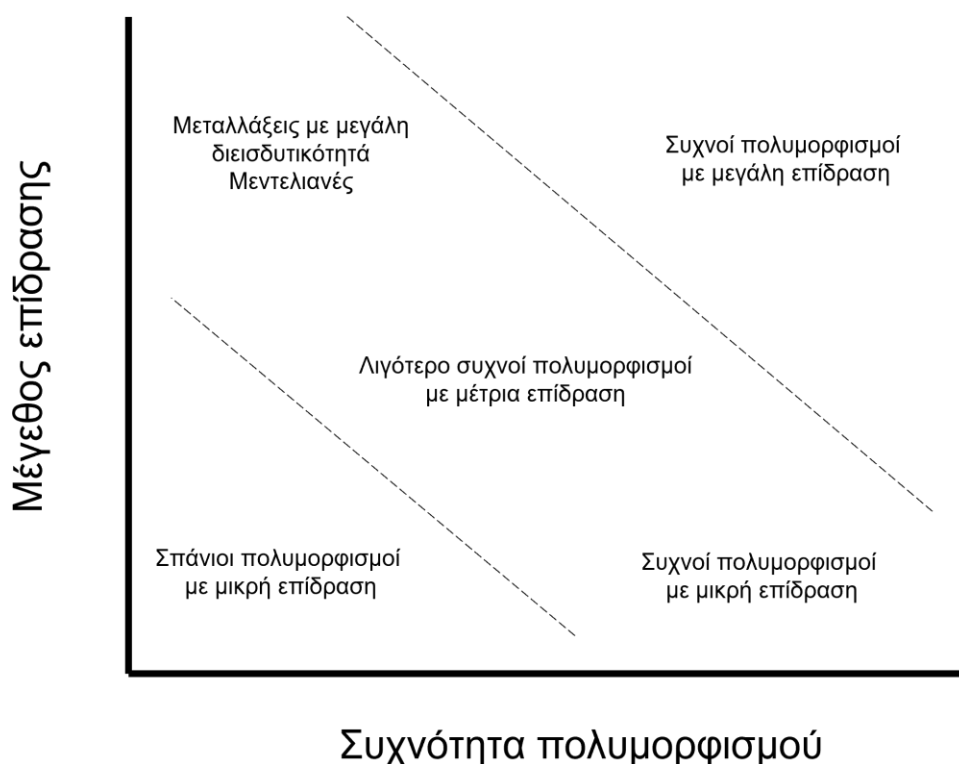
## 7. ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ

Πρώτο βήμα για την διεξαγωγή μιας μελέτης γενετικής συσχέτισης είναι η δημιουργία ενός αρχικού διερευνητικού πάνελ όπου όλοι οι πολυμορφισμοί θα μελετηθούν και από αυτούς θα επιλεγούν οι κατάλληλοι tSNPs μετά από τον υπολογισμό της LD. Το πάνελ αυτό ονομάζεται discovery panel και μπορεί να περιλαμβάνει γονοτύπωση όλων των πολυμορφισμών ή μερική γονοτύπωση. Πλέον υπάρχουν διαθέσιμα τέτοια πάνελ από διάφορους πληθυσμούς που αξιοποιούν διαφορετικές μεθόδους για χρήση σε μελέτες. Το επόμενο βήμα περιλαμβάνει την επιλογή των λιγότερων αλλά και πιο αντιπροσωπευτικών tSNPs ώστε να καλύπτουν όλο το γονιδίωμα ενδιαφέροντος όπως περιεγράφηκε.

Η πιο απλή μορφή μελέτης γενετικής συσχέτισης βασίζεται σε έναν σχεδιασμό ασθενών-μαρτύρων. Μελετάται η παρουσία ή μη ενός πολυμορφισμού σε ένα σύνολο πληθυσμού το οποίο έχει τον φαινότυπο ενδιαφέροντος καθώς και σε έναν πληθυσμό μαρτύρων που δεν εμφανίζει τον φαινότυπο ενδιαφέροντος. Η επιλογή του πολυμορφισμού που θα μελετηθεί βασίζεται σε κάποια υπόθεση που προέρχεται από λειτουργικές μελέτες ή βιολογικές υποθέσεις (hypothesis-driven). Από την άλλη ο πληθυσμός που θα επιλεγεί θα πρέπει να είναι όσο πιο ομοιογενής μπορεί συνήθως με την εφαρμογή αυστηρών κλινικών κριτηρίων. Μετά την γονοτύπωση ασθενών και μαρτύρων για ένα SNP με δύο αλληλίες καταγράφονται οι συχνότητες εμφάνισης των πιθανών συνδυασμών των αλληλίων οι οποίες μπορούν να τοποθετηθούν σε έναν 2x3 πίνακα και να αξιολογηθούν στατιστικά με την χρήση του Pearson chi-square (goodness-of-fit) το οποίο μπορεί να απορρίψει ή όχι την μηδενική υπόθεση της μη συσχέτισης της συχνότητας των γονοτύπων με την κατάσταση του ατόμου δηλαδή εάν είναι ασθενής ή μάρτυρας. Θα έχει προηγηθεί η εξέταση εάν οι συχνότητες των γονοτύπων βρίσκονται σε HWE συγκρίνοντας τις παρατηρούμενες τιμές με τις αναμενόμενες από την HWE εξίσωση τιμές. Η αναγνώριση ενός πολυμορφισμού που μπορεί να σχετίζεται με έναν φαινότυπο εάν και πολύ αποκαλυπτική μας δίνει λίγες πληροφορίες για το είδος αυτής της σχέσης, εάν δηλαδή πρόκειται για άμεση σχέση ή έμμεση σχέση με την έννοια της παρουσίας ενός τόπου σε LD με αυτόν. Περισσότερες πληροφορίες μπορούν να αντληθούν, με έναν παρόμοιο τρόπο σε μελέτες ασθενών μαρτύρων αλλά με την γονοτύπωση ενός υποψήφιου γονιδίου. Σε αυτήν την περίπτωση ισχύουν οι ίδιες αρχές με την μελέτη συσχέτισης με έναν πολυμορφισμό αλλά γίνεται γονοτύπωση πολλαπλών SNPs σε μια συγκεκριμένη περιοχή και η στατιστική ανάλυση ακολουθεί την μεθοδολογία της λογαριθμικής παλινδρόμησης όπου λαμβάνονται υπόψη οι επιδράσεις του κάθε δείκτη στην πιθανότητα να εμφανίζεται ο φαινότυπος που μας ενδιαφέρει καθώς και η αλληλεπίδραση μεταξύ των δεικτών. Η κατά βήματα (stepwise) ανάλυση των μοντέλων αξιολογεί το κατά πόσο επηρεάζει την εμφάνιση του φαινοτύπου ο κάθε δείκτης ξεχωριστά ή απλά βρίσκονται σε LD μεταξύ τους (12).

## 8. ΜΕΛΕΤΕΣ ΕΥΡΕΙΑΣ ΓΟΝΙΔΙΩΜΑΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ

Στις μελέτες γενετικής σύνδεσης ο στόχος είναι η αναζήτηση γενετικών τόπων που εμφανίζονται σπάνια αλλά έχουν μεγάλη επίδραση στην εμφάνιση του φαινοτύπου που εξετάζεται και για αυτόν τον λόγο εξετάζεται σε οικογένειες με αρκετά μέλη που πάσχουν. Από την άλλη στις μελέτες συσχέτισης υποψήφιων SNP ή υποψήφιων γονιδίων η μελέτη καθοδηγείται από μία υπόθεση. Το 1996 με στόχο την διερεύνηση σύνθετων παθήσεων προτάθηκε η υπόθεση ότι συχνά εμφανιζόμενοι σε πληθυσμούς πολυμορφισμοί μπορεί να επηρεάζουν κοινές παθήσεις (common variant-common disease theory)(Εικόνα 5).



Εικόνα 5. Σχηματική αναπαράσταση της θεωρίας common variant-common disease. Στόχος των GWAS είναι οι πολυμορφισμοί στην κάτω δεξιά γωνία του διαγράμματος με μεγάλη συχνότητα στον πληθυσμό και μικρή επίδραση. Προσαρμοσμένο από (2)

Για να μελετηθούν αυτοί οι κοινοί πολυμορφισμοί θα πρέπει να μελετηθούν σε μεγάλους πληθυσμούς ασθενών και μαρτύρων όπου οι σπάνιοι με σημαντική επίδραση πολυμορφισμοί θα εμφανίζονται σπάνια οπότε θα είναι δύσκολο να αποδειχθεί η συμμετοχή τους αλλά οι κοινοί πολυμορφισμοί θα εμφανίζονται συχνά και θα μπορεί να υπολογισθεί η επίδρασή τους. Για να γίνει

κάτι τέτοιο θα πρέπει το σύνολο των πολυμορφισμών του γονιδιώματος να αποτυπωθεί ή τουλάχιστον ένα σετ που να είναι αντιπροσωπευτικό με βάση την γενετική σύνδεση (13). Έτσι δημιουργήθηκαν projects με σκοπό την χαρτογράφηση των όσο πιο αντιπροσωπευτικών SNPs σε ολόκληρο το γονιδίωμα με σκοπό την χρήση τους σε ανάλογες μελέτες (HarMap project, 1000 genomes project)(14, 15). Οι μελέτες αυτού του είδους, δηλαδή αυτές που στοχεύουν στην ανεύρεση κοινών πολυμορφισμών που σχετίζονται με το ρίσκο εμφάνισης κοινών σύνθετων παθήσεων σε ολόκληρο το γονιδίωμα ονομάζονται μελέτες ευρείας γονιδιωματικής συσχέτισης ή ολιγονιδιωματικές μελέτες συσχέτισης (Genome-Wide Association Studies – GWAS).

Με την ανάπτυξη της τεχνολογίας και την δυνατότητα πλέον της γονοτύπωσης με τρόπο οικονομικά και χρονικά αποδοτικό με την χρήση ειδικών chips η διενέργεια GWAS τα τελευταία 15 χρόνια αποτέλεσε βασικό εργαλείο αναζήτησης της γενετικής των σύνθετων παθήσεων. Η βασική αρχή των μελετών είναι απλή. Επιλέγεται ένας πληθυσμός με την πάθηση που μελετάται και ένας παρόμοιος πληθυσμός ως μάρτυρες. Γίνεται συλλογή γενετικού υλικού και γονοτύπωση του κατάλληλου σετ πολυμορφισμών. Εάν κάποιο SNP εμφανίζεται συχνότερα στους ασθενείς από ότι στους μάρτυρες τότε λέμε ότι υπάρχει συσχέτιση του SNP (ή κάποιου άλλου γενετικού τόπου που συνδέεται με αυτόν) με την πάθηση. Προφανώς όμως και το σύνολο των διαδικασιών δεν είναι και τόσο απλό.

## **8.1 Τεχνολογία γονοτύπωσης**

Οι GWAS έγιναν εφικτές διότι εφευρέθηκε η κατάλληλη τεχνολογία γονοτύπωσης με βάση μικροαλληλουχίες πάνω σε ειδικά chip με την δυνατότητα τύπωσης περισσότερων από ένα εκατομμύριο SNPs. Υπάρχουν δύο βασικές πλατφόρμες που έχουν χρησιμοποιηθεί για αυτόν τον σκοπό των εταιρειών Illumina (San Diego, CA) και Affymetrix (Santa Clara, CA). Οι δύο αυτές πλατφόρμες χρησιμοποιούν διαφορετική τεχνολογία. Ένα ακόμη χαρακτηριστικό που θα πρέπει να έχουν οι πλατφόρμες είναι η εξειδίκευση τους σχετικά με τον πληθυσμό. Ανάλογα με τον πληθυσμό που μελετάται θα πρέπει να

χρησιμοποιηθεί αντίστοιχη πλατφόρμα που θα έχει καλύτερη γενομική κάλυψη σε αυτόν εφόσον διαφορετικοί πληθυσμοί και ανάλογα με την «παλαιότητα» τους έχουν διαφορετικό βαθμό ανασυνδυασμού και κατά συνέπεια μικρότερη LD μεταξύ διαφορετικών SNPs.

## 8.2 Σχεδιασμός μελέτης

Ένα βασικό στοιχείο των μελετών γενετικής συσχέτισης είναι ο ακριβής καθορισμός του φαινότυπου που μελετάται. Αυτός μπορεί να είναι μια ποσοτική τιμή η οποία θα προσθέσει δύναμη στα ευρήματα της μελέτης αλλά συχνά πρόκειται για μια κατηγορική τιμή. Εάν και η ποιοτική φύση των δευτέρων μειώνουν την δύναμη αυτό μπορεί να υπερκερασθεί με τον καθορισμό κριτηρίων με ακριβή στοιχεία που επιδεικνύουν μεγάλη ειδικότητα και μικρή διακύμανση μεταξύ διαφορετικών εξεταστών. Η εφαρμογή αυστηρών κριτηρίων παίζει σημαντικό ρόλο κυρίως σε πολυκεντρικές μελέτες οι οποίες είναι και ο κανόνας εφόσον απαιτείται μεγάλος αριθμός ασθενών για επιτευχθεί στατιστική σημαντικότητα σε γενωμικό επίπεδο.

## 8.3 Στατιστική ανάλυση

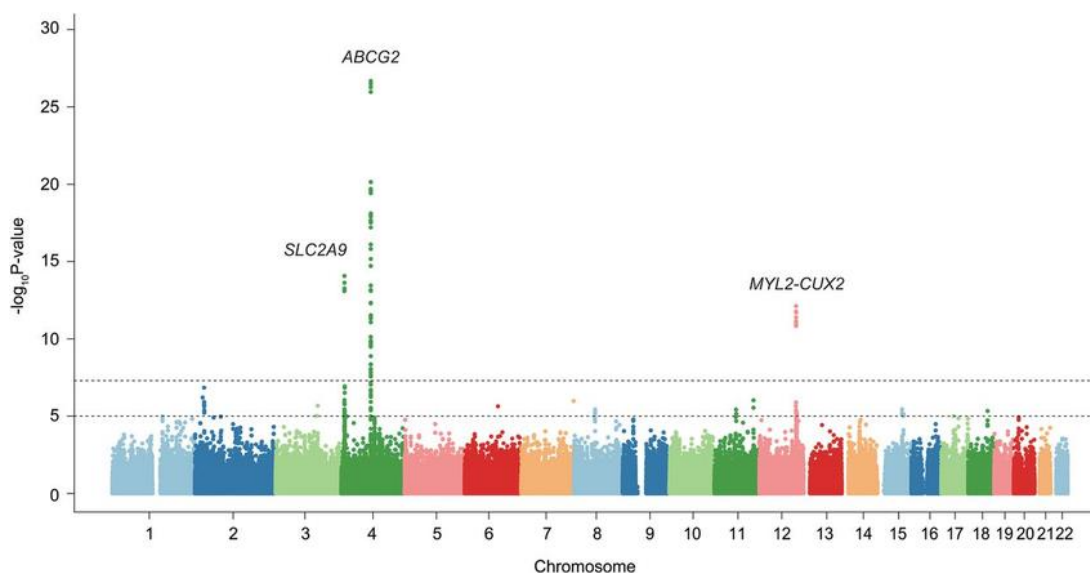
Μετά τον ακριβή καθορισμό του πληθυσμού και των μαρτύρων και την γονοτύπωση ακολουθεί η στατιστική ανάλυση των ευρημάτων που βασίζεται στην ανάλυση μιας σειράς στατιστικών τεστ για κάθε SNP ξεχωριστά και την συσχέτιση του με τον φαινότυπο. Η στατιστική μεθοδολογία που θα επιλεγεί εξαρτάται από το είδος του χαρακτηριστικού που μελετάται. Εάν πρόκειται για ποσοτική μεταβλητή η ανάλυση γίνεται με την χρήση γενικευμένου γραμμικού μοντέλου, συνήθως ανάλυση διακύμανσης ANOVA με κατηγορική μεταβλητή την γονοτυπική κλάση. Στην περίπτωση των κατηγορικών μεταβλητών ακολουθείται η χρήση πινάκων συνάφειας ή λογιστική παλινδρόμηση. Σε κάθε περίπτωση σημαντικό ρόλο παίζει και ο τρόπος που τα δεδομένα κωδικοποιούνται εάν δηλαδή θα αναλυθούν τα αλληλία ξεχωριστά ή ο κάθε γονότυπος ξεχωριστά (ή γονοτυπικές κλάσεις) και σε αυτήν την περίπτωση η ομαδοποίηση των γονοτύπων κάτω από ένα μοντέλο κληρονόμησης (υπολειπόμενο, επικρατές ή αθροιστικό). Με σκοπό την μείωση συστηματικών αποκλίσεων (bias) γίνεται ταυτόχρονα και έλεγχος για συμπαράγοντες που



μπορεί να εισάγουν bias στην μελέτη όπως η ηλικία το φύλο, κλινικοί παράγοντες που σχετίζονται με την πάθηση και η διαστρωμάτωση του πληθυσμού λόγω εθνικότητας. Συγκεκριμένα για το τελευταίο χρησιμοποιούνται λογισμικά που συγκρίνουν την συχνότητα των πολυμορφισμών σε γενωμικό επίπεδο του εξεταζόμενου πληθυσμού με αυτές διαφορετικών εθνικών ομάδων του HarMap. Τέλος λόγω των πολλαπλών συγκρίσεων σε μια GWAS θα πρέπει να γίνει διόρθωση για πολλαπλές συγκρίσεις η οποία είναι δυνατή με διάφορες μεθόδους όπως η διόρθωση Bonferroni όπου το επίπεδο του σφάλματος α διαιρείται με τον αριθμό των στατιστικών τεστ που διενεργούνται ή εφαρμόζοντας δοκιμασίες παραλλαγών (permutation testing) εναλλάσσοντας τυχαία και πολλαπλώς τους γονοτύπους με τα άτομα. Μία τελευταία προσέγγιση είναι αυτή της επίτευξης μιας σημαντικότητας σε γενωμικό επίπεδο η οποία προκαθορίζεται με βάση τον πληθυσμό ελέγχου και βασίζεται στην κατανομή της LD στο γονιδίωμα του συγκεκριμένου πληθυσμού. Το επίπεδο αυτό για Ευρωπαϊκούς πληθυσμούς έχει υπολογισθεί στο  $7.2e-8$ .

#### 8.4 Αποτύπωση αποτελεσμάτων

Τα αποτελέσματα συνήθως αποτυπώνονται σε διάγραμμα που ονομάζεται Manhattan plot. Στο διάγραμμα αυτό στον οριζόντιο άξονα τοποθετούνται τα χρωμοσώματα συνήθως σε σειρά και με διαφορετικό χρώμα με τρόπο ανάλογο του μεγέθους τους και στον κάθετο άξονα ο αρνητικός λογάριθμος του p-value της στατιστικής δοκιμασίας που έχει χρησιμοποιηθεί. Ο αρνητικός λογάριθμος του p-value κάθε SNP αποτυπώνεται με μια βούλα. Λόγω του μεγάλου αριθμού των SNPs αυτά μπορεί να μην ξεχωρίζουν μεταξύ τους όμως τα πιο σημαντικά θα έχουν το μικρότερο p-value και ο αρνητικός τους λογάριθμος θα έχει την μεγαλύτερη τιμή η οποία θα ξεχωρίζει στο διάγραμμα (Εικόνα 6).



Εικόνα 6. Παράδειγμα Manhattan plot (1)

## 9. ΜΕΤΑ ΑΝΑΛΥΣΗ ΓΕΝΕΤΙΚΩΝ ΜΕΛΕΤΩΝ

### 9.1 Μετα-ανάλυση ολογενωμικών μελετών σύνδεσης (GWLS)

Η μετα-ανάλυση μελετών, η ανασκόπηση δηλαδή μελετών με κοινό αντικείμενο και η κοινή τους ανάλυση με διάφορες στατιστικές μεθόδους με σκοπό την εξαγωγή κοινών συμπερασμάτων με στατιστική σημαντικότητα, έχει προσφέρει στις επιστήμες σημαντικό όφελος. Ομοίως στις μελέτες γενετικής σάρωσης από νωρίς δημιουργήθηκε η ανάγκη μετα-ανάλυσής τους. Με αυτή τη προσέγγιση εξετάζεται η επαναληψιμότητα των ευρημάτων και ενισχύεται ή απορρίπτεται η πραγματική συσχέτιση τους με την πάθηση. Η πρώτη μετα-ανάλυση μελετών γενετικής σύνδεσης έγινε το 1998 (16) με την χρησιμοποίηση της μεθόδου του Fisher (17) ο οποίος απέδειξε ότι ο φυσικός λογάριθμος των  $p$  values από  $k$  ανεξάρτητα τεστ σημαντικότητας ακολουθούν μια κατανομή  $\chi^2$  με  $2k$  βαθμούς ελευθερίας. Οι GWLS όμως έχουν ένα βασικό χαρακτηριστικό και αυτό είναι οι διαφορετικές μέθοδοι που μπορεί να χρησιμοποιηθούν σε διαφορετικές μελέτες. Σε κάθε μελέτη μπορεί να χρησιμοποιηθεί διαφορετικός αριθμός και δομή οικογενειών, διαφορετικοί markers και διαφορετική στατιστική ανάλυση. Επιπλέον δεν υπάρχει ένας κοινός τρόπος αναφοράς των ευρημάτων. Σε κάποιες μελέτες μπορεί να αναφέρονται όλα τα χρωμοσώματα ενώ σε άλλες μόνο αυτά που έχουν στατιστικά σημαντικά αποτελέσματα. Η αναφορά όλων των δεικτών και των σκορ τους δεν είναι ιδιαίτερα συχνή. Έχοντας αυτά υπόψη κατανοούμε ότι η μετα-ανάλυση των μελετών σύνδεσης με τον πιο αξιόπιστο τρόπο που θα ήταν η κοινή ανάλυση των δεικτών από όλες τις μελέτες μαζί είναι σχεδόν αδύνατη. Για να ξεπεραστούν αυτά τα εμπόδια περιγράφηκε το 1999 η μέθοδος Genome Scan Meta-Analysis (GSMA) (18). Σε γενικές γραμμές η μέθοδος κατατάσσει το LOD score (ή άλλο στατιστικό ελέγχου) σε μια σάρωση και στη συνέχεια συγκρίνει την κατάταξη μεταξύ των σαρώσεων.

Αρχικά έγινε ο διαχωρισμός του γονιδιώματος σε τμήματα που ονομάστηκαν Bins. Λαμβάνοντας υπόψη τους περιορισμούς που μπορεί να προκαλέσει η τμηματοποίηση του γονιδιώματος σε μικρότερες ή μεγαλύτερες περιοχές επιλέχθηκε το μέγεθος των 30cM ώστε να μειωθεί η πιθανότητα να καταγράφονται στο ίδιο Bin σημαντικά αποτελέσματα από διαφορετικές μελέτες αλλά και η πιθανότητα να μην υπάρχει marker σε κάποιο από τα διαστήματα. Μοιράζοντας το γονιδίωμα σε bins περίπου 30cM δημιουργούνται συνολικά

120 bins. Σε κάθε Bin καταγράφεται το υψηλότερο διαθέσιμο στατιστικό (LOD score, z, p-value) από κάθε μελέτη. Σε κάθε μελέτη ξεχωριστά το κάθε Bin λαμβάνει μια κατάταξη (rank) με το πιο σημαντικό να λαμβάνει την μεγαλύτερη. Η διαδικασία αυτή επαναλαμβάνεται για όλες τις σαρώσεις. Σε πολλές περιπτώσεις μπορεί να προκύψουν τιμές με μηδενικό (ή αρνητικό) LOD score ή απώλεια δεδομένων λόγω ελλιπούς αναφοράς τους. Σε αυτές τις περιπτώσεις τα bins λαμβάνουν ίδιους ranks σύμφωνα με την εξίσωση  $(120-\chi+1)/2$ , όπου  $\chi$  ο αριθμός των Bins που βρίσκονται πάνω από αυτούς στην κατάταξη. Στη συνέχεια οι βαθμοί κάθε Bin από όλες τις σαρώσεις αθροίζονται και εκτελείται στατιστική δοκιμασία όπου η μηδενική υπόθεση είναι ότι δεν υπάρχει γενετικός τόπος ρίσκου στο συγκεκριμένο bin και οι βαθμοί ανακατανέμονται τυχαία. Για  $m$  μελέτες και  $n$  bins η πιθανότητα το άθροισμα των βαθμών να είναι ίση με μια τιμή  $R$  περιγράφεται στην εικόνα 7 όπως καταγράφηκε στην αρχική αναφορά της μεθόδου.

$$\begin{aligned}
 P\left(\sum_{i=1}^m X_i = R\right) &= 0 \quad \text{for } R < m \\
 (18) \quad &= \frac{1}{n^m} \sum_{k=0}^d (-1)^k \binom{R-kn-1}{m-1} \\
 &\quad \times \binom{m}{k} \quad \text{for } m \leq R \leq mn \\
 &= 0 \quad \text{for } R > mn,
 \end{aligned}$$

Στην εφαρμογή της μεθόδου αργότερα από τον Levinson η ταξινόμηση των βαθμών  $R$  έγινε σε αύξουσα σειρά και χρησιμοποιήθηκε ο μέσος όρος των βαθμών του κάθε Bin ( $R_{avg}$ ) παρά το άθροισμα τους. Επιπλέον ως παράγοντας βαρύτητας των μελετών (weighting) χρησιμοποιήθηκε η ποσότητα  $\sqrt{N}$  (affected cases) όπου  $N$  ο αριθμός των ασθενών στην μελέτη (19).

Οι γενετικές σαρώσεις μπορεί να προκαλέσουν αντικρουόμενα αποτελέσματα με αποτέλεσμα να εμφανίζεται μεγάλη ετερογένεια στις μετα-αναλύσεις. Πηγές ετερογένειας είναι οι διαφορετικές μεθοδολογίες ή οι τρόποι γονοτύπωσης, ο σχεδιασμός η διαφορετική δομή των οικογενειών της μελέτης ή ακόμη και από τύχη. Για την μελέτη της ετερογένειας στις γενετικές GWLS με το σκεπτικό ότι η περιγραφή της υπάρχουσας ετερογένειας σε μια μετα-ανάλυση προσφέρει πολύ σημαντικές πληροφορίες για την πραγματική σχέση των αποτελεσμάτων με την νόσο προτάθηκε το 2005 η χρήση των στατιστικών  $Q$ ,  $H_a$  και  $B$ . Το

στατιστικό Q ορίζεται ως το άθροισμα των τετραγώνων των αποκλίσεων των ranks κάθε bin από κάθε μελέτη από τον μέσο όρο των ranks και δίνεται από τον τύπο:

$$Q = \sum_{i=1}^s w_i (R_i - R)^2, \quad (20)$$

Όπου  $R_i$  είναι ο rank του Bin που μελετάται για την μελέτη  $i$ ,  $R$  είναι ο μέσος όρος των ranks και  $w_i$  ο παράγοντας βαρύτητας για την μελέτη  $i$ . Το στατιστικό  $H_a$  ορίζεται ως το άθροισμα των απόλυτων διαφορών του κάθε rank από τον μέσο rank και περιγράφεται από τον τύπο:

$$H_a = \sum_{i=1}^s w_i (|R_i - R|). \quad (20)$$

Το στατιστικό B ορίζεται ως το άθροισμα των διακριτών απόλυτων διαφορών των ranks μεταξύ των μελετών λαμβάνοντας υπόψη όλα τα πιθανά ζεύγη bins μεταξύ των μελετών και περιγράφεται από τον τύπο:

$$B = \sum_{i,j=1}^s w_i w_j (|R_i - R_j|). \quad (20)$$

Για την αξιολόγηση της στατιστικής σημαντικότητας των στατιστικών τιμών Q,  $H_a$  και B προτάθηκε στην ίδια μελέτη η εφαρμογή μεθοδολογίας Monte Carlo κατά την οποία οι ranks των bins αντιμετωπίζονται τυχαία (permutation) για πολλαπλές φορές (50.000) και υπολογίζεται το στατιστικό ετερογένειας. Με τα αποτελέσματα σχηματίζεται η κατανομή αυτών. Το επίπεδο στατιστικής σημαντικότητας υπολογίζεται ως το ποσοστό των τυχαίων στατιστικών που υπερβαίνει το παρατηρούμενο (για υψηλή ετερογένεια) ή είναι μικρότερο από αυτό (για χαμηλή ετερογένεια). Την ίδια χρονιά παρουσιάσθηκε το λογισμικό HEGESMA (Heterogeneity and genome search meta-analysis) το οποίο διευκολύνει την

ανωτέρω διαδικασία μετα-ανάλυσης γενομικών σαρώσεων υπολογίζοντας και την ετερογένεια (21). Ως μετεξέλιξη του HEGESMA προτάθηκε το 2007 η μεθοδολογία και το αντίστοιχο λογισμικό METa-analysis in Ranked DIScovery datasets (METRADISC) ένα γενικευμένο μετα-αναλυτικό μοντέλο με στόχο την συνένωση δεδομένων από μελέτες που βασίζονται στην αναζήτηση-παραγωγή κάποιας θεωρίας και τον παράλληλο έλεγχο της ετερογένειας μεταξύ των μελετών (22). Η αρχική πρακτική περιγραφή της μεθόδου αφορούσε μελέτες μικροσυστοιχιών όπου κάθε μελέτη μπορεί να μην έχουν μελετήσει τις ίδιες ακριβώς βιολογικές παραμέτρους ή τον ίδιο αριθμό αυτών. Έτσι η μεθοδολογία βασίζεται στην προσαρμογή των βαθμών κατάταξης των βιολογικών παραμέτρων με τον πολλαπλασιασμό τους με τον παράγοντα  $n_{max}/n_i$  όπου  $n_{max}$  ο μέγιστος αριθμός παραμέτρων που εξετάστηκαν σε οποιαδήποτε μελέτη και  $n_i$  ο αριθμός παραμέτρων που εξετάστηκαν στην συγκεκριμένη μελέτη. Με βάση τους προσαρμοσμένους βαθμούς υπολογίζεται ο μέσος βαθμός ( $R^*$ ) και το στατιστικό της ετερογένειας ( $Q^*$ ) των οποίων η στατιστική σημαντικότητα συγκρίνεται με καμπύλες κατανομής που παράγονται μετά από πολλαπλές συγκρίσεις με την μέθοδο Monte Carlo. Με δεδομένο ότι σε κάθε μελέτη είναι πιθανό να έχουν εξετασθεί διαφορετικές βιολογικές παράμετροι (σε είδος και αριθμό) ο σχηματισμός μηδενικών κατανομών μετά την τυχαία ανακατανομή των βαθμών γίνεται για κάθε ξεχωριστή κλάση πληροφορίας (information class). Αυτό σημαίνει ότι κάθε παράγοντας που εξετάστηκε σε συγκεκριμένο αριθμό μελετών συγκρίνεται με την μηδενική κατανομή από αυτές τις συγκεκριμένες μελέτες. Ο μέγιστος αριθμός information classes μπορεί να είναι  $2^s - 1$  όταν το πλήθος των μελετών είναι  $s$ .

## **9.2 Μετα-ανάλυση ολογενωμικών μελετών συσχέτισης (GWAS)**

Η ολοκλήρωση μιας μελέτης GWA μπορεί να καταλήξει στην ανεύρεση ενός γενετικού τόπου που συσχετίζεται με την υπο μελέτη νόσο, όμως αυτό δεν είναι το τέλος. Η ερμηνεία αυτών των αποτελεσμάτων μπορεί να είναι προβληματική. Θα πρέπει να διερευνηθεί κατά πόσο αυτά τα αποτελέσματα είναι αναπαραγώγιμα, εάν μπορούν τα ευρήματα να γενικευθούν και σε ποιον πληθυσμό και φυσικά κατά πόσο προέκυψαν λόγω πραγματικής συσχέτισης ή απλά από τύχη. Αυτά τα ερωτήματα αρχικά προσεγγίζονται με

μελέτες αναπαραγωγής (replication studies) και σε δεύτερο επίπεδο με τις μετα-αναλύσεις των GWAS.

Οι GWAS είναι μελέτες ελεύθερες υπόθεσης. Μάλιστα παράγουν υποθέσεις. Η γενετική επίδραση των κοινών αλληλίων είναι μικρή γεγονός που καθιστά τις μεμονωμένες GWAS λιγότερο ισχυρές στην αναγνώριση τους ενώ από την άλλη η στατιστική σύνθεση των πληροφοριών από πολλαπλές ανεξάρτητες μελέτες είναι ικανή να αυξήσει την ισχύ των μελετώ και να μειώσει την πιθανότητα ανεύρεσης ψευδώς θετικών αποτελεσμάτων (23).

Η διενέργεια μια μετα-ανάλυσης GWAS είναι μια διαδικασία που απαιτεί πολύ καλή οργάνωση και σχεδιασμό και συνήθως συνεργασία μεταξύ ερευνητικών ομάδων. Ιδεατά κάθε μελέτη από διαφορετικές ομάδες θα στηρίζονταν στον σχεδιασμό της μελλοντικής τους σύνθεσης ενσωματώνοντας στον σχεδιασμό τους κοινά στοιχεία διατηρώντας ταυτόχρονα την ανεξαρτησία τους. Παρόλα αυτά η αναδρομική φύση των μετα-αναλύσεων εισάγει αναπόφευκτα μεθοδολογικούς περιορισμούς στις μετα-αναλύσεις (24). Μία επιπλέον πρόκληση για την διενέργεια μιας μετα-ανάλυσης είναι η αποφυγή εισαγωγής Bias στην μελέτη. Αυτό μπορεί να προκύψει από την επιλογή των μελετών (selection bias), την περιορισμένη διαθεσιμότητα των δεδομένων από τις ξεχωριστές μελέτες, την επιλεκτική αναφορά των αποτελεσμάτων και κάποιες φορές την μη δημοσίευση κάποιων μελετών (publication bias). Με σκοπό την αποφυγή εισαγωγής μεροληψίας (bias) στην μελέτη θα πρέπει να υπάρχει ένα αναλυτικό πρωτόκολλο που θα προβλέπει τα κριτήρια επιλογής και απόρριψης των μελετών, τον καλά ορισμένο φαινότυπο της νόσου που μελετάται και τα βήματα που θα ακολουθηθούν ώστε να μην υπάρξει επιλογή των πληροφοριών που θα μελετηθούν και θα αναλυθούν. Επιπλέον θα πρέπει να προβλέπονται οι μέθοδοι που θα χρησιμοποιηθούν για την στατιστική ανάλυση των δεδομένων.

Υπάρχουν αρκετές μέθοδοι που χρησιμοποιούνται επιτυχώς τα τελευταία χρόνια με πιο δημοφιλή την μέθοδο των σταθερών επιδράσεων (fixed effects). Σε αυτό το μοντέλο υπάρχει η υπόθεση ότι η αληθινή επίδραση κάθε αλληλίου είναι η ίδια σε κάθε διαφορετικό σύνολο δεδομένων (από τις διαφορετικές μελέτες) και χρησιμοποιεί την μέθοδο της αναστροφής διακύμανσης για τον καθορισμό της βαρύτητας του κάθε αλληλίου (25). Ως παράγοντας βαρύτητας της κάθε μελέτης σε αυτήν την περίπτωση χρησιμοποιείται ο αντίστροφος της τετραγωνικής ρίζας της διακύμανσης της. Από την άλλη το μοντέλο των τυχαίων επιδράσεων χρησιμοποιεί τις περισσότερες φορές τον εκτιμητή των DerSimonian και Laird και είναι πιο κατάλληλο για την περιγραφή της γενίκευσης των αποτελεσμάτων

(26). Αυτές και επιπλέον μέθοδοι έχουν πλέον ενσωματωθεί σε λογισμικά πακέτα που απαιτούν εξειδικευμένους χρήστες αλλά υποβοηθούν σημαντικά την στατιστική ανάλυση των μετα-αναλύσεων των GWAS εφόσον αυτές έχουν σχεδιαστεί και εκτελεστεί με ακρίβεια καταναλώνοντας σημαντικό επιστημονικό και υπολογιστικό δυναμικό.

Εκτός των κλασικών παραμετρικών μεθοδολογιών ως μετα-αναλυτική μεθοδολογία μπορεί να χρησιμοποιηθεί και η μη παραμετρική προσέγγιση της meta-analysis for ranked discovery datasets με την χρήση του λογισμικού METRADISC (22) και METRADISC-XL (27) το οποίο είναι επέκταση του πρώτου και δεν έχει περιορισμό στον αριθμό των εισαγόμενων μελετών και τον αριθμό των προσομοιώσεων. Η δυνατότητα παραγωγής ενός μέσου rank με βάση προσαρμοσμένους ranks ανάλογα με τον αριθμό μελετών που αξιολόγησαν ένα SNP από το σύνολο αυτών και η αναζήτηση της στατιστικής σημαντικότητας σε ξεχωριστές information classes καθιστά το λογισμικό ικανό να διαχειρισθεί πολλαπλές GWAS με απώλεια δεδομένων.

## 10.ΕΠΙΓΕΝΕΤΙΚΗ

Όπως αναφέρθηκε στο προηγούμενο κεφάλαιο ένα ακόμη επίπεδο ελέγχου της γονιδιακής έκφρασης είναι ο επιγενετικός έλεγχος. Το DNA είναι συμπυκνωμένο στη μορφή της χρωματίνης στον πυρήνα σε δύο μορφές. Την πιο «σφιχτή» δομή της ετεροχρωματίνης και την πιο «χαλαρή» της ευχρωματίνης με βασικό δομικό στοιχείο τα νουκλεοσώματα που αποτελούνται από βάσεις DNA τυλιγμένες γύρω από οκταμερή ιστονών. Αυτές οι δομές καθιστούν το γενετικό υλικό μη προσβάσιμο σε ρυθμιστικούς παράγοντες και οι βασικοί μηχανισμοί που ρυθμίζουν τη πρόσβαση είναι η μεθυλίωση του DNA και η τροποποίηση των ιστονών. Ο επιγενετικός αυτός έλεγχος επιτρέπει την έκφραση γονιδίων σε διαφορετικούς ιστούς σε διάφορες φάσεις της ανάπτυξης και τη σίγαση τους αργότερα ενώ είναι επιρρεπής και σε περιβαλλοντικούς παράγοντες.

Η μεθυλίωση του DNA γίνεται σε δινουκλεοτίδια CpG τα οποία βρίσκονται συχνότερα σε κάποιες περιοχές που ονομάζονται νησίδες CpG και σχετίζονται με το μεγαλύτερο ποσοστό των υποκινητών γονιδίων. Αυτή η συσχέτιση υπονοεί έναν σημαντικό μεταγραφικό ρόλο των νησίδων CpG. Η μεθυλίωση του DNA αποτελεί έναν καλό βιοδείκτη για την μελέτη της επιγενετικής ρύθμισης και της ποικιλότητας μεταξύ ατόμων. Οι μελέτες αυτές ξεκίνησαν κυρίως σε ογκολογικούς ασθενείς όπου περιεγράφηκαν επιγενετικές

μεταβολές σχεδόν σε κάθε είδος καρκίνου με χαρακτηριστική την υπερμεθυλίωση ογκοκατασταλτικών γονιδίων. Ταυτόχρονα όμως η ανεύρεση και γονιδιακών μεταλλάξεων στις αντίστοιχες περιοχές καθιστά δύσκολη την τεκμηρίωση αιτιότητας αυτών των αλλαγών και οδηγεί στην υπόθεση ότι ο φαινότυπος είναι αποτέλεσμα διάδρασης των γενετικών και επιγενετικών επιδράσεων. Οι μεθοδολογίες για την ανίχνευση της μεθυλίωσης του DNA εμφανίζουν ποικιλομορφία και εξαρτώνται από το είδος του ιστού που θα καλλιεργηθεί, εάν η αλληλουχία του DNA που χρησιμοποιείται είναι γνωστή ή όχι, την ποσότητα της μεθυλίωσης που απαιτείται για να ανιχνευθεί (η ευαισθησία της μεθόδου) καθώς και εάν η μελέτη θα αφορά υποψήφια γονίδια ή όλο το γονιδίωμα. Σε αυτήν την δεύτερη περίπτωση σημαντικό ρόλο έπαιξε η ανάπτυξη της τεχνολογίας των μικροσυστοιχιών (microarrays).

Οι μικροσυστοιχίες είναι ειδικά σχεδιασμένα chips που περιέχουν χιλιάδες μικροσκοπικά τμήματα DNA ή ολιγονουκλεοτίδια με γνωστή αλληλουχία. RNA από τον ιστό που μελετάται επισημαίνεται και μετατρέπεται σε cDNA ή cRNA και τοποθετείται στο chip. Εάν μια περιοχή του μεταγραφώματος είναι συμπληρωματική μιας αλληλουχίας στο chip θα ενωθεί με αυτό και θα γίνει ορατό και μάλιστα με τρόπο ανάλογο της αφθονίας στην οποία υπάρχει. Η θέση της ορατούς σήμανσης αποκαλύπτει και το ολιγονουκλεοτίδιο που αφορά. Με αυτήν την ανάλυση μικροσυστοιχιών βρίσκουμε λοιπόν εάν μια συγκεκριμένη αλληλουχία είναι παρούσα στο γονιδίωμα και μάλιστα και σε μια ποσοτικοποιημένη μορφή. Έτσι μπορούμε να συγκρίνουμε την έκφραση και την ποσότητα αυτής σε διαφορετικά πειραματικά περιβάλλοντα ή μεταξύ καταστάσεων ασθενείας ή υγείας (ασθενών-μαρτύρων).

Η στατιστική ανάλυση των δεδομένων από μελέτες μικροσυστοιχιών αποτελεί μια από τις πιο σημαντικές στιγμές της διάδρασης μεταξύ της στατιστικής και της βιοπληροφορικής και περιλαμβάνει πολυσύνθετα μοντέλα ανάλυσης μέσω λογισμικών πακέτων συνήθως του προγράμματος R.



# ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ: Η ΝΟΣΟΣ PARKINSON

## 1. Εισαγωγή

Η νόσος Parkinson αποτελεί μία από τις συχνότερες νευροεκφυλιστικές νόσους της μεγάλης ηλικίας. Επηρεάζει αρκετά εκατομμύρια ανθρώπους παγκοσμίως και αποτελεί μείζον πρόβλημα τόσο της πρωτοβάθμιας φροντίδας υγείας, όσο και της καθημερινής νευρολογικής κλινικής πρακτικής. Τα τελευταία χρόνια έχει μεγάλη πρόοδος ως προς την κατανόηση της παθοφυσιολογίας της νόσου, ενώ μελέτες έχουν εστιαστεί περισσότερο στην αιτιολογική θεραπεία της νόσου, καθώς οι μέχρι σήμερα θεραπείες αποτελούν συμπτωματική αντιμετώπιση της νόσου.

## 2. Διάγνωση – Κλινική εικόνα

Η νόσος Parkinson αποτελεί μία χρόνια, προοδευτικά εξελισσόμενη νευροεκφυλιστική νόσο του ΚΝΣ, με κύριο χαρακτηριστικό της την συμπτωματολογία από την δυσλειτουργία του εξωπυραμιδικού συστήματος. Νεότερες μελέτες έχουν καθιερώσει τρία στάδια στην νόσο, το προκλινικό στάδιο κατά το οποίο άρχεται η νευροεκφύλιση, το πρόδρομο στάδιο κατά το οποίο υπάρχει συμπτωματολογία η οποία δεν είναι αρκετή για την θεμελίωση της διάγνωσης και τέλος το κλινικό με πλήρη συμπτωματολογία της νόσου. Η νόσος χαρακτηρίζεται από κινητικά και μη κινητικά συμπτώματα με προεξάρχοντα τα κινητικά (28).

Τα κινητικά συμπτώματα της νόσου Parkinson οφείλονται κατ' εξοχήν στην δυσλειτουργία του εξωπυραμιδικού συστήματος. Τρόμος ηρεμίας, βραδυκινησία, δυσκαμψία και αστάθεια βάδισης αποτελούν τα μείζονα κινητικά συμπτώματα της νόσου, για τα οποία συνήθως ο ασθενής ζητάει ιατρική βοήθεια. Σε αυτά προστίθενται η υπομιμία του προσώπου και η υποφωνία τα οποία είναι ελάσσονα σε σχέση τα προηγούμενα. Το κύριο σύμπτωμα που οδηγεί συνήθως τον ασθενή στην αναζήτηση ιατρικής βοήθειας είναι ο τρόμος. Πρόκειται για ρυθμικό τρόμο ηρεμίας που προσβάλλει συνήθως κάποιο από τα άκρα, συνηθέστερα τα άνω. Σε συνδυασμό με τον τρόμο οι ασθενείς αναφέρουν βραδύτητα στην εκτέλεση κινήσεων και διαταραχή στη βάδιση. Κατά την εξέταση γίνεται αντιληπτός ο τρόμος ηρεμίας, ενώ η παθητική κίνηση των άκρων αναδεικνύει δυσκαμψία του τύπου του

οδοντωτού τροχού και του μολυβδοσωλήνα (29). Απότοκο της δυσκαμψίας είναι η βραδυκινησία, η οποία συνήθως γίνεται αντιληπτή από την εκτέλεση επαναλαμβανόμενων κινήσεων από τον ασθενή.

Η βάδιση των ασθενών με νόσο Parkinson είναι διαταραγμένη. Οι ασθενείς παρουσιάζουν καμπτοκορμία, κάμψη του κορμού προς τα εμπρός, και μικρά συρόμενα βήματα, τα οποία αποτελούν τα χαρακτηριστικά του παρκινσονικού τύπου βάδισης.

Πολλοί ασθενείς παρουσιάζουν δυστονία, με ανώμαλες κάμψεις αρθρώσεων, κυρίως των άκρων, λόγω ανώμαλων συσπάσεων μυϊκών ομάδων που οφείλονται στα μειωμένα επίπεδα ντοπαμίνης.

Τα κινητικά συμπτώματα, συνήθως με την πρόοδο της νόσου, έχουν διακυμάνσεις και οι ασθενείς βιώνουν περιόδους με καλύτερη κινητικότητα, τις λεγόμενες On περιόδους, και αντίστοιχες με χειρότερη, τις off. Χαρακτηριστικό των off περιόδων είναι το λεγόμενο πάγωμα, κατά το οποίο οι ασθενείς αδυνατούν να κινηθούν.

<b>Κινητικά συμπτώματα</b>	<b>Μη κινητικά συμπτώματα</b>
<ul style="list-style-type: none"><li>• Τρόμος ηρεμίας</li><li>• Δυσκαμψία</li><li>• Βραδυκινησία</li><li>• Διαταραχή της βάδισης</li><li>• Δυστονία</li></ul>	<ul style="list-style-type: none"><li>• Ανοσμία</li><li>• Δυσκοιλιότητα</li><li>• Ναυτία</li><li>• Σιελόρροια</li><li>• Ορθοστατική υπόταση</li><li>• Υπεριδρωσία</li><li>• Αυξημένη παραγωγή σμήγματος</li><li>• Διαταραχές του ύπνου</li><li>• Αιμωδίες</li><li>• Διάχυτα άλγη</li><li>• Κατάθλιψη</li><li>• Άγχος</li><li>• Απάθεια</li><li>• Αβουλησία</li><li>• Άνοια</li></ul>

Πίνακας 3: Συμπτώματα της νόσου Parkinson

Μαζί τα κινητικά συμπτώματα που αναφέρθηκαν, η νόσος Parkinson εμφανίζει και μια σειρά μη κινητικών συμπτωμάτων που περιλαμβάνουν αισθητικές διαταραχές, διαταραχές από τον ύπνο, ανοσμία, συμπτώματα από δυσλειτουργία του αυτόνομου και διαταραχές των ανώτερων νοητικών λειτουργιών με σημαντικότερες την κατάθλιψη και την άνοια. Στον πίνακα που ακολουθεί φαίνονται συνοπτικά τα συμπτώματα της νόσου (30).

### **3. Επιδημιολογία**

Η νόσος Parkinson είναι μία νόσος κατ' εξοχήν της μεγάλης ηλικίας. Επηρεάζει άτομα συνήθως άνω των 65 ετών με ετήσια επίπτωση στις ανεπτυγμένες χώρες 160/100.000 με μικρή υπεροχή στους άνδρες. Η νόσος εκδηλώνεται και σε άτομα μικρότερα των 65 ετών καθώς και σε νεαρούς ενήλικες, οπότε εκεί πρόκειται για νεανικές μορφές με υψηλή την πιθανότητα γενετικού υποβάθρου. Γενικά η εμφάνιση της νόσου είναι συχνότερη με την αύξηση της ηλικίας (31).

### **4. Παθοφυσιολογία**

Το κύριο παθοφυσιολογικό χαρακτηριστικό της νόσου είναι μειωμένη ντοπαμινεργική μετάδοση. Η ντοπαμίνη παράγεται κατά κύριο λόγο στους ντοπαμινεργικούς νευρώνες της μέλαινας ουσίας του εγκεφάλου και από εκεί με την μελαινοραβδωτή οδό μεταφέρεται στο ραβδωτό σώμα. Στη νόσο Parkinson παρατηρείται εκφύλιση και κυτταρικός θάνατος των ντοπαμινεργικών νευρώνων της μέλαινας ουσίας και επομένως ανεπάρκεια στην παραγωγή της ντοπαμίνης. Η εκφύλιση των νευρώνων είναι απότοκο της συσσώρευσης α-συνουκλεΐνης, λόγω διαταραχής στην αποδόμηση της από τα λυσοσώματα. Η συσσώρευση αυτή οδηγεί σε δυσλειτουργία των νευρώνων και δευτεροπαθώς σε κυτταρικό θάνατο. Οι συσσωρεύσεις της α-συνουκλεΐνης είναι ορατές σε ιστοπαθολογικές τομές νεκροτομικών παρασκευασμάτων. Είναι χαρακτηριστικό ότι όταν εμφανίζεται η συμπτωματολογία και αναζητά ο ασθενής ιατρική βοήθεια έχει απολεσθεί το 80% περίπου των ντοπαμινεργικών νευρώνων (32).

Αιτιολογικά η νόσος φαίνεται ότι αποτελεί τυπικό παράδειγμα πολυπαραγοντικής νόσου και σύνθετης γενετικής πάθησης. Η συχνότητα της νόσου αυξάνεται με την αύξηση της ηλικίας άρα η ηλικία αποτελεί έναν παράγοντα κινδύνου ενώ ταυτόχρονα έχουν αναγνωρισθεί περιβαλλοντικοί παράγοντες που μπορεί να προκαλέσουν συμπτώματα

συμβατά με τη νόσο με πιο γνωστή τη δράση χημικών παρασιτοκτόνων. Την ίδια στιγμή έχουν αναγνωρισθεί γενετικοί παράγοντες στην νόσο Parkinson.

## 5. Διαφορική διάγνωση

Η διαφορική διάγνωση της νόσου του Parkinson περιλαμβάνει μία πλειάδα νοσημάτων, τα οποία έχουν στην συμπτωματολογία-σημειολογία τους παρκινσονισμό. Η νόσος έχει το χαρακτηριστικό ότι συνήθως στα αρχικά στάδια πλαγιώνει την συμπτωματολογία σε μία πλευρά και φυσικά δεν περιλαμβάνει γνωσιακές διαταραχές εξ αρχής. Επίσης δεν υφίστανται διαταραχές από άλλα λειτουργικά συστήματα. Οι πτώσεις εμφανίζονται στην πορεία της νόσου καθώς και η δυσλειτουργία του αυτόνομου. Τα παραπάνω είναι σημαντικά για την διαφοροδιάγνωση από τα άτυπα εξωπυραμιδικά σύνδρομα. Σε αντίθεση με την νόσο Parkinson τα άτυπα σύνδρομα εμφανίζουν από νωρίς γνωσιακές διαταραχές, η προϊούσα υπερπυρηνική παράλυση και η φλοιοβασική εκφύλιση, ενώ συνυπάρχουν παρεγκεφαλιδικές διαταραχές και πτώσεις εξ αρχής, η ατροφία πολλαπλών συστημάτων και η προϊούσα υπερπυρηνική παράλυση. Στην διαφορική διάγνωση και κυρίως σε νεότερα άτομα περιλαμβάνεται η νόσος Wilson και άλλες νόσοι εναπόθεσης μετάλλων στα βασικά γάγγλια (29). Επίσης θα πρέπει να αποκλειστεί το ενδεχόμενο αγγειακού παρκινσονισμού με μαγνητική τομογραφία εγκεφάλου. Στην διαφορική διάγνωση επίσης περιλαμβάνονται και ανοϊκές διαταραχές στις οποίες εκδηλώνεται εξωπυραμιδική συμπτωματολογία όπως η νόσος Alzheimer, η μετωποκροταφική άνοια και η άνοια με σωμάτια Lewy, καθώς και άλλα σπανιότερα νευροεκφυλιστικά νοσήματα όπως κληρονομικές αταξίες και μεταβολικές διαταραχές όπως η νόσος Gaucher. Στον παρακάτω πίνακα φαίνεται συνοπτικά η διαφορική διάγνωση της νόσου Parkinson.

- Προϊούσα υπερπυρηνική παράλυση
- Φλοιοβασική εκφύλιση
- Ατροφία πολλαπλών συστημάτων
- Αγγειακός παρκινσονισμός
- Νόσος Wilson
- Νόσος Hallevorden Spatz
- Νόσος Alzheimer

- Μετωποκροταφική άνοια
- Άνοια με σωμάτια Lewy
- Κληρονομικές αταξίες
- Μεταβολικά νοσήματα

---

Πίνακας 4: Διαφορική διάγνωση της νόσου Parkinson

## **6. Διάγνωση**

Η διάγνωση της νόσου Parkinson είναι κατά βάση κλινική και συνεπικουρείται από εργαστηριακές εξετάσεις. Η κλινική εικόνα και το ιστορικό της έναρξης της συμπτωματολογίας είναι σημαντικά και στην πλειονότητα των περιπτώσεων αρκούν για να τεθεί η διάγνωση. Από το τέλος της δεκαετίας του 90 θεσπίστηκαν κλινικά κριτήρια για την διάγνωση της νόσου με βασικό απαιτούμενο την παρουσία παρκινσονισμού τα οποία δέχτηκαν διάφορες τροποποιήσεις κατά την διάρκεια των ετών. Τα πιο συχνά χρησιμοποιούμενα κριτήρια πλέον είναι αυτά της Movement Disorders Society τα οποία περιλαμβάνουν απαιτούμενα και υποστηρικτικά κριτήρια (33). Αντίστοιχα υπάρχουν κλινικές κλίμακες αξιολόγησης της βαρύτητας της νόσου που χρησιμοποιούνται για την παρακολούθηση και εκτίμηση της αποτελεσματικότητας της θεραπείας όπως η MDS-UPDRS (34). Σε κάθε περίπτωση όμως οι ασθενείς πρέπει να υποβάλλονται σε νευροαπεικονιστικό έλεγχο, καλύτερα με μαγνητική τομογραφία, για αποκλεισμό αγγειακών και άλλου είδους βλαβών. Η απεικόνιση με σπινθηρογράφημα βασικών γαγγλίων, στο οποίο αναδεικνύεται η δυσλειτουργία στο ραβδωτό σώμα είναι απαραίτητη μόνο σε περιπτώσεις αμφιβολίας για την διάγνωση. Δεν αποτελεί εξέταση ρουτίνας και φυσικά δεν είναι απαραίτητη όταν είναι έκδηλος ο παρκινσονισμός. Τέλος σε υποψία νόσου Wilson θα πρέπει να μετρώνται τα επίπεδα χαλκού σε ορό και συλλογή ούρων 24ώρου και σερουλοπλασμίνης ορού (35).

## **7. Γενετική της νόσου Parkinson**

Μπορούμε να διαχωρίσουμε τους γενετικούς παράγοντες που σχετίζονται με τη νόσο Parkinson σε δύο κατηγορίες. Η μία αφορά τους σπάνιους γενετικούς παράγοντες με Μεντελιανή κληρονομηση που όπως αναφέρθηκε στο πρώτο μέρος μπορούν να αναζητηθούν με μελέτες γενετικής σύνδεσης και η δεύτερη πιο συχνά αλληλία ρίσκου που

για την αναγνώριση τους απαιτούνται μεγάλες μελέτες και εκτεταμένη στατιστική ανάλυση (36).

Ορόσημο στις γενετικές μελέτες στη νόσο Parkinson αποτέλεσε η ανακάλυψη του γονιδίου της συνουκλεΐνης (SNCA) (37) αποτελώντας την αρχή πολλών ερευνών που βασίζονται σε οικογένειες με αποτέλεσμα πλέον περισσότερα από είκοσι γονίδια να θεωρούνται ότι προκαλούν την νόσο (Πίνακας 5). Βέβαια το να ονομάσουμε αυτές τις μορφές της νόσου μονογονιδιακές, αν και το κάνουμε, είναι κάπως άτοπο εφόσον βλέπουμε πολλαπλές διαφορετικές κλινικές εικόνες ακόμη και σε άτομα με την ίδια μετάλλαξη, διαφορετικές ηλικίες έναρξης και την επίδραση περιβαλλοντικών παραγόντων ακόμη να διαδραματίζει κάποιο ρόλο. Παρόλα αυτά η προσφορά αυτών των ανακαλύψεων είναι πολύ σημαντική γιατί προσφέρει την δυνατότητα της μελέτης και αναπαραγωγής των αποτελεσμάτων των μεταλλάξεων σε πειραματόζωα με σκοπό την ανεύρεση των παθοφυσιολογικών μονοπατιών που μπορεί να εμπλέκονται. Πρόκληση στο συγκεκριμένο είδος μελετών εξαιτίας της σπανιότητας των μεταλλάξεων αυτών είναι η συγκέντρωση του κατάλληλου δείγματος για γονοτύπωση.

Γονίδιο PARK	Τοποθεσία	Γονίδιο
PARK1/4	4q21.3-q22	SNCA
PARK2	6q25.2-27	PRKN
PARK3	2p13	unknown
PARK5	4p13	UCH-L1
PARK6	1p36.12	PINK1
PARK7	1p36.23	DJ-1
PARK8	12q12	LRRK2
PARK9	1p36	ATP13A2
PARK10	1p32	unknown
PARK11	2q36-q37	GIGYF2
PARK12	Xq21-q25	unknown
PARK13	2p13.1	HTRA2
PARK14	22q13.1	PLA2G6
PARK15	22q11.2-qter	FBXO7
PARK16	1q32	unknown
PARK17	16q12	VPS35
PARK18	3q27-qter	EIF4G1
PARK19	1p31.3	DNAJC6
PARK20	21q22.11	SYNJ1
PARK21	3q22.1	DNAJC13
PARK22	7p11.2	CHCHD2

Πίνακας 5: Γονίδια σχετιζόμενα με τη νόσο Parkinson

Στα περιστατικά της λεγόμενης σποραδικής μορφής της νόσου η διερεύνηση ενός γενετικού παράγοντα ρίσκου ήταν για πολλά χρόνια κάτι όχι τόσο προφανές και μάλιστα ιδιαίτερα δύσκολο. Με βάση υποθέσεις που βασίζονταν σε παθοφυσιολογικά μονοπάτια στοχεύονταν αλληλία (υποψήφια αλληλία) στα οποία γινόταν γονοτύπωση σε μελέτες συσχέτισης. Ήδη έως το 2012 όπου δημιουργήθηκε η γενετική βάση PD Gene (<http://www.pdgene.org/>) (38) περισσότερα από 800 δημοσιευμένα άρθρα με μελέτες γενετικής συσχέτισης είχαν αναγνωρισθεί και σε συνδυασμό με δεδομένα από GWAS περισσότερα από 7 εκατομμύρια αλληλία είχαν μετα-αναλυθεί. Οι μελέτες GWAS φυσικά προχώρησαν τις γενετικές μελέτες στη νόσο Parkinson πολλά βήματα παραπέρα. Η πρώτη GWAS στη νόσο Parkinson έγινε το 2009 με δεδομένα από 5000 ασθενείς και 9000 μάρτυρες (39) ενώ δέκα χρόνια αργότερα έχουν αναγνωρισθεί τουλάχιστον 90 τέτοια αλληλία παλαιά και νέα με δεδομένα από 37000 ασθενείς και 1,4 εκατομμύρια μάρτυρες εξηγώντας με αυτόν τον τρόπο μόνο ένα μέρος της υπολογιζόμενης γενετικής συμμετοχής στην νόσο (40).

## 8. Θεραπεία

Μέχρι σήμερα δεν υπάρχει αιτιολογική θεραπεία της νόσου Parkinson. Η θεραπεία είναι συμπτωματική και στοχεύει στην βελτίωση της ποιότητας ζωής των ασθενών και την επιβράδυνση της νευροεκφύλισης.

Ακρογωνιαίος λίθος στην θεραπεία της νόσου είναι η υποκατάσταση της ελλείπουσας ντοπαμίνης. Η υποκατάσταση αυτή γίνεται με χορήγηση λεβοντόπας από του στόματος. Η λεβοντόπα διέρχεται τον αιματοεγκεφαλικό φραγμό και μεταβολίζεται σε ντοπαμίνη και δρα αγωνιστικά στους ντοπαμινεργικούς υποδοχείς. Χορηγείται σε πολλές διασπαρμένες δόσεις ημερησίως οι οποίες εξαρτώνται από τη βαρύτητα της νόσου. Η λεβοντόπα μπορεί να χορηγείται μαζί με καρβιντόπα για πρόληψη της περιφερικής αποδόμησης της λεβοντόπας.

Η μακροχρόνια χορήγηση λεβοντόπας έχει ως αποτέλεσμα την εμφάνιση υπερκινησιών οι οποίες είναι χορειακόμορφες κινήσεις των άκρων, του κορμού και της

κεφαλής. Συνήθως εμφανίζονται στη μέγιστη στάθμη στο πλάσμα της λεβοντόπας οπότε και χαρακτηρίζονται ως peak dose υπερκινησίες. Η υπερκινησίες αποτελούν σημαντικό πρόβλημα στην θεραπευτική της νόσου καθώς επηρεάζουν σημαντικά την ποιότητα ζωής του ασθενούς. Στην στρατηγική της αντιμετώπισης τους περιλαμβάνονται η χορήγηση του αντιϊκού αμανταδίνη και η μεγαλύτερη διασπορά των δόσεων στον χρόνο. Οι ανεπιθύμητες ενέργειες από την χορήγηση της λεβοντόπας αφορούν ψυχωσικόμορφες εκδηλώσεις και διαταραχές της αρτηριακής πίεσης με την μορφή της υπότασης.

Σε προχωρημένα στάδια της νόσου και στην προσπάθεια για μεγαλύτερο θεραπευτικό όφελος από την χορηγούμενη δόση λεβοντόπας χορηγούνται και αναστολείς της COMT, ενζύμου που αποδομεί την ντοπαμίνη στην συναπτική σχισμή. Παρόμοιας δράσης είναι και οι αναστολείς της MAO-B που και αυτοί δύνανται να χορηγηθούν.

Σε πολύ προχωρημένα στάδια της νόσου, όταν οι υπερκινησίες και τα φαινόμενα off είναι μη αντιμετωπίσιμα με την από του στόματος θεραπεία μπορεί η λεβοντόπα να χορηγηθεί μέσω γαστροστομίας σε συνεχή έγχυση στην νήστιδα. Η χορήγηση με αυτό τον τρόπο εξαλείφει τις υπερκινησίες και μειώνονται δραστικά τα off φαινόμενα. Παρόμοια αποτελέσματα έχει και η συνεχής υποδόρια έγχυση απομορφίνης, η οποία είναι αγωνιστής της ντοπαμίνης.

Σε νεότερους ασθενείς προτιμάται η χορήγηση αγωνιστών της ντοπαμίνης. Οι αγωνιστές της ντοπαμίνης έχουν παρόμοια δράση με την λεβοντόπα, αλλά δεν προτιμώνται σε μεγαλύτερες ηλικίες λόγω αυξημένων ψυχωσικών παρενεργειών.

Τέλος στην θεραπεία της νόσου Parkinson έχει θέση η εν τω βάθει εγκεφαλική διέγερση. Πρόκειται για παρεμβατική μέθοδο εισαγωγής ηλεκτροδίων στα βασικά γάγγλια μέσω των οποίων δίδονται ηλεκτρικά ερεθίσματα. Η αποτελεσματικότητα της μεθόδου είναι μεγάλη και απαιτείται μικρή μόνο δόση λεβοντόπας για την συμπλήρωση της θεραπείας.



## **Β ΜΕΡΟΣ: ΕΙΔΙΚΟ**

# ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ: ΜΕΤΑ ΑΝΑΛΥΣΗ ΟΛΟΓΕΝΩΜΙΚΩΝ ΜΕΛΕΤΩΝ ΣΥΝΔΕΣΗΣ (GWLS)

## 1. ΕΙΣΑΓΩΓΗ

Στο πρώτο μέρος του πειραματικού τμήματος της διατριβής διενεργήθηκε μετα-ανάλυση των ολογενωμικών μελετών γενετικής σύνδεσης (GWLS) που αφορούν την νόσο Parkinson. Τα στάδια που ακολουθήθηκαν είναι αρχικά συστηματική αναζήτηση των σχετικών μελετών, εφαρμογή προκαθορισμένων κριτηρίων ένταξης και αποκλεισμού, εξαγωγή των απαραίτητων στοιχείων (με βασικά δεδομένα τα διαγράμματα των στατιστικών (LOD scores) από κάθε μελέτη, ο διαχωρισμός του γονιδιώματος σε bins ίσου περιόδου μεγέθους, η εξαγωγή των πιο σημαντικών στατιστικών τιμών σε κάθε bin από τα διαθέσιμα διαγράμματα και τέλος η εισαγωγή των διαθέσιμων τιμών σε λογισμικό μη παραμετρικής ανάλυσης. Τα ανωτέρω στάδια θα αναλυθούν στη συνέχεια με λεπτομέρεια.

## 2. ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΣ

### 2.1 Συστηματική αναζήτηση μελετών

Έγινε συστηματική αναζήτηση των ολογενωμικών μελετών γενετικής σύνδεσης με την χρήση συνδυασμού λέξεων κλειδίων στην βάση δεδομένων PubMed. Ο συνδυασμός λέξεων που χρησιμοποιήθηκε ήταν : ((Parkinson's) OR Parkinson) AND ("whole genome scan" OR "genomic screen" OR "genome wide" OR "genome-wide" OR "genome scan" OR "genomic scan" OR "genomic screen" OR "LOD score" OR "NPL score" OR "genome-wide linkage analysis"). Τα κριτήρια που επιπλέον επιλέχθηκαν στην αναζήτηση ήταν «Human» και «English». Το χρονικό διάστημα αναζήτησης ήταν έως 31 Δεκ 2019. Επιπλέον αναζητήθηκαν οι βιβλιογραφικές αναφορές των εξαγώγιμων μελετών καθώς και άρθρων ανασκόπησης επι του θέματος για αναζήτηση επιπλέον μελετών.

### 2.2 Εφαρμογή κριτηρίων ένταξης και αποκλεισμού

Ως κριτήρια ένταξης των μελετών καθορίστηκαν τα εξής: α. διεξαγωγή σε ανθρώπινο πληθυσμό, β. η μελέτη αφορούσε αποκλειστικά ασθενείς με νόσο Parkinson

σύμφωνα με σαφή και αναγνωρισμένα κλινικά κριτήρια γ. έγινε γονοτύπωση πολυμορφικών δεικτών με πυκνότητα τουλάχιστον 30cM κατά μήκος του γονιδιώματος, δ. υπήρχε διαθεσιμότητα των απαραίτητων δεδομένων σύνδεσης (δημοσίευση στο διαδίκτυο, παροχή από τους συγγραφείς της μελέτης ή εξαγωγή από δημοσιευμένα γραφήματα), ε. δεν υπήρχε αλληλοεπικάλυψη των δειγμάτων.

Ως κριτήρια αποκλεισμού ορίστηκαν: α. τμηματικές σαρώσεις του γονιδιώματος, β. μελέτες υποψήφιων περιοχών, γ. μη καλά ορισμένη νόσος (παρκινσονισμός, συσχέτιση με συγκεκριμένα κλινικά κριτήρια ή ηλικία έναρξης), δ. σε περίπτωση αλληλοεπικάλυψης των δειγμάτων επιλέχθηκε η μελέτη με το μεγαλύτερο δείγμα ή πιο πρόσφατη ή αυτή με τα περισσότερα διαθέσιμα δεδομένα.

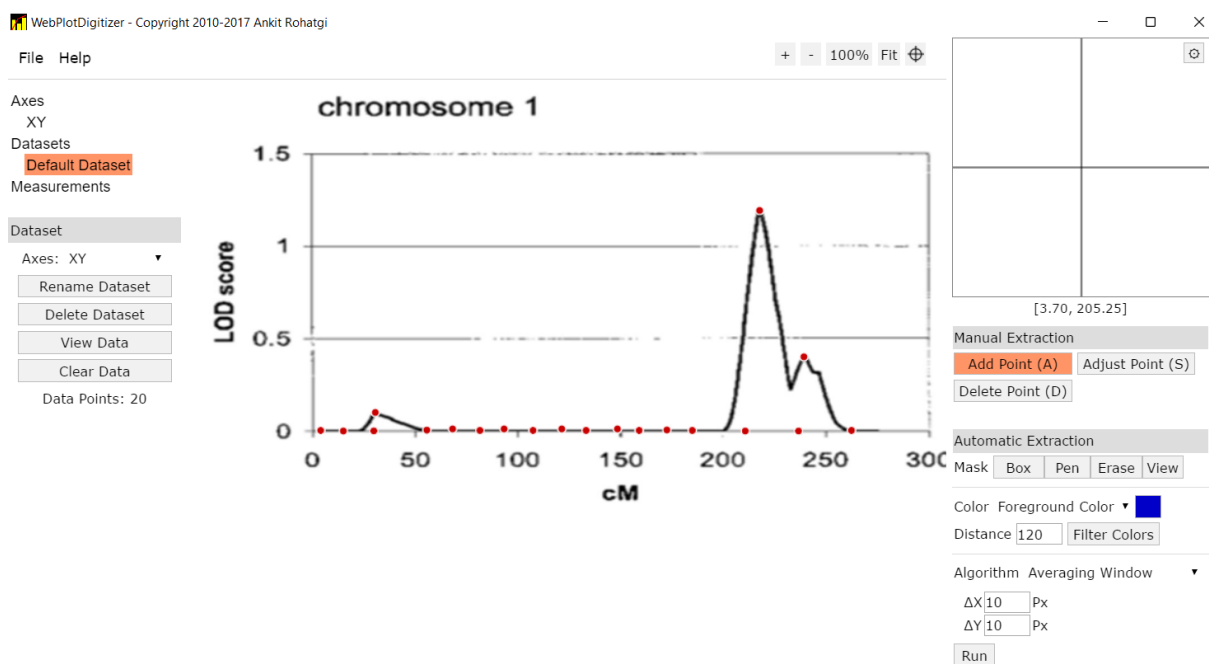
## 2.3 Εξαγωγή δεδομένων

Από κάθε μελέτη έγινε εξαγωγή των παρακάτω δεδομένων: πρώτος συγγραφέας, έτος δημοσίευσης, καταγωγή πληθυσμού, διαγνωστικά κριτήρια, αριθμός ασθενών, αριθμός δεικτών, μέση απόσταση δεικτών, τύπος ανάλυσης/λογισμικό και δεδομένα της ανάλυσης. Σε περιπτώσεις που δεν υπήρχαν διαθέσιμα δεδομένα έγινε προσπάθεια ανάκτησης τους από τους υπεύθυνους επικοινωνίας των δημοσιευμένων μελετών και επι αποτυχίας εξαγωγή αυτών από τα διαθέσιμα διαγράμματα με την χρήση του προγράμματος Web Plot Digitizer (41, 42).

Για την εξαγωγή των δεδομένων σύνδεσης από τα διαγράμματα το πλήρες γονιδίωμα χωρίστηκε σε 120bins ίσου περίπου μεγέθους των 30cM (Πίνακας 6. Πλήθος και μέγεθος bins ανα χρωμόσωμα (Chr)

Σε κάθε LOD score plot από κάθε χρωμόσωμα έγινε αρχικά η βαθμονόμηση του διαγράμματος ορίζοντας το σημείο έναρξης του χρωμοσώματος ( $x=0$ ) και η βασική γραμμή των LOD scores ( $y=0$ ) καθώς και από ένα δεύτερο γνωστό σημείο σε κάθε άξονα. Στη συνέχεια τέθηκε ως βηματισμός στον άξονα  $x$  το μέγεθος του bin από την τιμή μηδέν (έναρξη χρωμοσώματος) έως την τιμή που δηλώνει το μέγεθος του χρωμοσώματος ( $x \max$ ). Αντίστοιχα στον άξονα  $y$  τέθηκαν ως  $y \min$  και  $y \max$  τιμές κοντά στο μηδέν. Με την εντολή

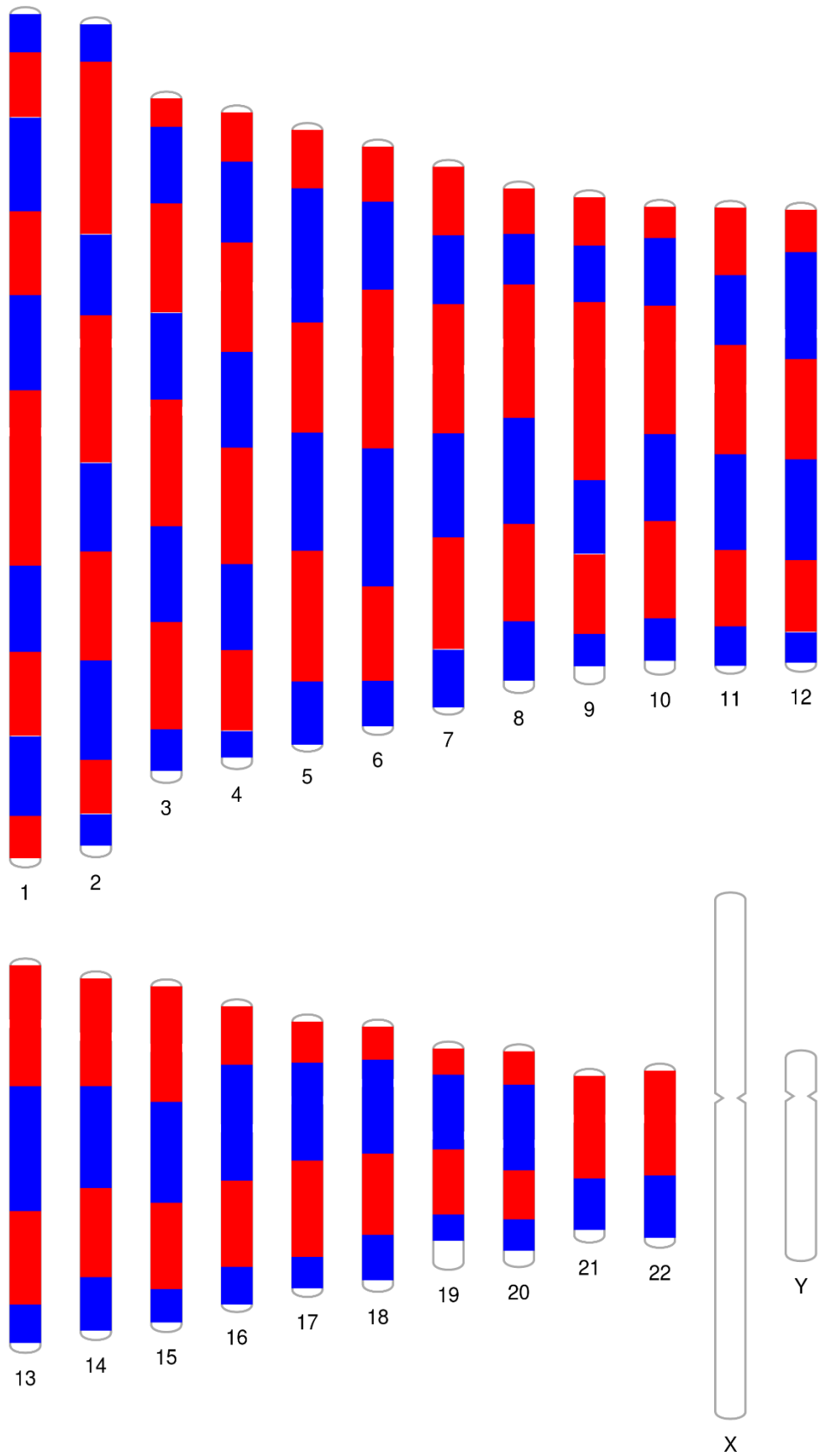
X step with interpolation το λογισμικό μαρκάρει αυτόματα τα σημεία διαχωρισμού των bins σε συνέχεια, πάνω στον άξονα των x από την αρχή έως το τέλος του χρωμοσώματος . Στη συνέχεια με την χρήση του κέρσορα επιλέχθηκε το υψηλότερο σημείο της καμπύλης εντός του κάθε bin που αντιστοιχεί στο υψηλότερο LOD score. Αρνητικές τιμές έλαβαν την τιμή μηδέν (Εικόνα 7). Με την επιλογή View Data εμφανίζονται τα επιλεγμένα σημεία από τα οποία αγνοώντας τα σημεία που ορίζουν την έναρξη των bins έγινε εξαγωγή των μέγιστων LOD scores που αντιστοιχούν σε κάθε bin με την σειρά σε κάθε διαθέσιμο χρωμόσωμα από κάθε μελέτη.



Εικόνα 7. Στιγμιότυπο από το λογισμικό Web Plot Digitizer. Διακρίνονται τα σημεία διαχωρισμού των Bins καθώς και τα υψηλότερα σημεία της καμπύλης που αντιστοιχούν στο μεγαλύτερο LOD score σε κάθε bin (42).

Chr	Μήκος Chr	bins/Chr	Μήκος Bin
1	289,66	10	28,97
2	269,07	10	26,91
3	228,14	8	28,52
4	211,65	8	26,46
5	197,54	6	32,92
6	193,14	6	32,19
7	181,97	6	30,33
8	167,9	6	27,98
9	168,98	6	28,16
10	173,13	6	28,86
11	147,77	6	24,63
12	170,66	6	28,44
13	114,98	4	28,75
14	138,18	4	34,55
15	122,14	4	30,54
16	134,12	4	33,53
17	126,46	4	31,62
18	126	4	31,50
19	105,02	4	26,26
20	101,22	4	25,31
21	57,77	2	28,89
22	62,31	2	31,16

Πίνακας 6. Πλήθος και μέγεθος bins ανα χρωμόσωμα (Chr)



Εικόνα 8. Phenogram των χρωμοσωμάτων και των bins σε αυτά. (43)

Η εξαγωγή των δεδομένων έγινε σε λογισμικό φύλλο όπου τα bins ονομάστηκαν ως c.n όπου c το χρωμόσωμα (Chr) και n το bin κατά σειρά σε κάθε χρωμόσωμα (παράδειγμα το bin 1.1 σηματοδοτεί το πρώτο bin του Chr 1) και σε κάθε bin αντιστοιχίστηκε το max LOD score από κάθε μελέτη.

## 2.4 Ανάλυση δεδομένων

Η μετα ανάλυση των δεδομένων έγινε με βάση την μεθοδολογία GSMA όπως περιγράφηκε στο γενικό μέρος και με την χρήση του λογισμικού HEGESMA (20, 21).

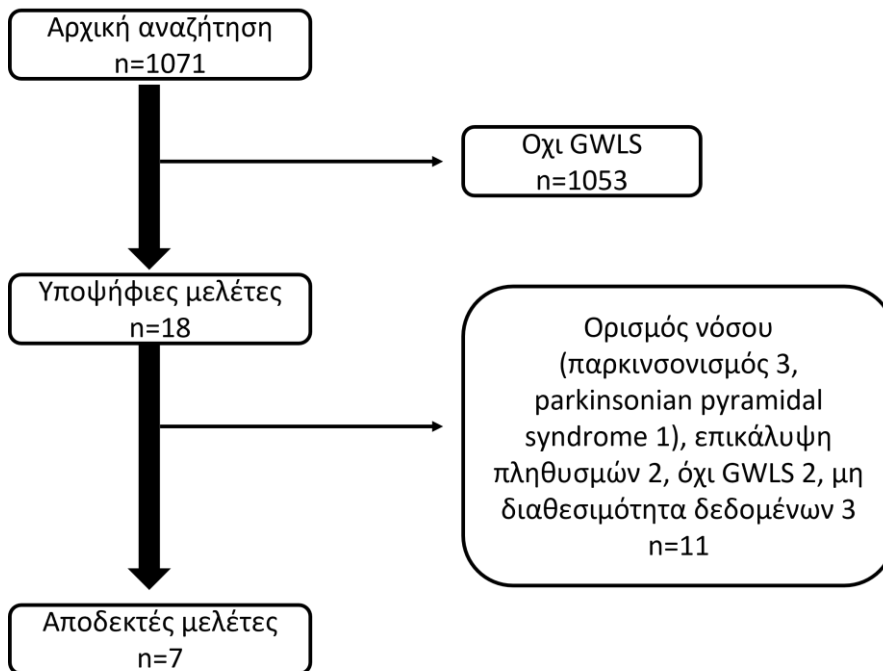
Τα μέγιστα LOD scores του κάθε bin από κάθε χρωμόσωμα και μελέτη όπως προέκυψαν από την εξαγωγή των δεδομένων τέθηκαν σε σειρά από το μικρότερο στο μεγαλύτερο και τους ανατέθηκαν βαθμοί (ranks) έτσι ώστε το bin με το μεγαλύτερο LOD score να έχει την μεγαλύτερη βαθμολογία. Στις περιπτώσεις έλλειψης δεδομένων στα αντίστοιχα bins ανατέθηκε LOD score ίσο με μηδέν με το σκεπτικό ότι εφόσον δεν δημοσιεύονται δεν περιέχουν αξιοποιήσιμες πληροφορίες. Στις περιπτώσεις κοινών τιμών ανατέθηκαν ίδιες βαθμολογίες με την μέθοδο του μέσου όρου. Παράδειγμα, εάν 100 bins από μια μελέτη έχουν LOD score μηδέν τότε και τα 100 bins θα λάβουν rank 50.

Έγινε ανάλυση με στάθμιση και χωρίς στάθμιση. Για την ανάλυση με στάθμιση χρησιμοποιήθηκε ως παράγοντας στάθμισης η ποσότητα που δίνεται από τον λόγο της τετραγωνικής ρίζας των προσβεβλημένων ασθενών προς τον μέσο όρο των προσβεβλημένων.

Οι ranks των bins από κάθε μελέτη και ο παράγοντας στάθμισης της κάθε μελέτης αποτελούν τα δεδομένα εισαγωγής στο λογισμικό HEGESMA όπου έγινε επιλογή ανάλυσης για 7 μελέτες 120 bins και 50000 runs σε σταθμισμένη και μη σταθμισμένη ανάλυση.

## 3. ΑΠΟΤΕΛΕΣΜΑΤΑ

Η συστηματική βιβλιογραφική αναζήτηση επέστρεψε 1071 άρθρα από τα οποία επιλέχθηκαν 18 μελέτες ως υποψήφιες για εισαγωγή. Μετά την εφαρμογή των κριτηρίων ένταξης και αποκλεισμού 7 μελέτες κρίθηκαν ως κατάλληλες για εισαγωγή (Εικόνα 9).



Εικόνα 9. Διάγραμμα ροής βιβλιογραφικής αναζήτησης

Τα στοιχεία των προς εξέταση μελετών φαίνονται στον Πίνακας 7.

Author	DeStefano (44)	Hicks (45)	Pankratz (46)	Martinez (47)	Gao (48)	Cummings (49)	Bruggemann (50)
Year	2001	2002	2003	2004	2009	2011	2012
Source population	unknown	Iceland	North America (primarily White)		European Non- Hispanic descent		Northern German
Diagnostic criteria	UK Brain Bank Criteria	2 cardinal signs no other causes	UPDRS + Diagnostic Checklist		2 cardinal signs	Gelb 1999	UK BBC (except positive family history)
No of affected	113	117	399	471	607	33	5
No of markers	339	781	400	391	5822	5944	423
Διαθέσιμα Chr	1, 9, 10, 16	1-22	1-22	1-22	1-22	1, 2, 4, 6, 7, 10, 12, 15, 16, 19, 20, 21, 22	1-22

Πίνακας 7. Στοιχεία προς ανάλυση μελετών



Τα μέγιστα στατιστικά σύνδεσης όπως προέκυψαν από την διαδικασία εξαγωγής από τα διαθέσιμα διαγράμματα φαίνονται στον Πίνακα 8.

CHR	BIN	DeStefano	Hicks	Pankratz	Martinez	Gao	Cummings	Brugeman
1	1	0	0,55	0	0	0	0,897	0
1	2	0,1	1,16	0	0,026	0,07	0,541	0
1	3	0	3,75	0,044	0,048	0,4	0,026	0
1	4	0	3,71	0	0,048	0,23	0	0,946
1	5	0	0,23	0,04	0,494	0,98	1,015	1,92
1	6	0	1,01	0	0	0,27	0,552	1,326
1	7	0	0,31	0,11	0,177	0,22	0,026	0
1	8	1,19	0,24	0,122	0	0,12	0,995	0,644
1	9	0,398	0,44	0,706	0,235	1,45	0,119	0,498
1	10	0,318	1,13	0	0,102	1,82	0,036	0
2	1		0,45	0	0,167	0	1,562	0
2	2		0,03	0	0	0,08	1,371	0
2	3		0	0	0	0,14	0,392	0
2	4		0	0	0,873	0,19	0,531	0
2	5		0,14	0	2,068	0,12	0,655	0
2	6		0,01	0,132	1,16	0,2	0,325	0
2	7		0,02	0,293	0	0,54	0	0
2	8		0,03	0,239	0,338	0,37	0,015	0
2	9		0,19	3,179	0,416	1,22	0	0
2	10		0,07	0	0,435	1,05	0,222	0
3	1		0	0	0,495	0	0,84	0
3	2		0	0	0,136	0,08		0
3	3		0	0,236	0	0,05		1,692
3	4		0	0	0,221	0		0
3	5		0	0,172	0	0		0
3	6		0	0,087	0,214	1,27		0
3	7		0	0,236	0,97	0,44	0,76	0
3	8		0,19	0	0,19	0,11	0,04	0,562
4	1		0,05	0,4	0	0,61	0,42	1,682
4	2		0,74	0	0	0,29	1,164	0
4	3		0,2	0,15	0,505	0	0	0
4	4		0,36	0,295	0,324	0,19	0,13	0

4	5		0,15	0,263	0,391	1,06	0,225	0
4	6		0,03	0	0	0,02	0,251	0
4	7		0,25	0,317	0,514	0,01	0	0,52
4	8		0,22	0,568	0,1	0,46	0,084	0
5	1		0,17	0,267	0,41	0,07		1,115
5	2		0,07	0,097	0,824	0,31		0
5	3		0,04	1,034	0,315	0		0
5	4		1,39	0,766	1,065	0,31	0,97	0
5	5		1,61	1,326	0,358	0,65		0,43
5	6		0,29	0	0	0	0,03	0
6	1		0	0	0,386	0	0,089	0,359
6	2		0	0	0,303	0,3	0,06	0
6	3		0	0,202	1,453	0	0,05	0
6	4		0	0,054	0,007	0	0,063	0,381
6	5		0	0,516	0,21	0	2,424	0
6	6		0	0	0,68	0	0,851	0
7	1		0	0	1,535	0,04	0,048	0,189
7	2		0	0,246	0,372	0	0,705	1,77
7	3		0	0,41	0	0,12	1,251	0,448
7	4		0	0,015	0,519	1,72	1,029	0
7	5		0,37	0	0,529	0,45	0,966	0
7	6		1,2	0,269	0,558	0,1	0,574	0
8	1		0,31	0,358	0	0,49		0
8	2		0,62	1,059	0	0,83		0
8	3		0	0,146	0,153	1,17		0,448
8	4		0,14	0	0,052	0,08		0
8	5		0	0	0	0,06		0
8	6		0	0,105	0,211	1,34	0,66	0
9	1	0,353	0	0,307	0,395	0,13		0,723
9	2	0,079	0	0,464	0,462	0,7		0
9	3	0,05	0	0	0,723	0,57		0
9	4	0,73	0,1	0,374	0,53	0,32		0
9	5	1,279	0	0	0,309	0,06		0
9	6	0,379	0	0,236	0,2	0,29		0
10	1	0,766	0	0,521	0,593	0,29	0	0,067
10	2	0	0	0,114	0,739	0,94	0,072	0

10	3	0	0,1	0,427	0,131	0,13	0,205	0,414
10	4	1,062	0	0,046	0,047	0,76	0,359	0,499
10	5	0	0	0,185	0	0,38	1,138	0
10	6	0	0	0,181	0	1,16	1,108	0
11	1		0	0	0,093	0		0,469
11	2		0	0,02	0,075	0		0,375
11	3		0	0,019	0,9	0,03		0,308
11	4		0	0,167	1,604	0,07		0
11	5		0	1,254	0,692	0,5		0
11	6		0	0	0,589	0,05		0
12	1		0	0	0,062	0	0	0,156
12	2		0	1,345	0	0	0,169	0,267
12	3		0	0	0	0,41	1,188	0
12	4		0,08	0	0	0	0,774	0
12	5		0	0	0,655	0,03	0	0
12	6		0,39	0	0,748	0,22	0,513	0
13	1		0	0,16	0	0,2		0
13	2		0	0,28	0,663	0,44		0
13	3		0,88	0,978	0,663	0	0,18	0
13	4		0,88	2,013	0,09	0		0
14	1		0	1,178	0,135	0,21	0,34	0
14	2		0,72	2,121	0,076	0,01		0
14	3		0,84	0	0,087	0,05		0,022
14	4		0,71	0,4	0,155	0,07		0
15	1		0	0	0	0,21	0,453	0
15	2		0,32	0,215	0,351	0,54	0,117	0
15	3		0,16	0,296	0,227	0,2	0,859	0
15	4		0	0,31	0	0,11	1,952	0,433
16	1	0,332	0	0,032	0	0,02	0,694	0,159
16	2	0,694	0	0	0,212	0,05	0,49	0,58
16	3	0,043	0,33	0	0,045	0,06	1,122	1,127
16	4	0,921	0,53	0,218	0,43	0,29	1,214	0
17	1		0,29	0	0,847	0,36		0,659
17	2		0,39	0,437	0,938	0,63		0
17	3		0,24	0	0	0,79		0
17	4		0	0	0,345	0,04		0

18	1	0	0,039	0,079	0,18	0,34	0
18	2	0,07	0	0	1,1		1,3
18	3	0,31	0,289	0	0,38		0
18	4	0	0,348	0	0	0,14	0
19	1	0,08	0	0	0,87	0,277	0
19	2	0	0,024	0	0	0,206	0
19	3	0,19	0	1,085	0	1,845	0
19	4	0	0,649	0	0	2,4	0,38
20	1	0	0	0	1,22	1,673	0,4
20	2	0	0,416	0	0,23	1,407	0
20	3	0	0,289	0	0,04	0,974	0
20	4	0	0	0	0,06	0	0
21	1	0,23	0	0	0,49	2,21	0,05
21	2	0,22	0,55	0	1,03	2,448	0
22	1	0	0	0,33	0,02	0,376	0
22	2	0	0,236	0	0,14	2,019	0,456

Πίνακας 8. Μέγιστο στατιστικό σύνδεσης σε κάθε bin απο κάθε μελέτη.

Αντίστοιχα στον Πίνακα 9 φαίνονται οι ranks που ανατέθηκαν σε κάθε bin.

bin	DeStefano	Hicks	Pankratz	Martinez	Gao	Cummings	Brugeman
1.1	52	105	24	19,5	12,5	96	41,5
1.2	108	115	24	41	46	84	41,5
1.3	52	120	56	45	86	50	41,5
1.4	52	119	24	44	72	23,5	112
1.5	52	86	55	90	107	101	120
1.6	52	113	24	19,5	73	85	116
1.7	52	93	62	62	69	49	41,5
1.8	119	88	64	19,5	54	100	109
1.9	113	102	109	71	118	62	104
1.10	109	114	24	55	120	52	41,5
2.1	52	103	24	61	12,5	112	41,5
2.2	52	65	24	19,5	49	110	41,5
2.3	52	30	24	19,5	60	78	41,5
2.4	52	30	24	110	62	83	41,5

2.5	52	76	24	120	56	87	41,5
2.6	52	61	65	116	65	73	41,5
2.7	52	62	88	19,5	96	23,5	41,5
2.8	52	63	80	77	83	48	41,5
2.9	52	81	120	86	115	23,5	41,5
2.10	52	70	24	88	109	69	41,5
3.1	52	30	24	91	12,5	93	41,5
3.2	52	30	24	58	48	23,5	41,5
3.3	52	30	79	19,5	37	23,5	118
3.4	52	30	24	69	12,5	23,5	41,5
3.5	52	30	70	19,5	12,5	23,5	41,5
3.6	52	30	59	68	116	23,5	41,5
3.7	52	30	78	113	89	91	41,5
3.8	52	82	24	63	52	53	107
4.1	52	67	98	19,5	98	79	117
4.2	52	109	24	19,5	76	106	41,5
4.3	52	83	67	92	12,5	23,5	41,5
4.4	52	98	89	75	63	63	41,5
4.5	52	77	82	83	110	70	41,5
4.6	52	64	24	19,5	30	71	41,5
4.7	52	90	93	93	26	23,5	106
4.8	52	85	107	54	91	59	41,5
5.1	52	79	83	85	44	23,5	113
5.2	52	69	60	108	79	23,5	41,5
5.3	52	66	112	74	12,5	23,5	41,5
5.4	52	117	110	114	80	98	41,5
5.5	52	118	116	80	100	23,5	98
5.6	52	92	24	19,5	12,5	51	41,5
6.1	52	30	24	82	12,5	60	92
6.2	52	30	24	72	78	56	41,5
6.3	52	30	73	117	12,5	55	41,5
6.4	52	30	58	40	12,5	57	95
6.5	52	30	104	65	12,5	119	41,5
6.6	52	30	24	103	12,5	94	41,5
7.1	52	30	24	118	35	54	89
7.2	52	30	81	81	12,5	90	119

7.3	52	30	99	19,5	55	109	101
7.4	52	30	49	94	119	102	41,5
7.5	52	99	24	95	90	97	41,5
7.6	52	116	84	97	51	86	41,5
8.1	52	95	95	19,5	93	23,5	41,5
8.2	52	106	113	19,5	104	23,5	41,5
8.3	52	30	66	59	113	23,5	100
8.4	52	75	24	46	50	23,5	41,5
8.5	52	30	24	19,5	42	23,5	41,5
8.6	52	30	61	66	117	88	41,5
9.1	111	30	91	84	58	23,5	111
9.2	107	30	103	89	101	23,5	41,5
9.3	106	30	24	105	97	23,5	41,5
9.4	115	74	96	96	81	23,5	41,5
9.5	120	30	24	73	43	23,5	41,5
9.6	112	30	77	64	77	23,5	41,5
10.1	116	30	105	99	74	23,5	86
10.2	52	30	63	106	106	58	41,5
10.3	52	73	101	56	57	67	97
10.4	118	30	57	43	102	76	105
10.5	52	30	72	19,5	84	105	41,5
10.6	52	30	71	19,5	112	103	41,5
11.1	52	30	24	53	12,5	23,5	103
11.2	52	30	51	48	12,5	23,5	93
11.3	52	30	50	111	32	23,5	91
11.4	52	30	69	119	45	23,5	41,5
11.5	52	30	115	104	94	23,5	41,5
11.6	52	30	24	98	39	23,5	41,5
12.1	52	30	24	47	12,5	23,5	87
12.2	52	30	117	19,5	12,5	65	90
12.3	52	30	24	19,5	87	107	41,5
12.4	52	71	24	19,5	12,5	92	41,5
12.5	52	30	24	100	31	23,5	41,5
12.6	52	100	24	107	70	82	41,5
13.1	52	30	68	19,5	64	23,5	41,5
13.2	52	30	85	102	88	23,5	41,5

13.3	52	112	111	101	12,5	66	41,5
13.4	52	111	118	52	12,5	23,5	41,5
14.1	52	30	114	57	67	75	41,5
14.2	52	108	119	49	27	23,5	41,5
14.3	52	110	24	51	36	23,5	84
14.4	52	107	97	60	47	23,5	41,5
15.1	52	30	24	19,5	68	80	41,5
15.2	52	96	74	79	95	61	41,5
15.3	52	78	90	70	66	95	41,5
15.4	52	30	92	19,5	53	115	99
16.1	110	30	53	19,5	28	89	88
16.2	114	30	24	67	38	81	108
16.3	105	97	24	42	40	104	114
16.4	117	104	75	87	75	108	41,5
17.1	52	91	24	109	82	23,5	110
17.2	52	101	102	112	99	23,5	41,5
17.3	52	89	24	19,5	103	23,5	41,5
17.4	52	30	24	78	33	23,5	41,5
18.1	52	30	54	50	61	74	41,5
18.2	52	68	24	19,5	111	23,5	115
18.3	52	94	87	19,5	85	23,5	41,5
18.4	52	30	94	19,5	12,5	64	41,5
19.1	52	72	24	19,5	105	72	41,5
19.2	52	30	52	19,5	12,5	68	41,5
19.3	52	80	24	115	12,5	114	41,5
19.4	52	30	108	19,5	12,5	118	94
20.1	52	30	24	19,5	114	113	96
20.2	52	30	100	19,5	71	111	41,5
20.3	52	30	86	19,5	34	99	41,5
20.4	52	30	24	19,5	41	23,5	41,5
21.1	52	87	24	19,5	92	117	85
21.2	52	84	106	19,5	108	120	41,5
22.1	52	30	24	76	29	77	41,5
22.2	52	30	76	19,5	59	116	102

Πίνακας 9. Ranking των bins

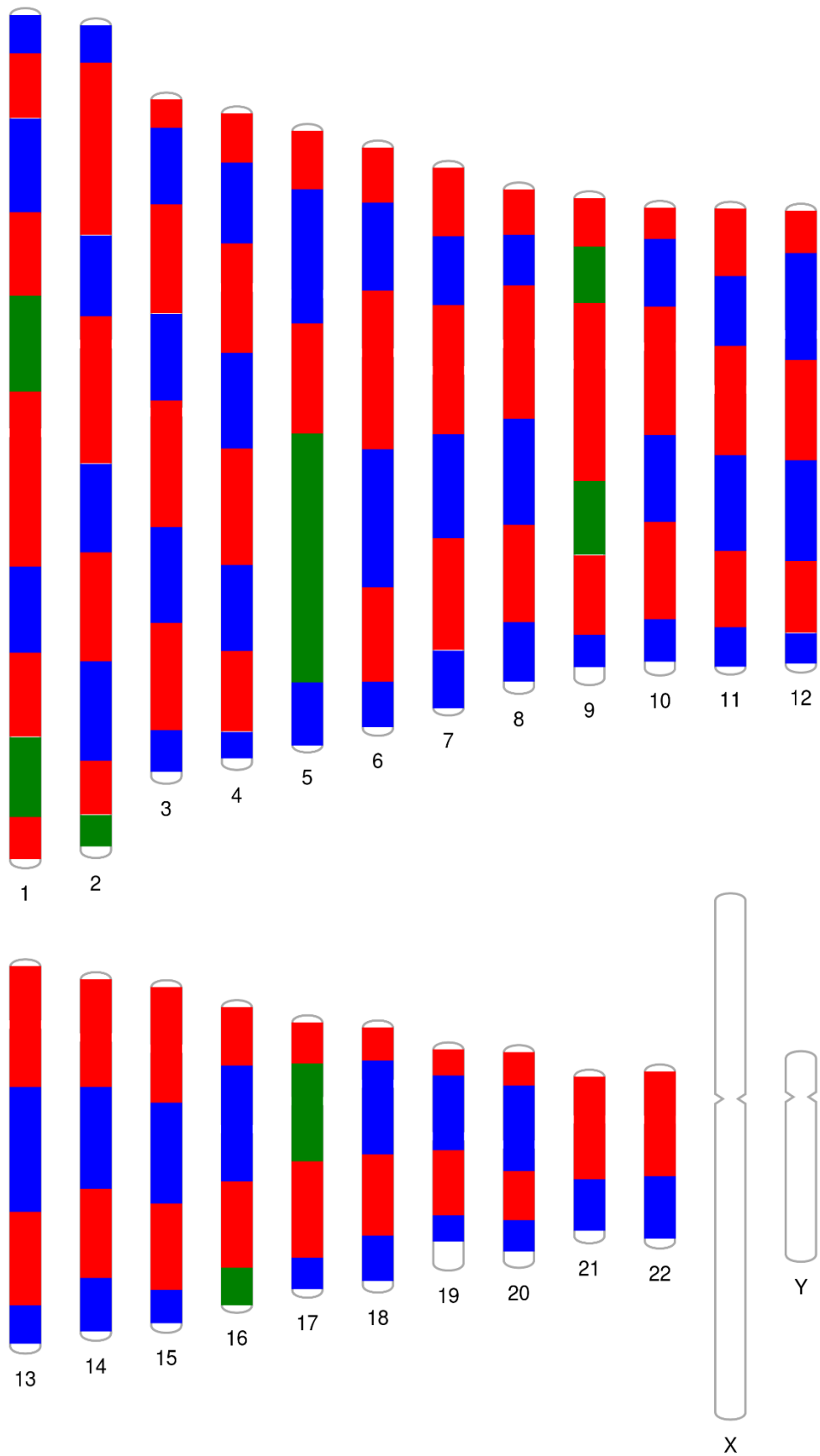
Τα αποτελέσματα της ανάλυσης όπου φαίνονται οι μέσοι όροι των ranks τα στατιστικά  $H_a$ ,  $B$  και  $Q$  καθώς και τα αντίστοιχα  $p$ -values των bins φαίνονται για την μη σταθμισμένη και τη σταθμισμένη ανάλυση στο παράρτημα στην Εικόνα 16 και την Εικόνα 17 αντίστοιχα.

Στον Πίνακα 10 φαίνονται τα σημαντικότερα ευρήματα της μετα ανάλυσης και οπτικά στην Εικόνα 10.

Bin	1.9	5.4	1.5	16.4	5.5	17.2	2.9	9.4	9.2
Mean Rank	97,0	87,5	87,3	86,8	83,9	75,9	74,1	75,3	70,7
DeStefano	113	52	52	117	52	52	52	115	107
Hicks	102	117	86	104	118	101	81	74	30
Pankratz	109	110	55	75	116	102	120	96	103
Martinez	71	114	90	87	80	112	86	96	89
Gao	118	80	107	75	100	99	115	81	101
Cummings	62	98	101	108	23,5	23,5	23,5	23,5	23,5
Brugeman	104	41,5	120	41,5	98	41,5	41,5	41,5	41,5
$P_u$	0,001	0,0128	0,0135	0,0141	0,0269	0,099	0,127	0,109	0,1939
$P_w$	0,002	0,0066	0,0571	0,0341	0,0167	0,0148	0,016	0,0402	0,0458
$p$ value for $Q_u$	0,101	0,4489	0,2129	0,2161	0,7169				
$p$ value for $Q_w$						0,5284	0,717	0,2303	0,6194

Πίνακας 10. Σημαντικότερα αποτελέσματα ανάλυσης GWLS με HEGESMA.  $p_u$ :  $p$ -value unweighted,  $p_w$ :  $p$ -value weighted,  $Q_u$ :  $Q$  unweighted,  $Q_w$ :  $Q$  weighted.





Εικόνα 10. Phenogram των σημαντικότερων bins του Πίνακα 7 (πράσινο)

# ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ: ΜΕΤΑ ΑΝΑΛΥΣΗ ΟΛΟΓΕΝΩΜΙΚΩΝ ΜΕΛΕΤΩΝ ΣΥΧΕΤΙΣΗΣ (GWAS)

## 1. ΕΙΣΑΓΩΓΗ

Στο δεύτερο κεφάλαιο του ειδικού μέρους πραγματοποιήθηκε μετα ανάλυση των ολογενωμικών μελετών συσχέτισης (GWAS) στη νόσο Parkinson με την αξιοποίηση του λογισμικού METanalysis of RAnked DIScovery datasets (METRADISC-XL). Η μετα ανάλυση GWAS είναι μια πολύ εξειδικευμένη και δύσκολη εργασία η οποία απαιτεί την χρήση ειδικού λογισμικού, την συνεργασία μεγάλων ερευνητικών ομάδων με διεθνείς συνεργασίες και διαθεσιμότητα των δεδομένων. Επιπλέον εφόσον κάθε GWAS μπορεί να χρησιμοποιεί διαφορετικά σετ γονοτύπων η σύνθεση τους μπορεί να απαιτήσει την προσομοίωση γονοτύπων με αμφίβολα αποτελέσματα στην αξιοπιστία των αποτελεσμάτων.

Στόχος της παρούσας ήταν να ξεπεραστούν τα ανωτέρω προβλήματα με μια μεθοδολογία πιο βατή με την δυνατότητα να συνθέτει δεδομένα από διαφορετικές πηγές και διαφορετικές μεθοδολογίες και να παράγει δεδομένα συγκρίσιμα με αντίστοιχη ανάλυση άλλου είδους μελετών όπως οι GWLS στο πρώτο κεφάλαιο.

## 2. ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΣ

### 2.1 Βιβλιογραφική αναζήτηση

Η βιβλιογραφική αναζήτηση διενεργήθηκε στην βάση δεδομένων PubMed με στόχο τις μελέτες GWAS που αφορούσαν την νόσο του Parkinson με την χρήση λέξεων κλειδιών και συνδυασμού αυτών όπως: «Parkinson's disease», «genome-wide association study», «GWAS», «linkage disequilibrium», «whole genome association». Για να ενισχυθεί το βάθος και η αξιοπιστία της αναζήτησης τα ευρήματα διασταυρώθηκαν με τα ευρήματα αναζήτησης στις βάσεις HuGe navigator/GWAS integrator (51) και GWAS catalog (52) όπου χρησιμοποιήθηκε ως κριτήριο ο όρος του φαινοτύπου «Parkinson disease».

## 2.2 Κριτήρια ένταξης

Ως κατάλληλες για ένταξη ορίσθηκαν οι μελέτες που ήταν στην Αγγλική γλώσσα, ακολούθησαν τυπική μεθοδολογική προσέγγιση GWAS με σαφή ορισμό του υπο εξέταση φαινοτύπου και διαθεσιμότητα γονοτυπικών δεδομένων πλήρων ή τουλάχιστον σε επίπεδο των πιο σημαντικών. Μελέτες οι οποίες εξέτασαν άλλες μορφές παρκινσονισμού (νεανικός ή πρώιμης έναρξης παρκινσονισμός) ή περιέγραφαν συσχετίσεις με κλινικά χαρακτηριστικά της νόσου (ηλικία έναρξης, κινητικές γνωσιακές δεξιότητες) ή συσχετίσεις (γονιδίου-περιβάλλοντος, κατανάλωση καφέ) αποκλείσθηκαν από την ανάλυση. Στις περιπτώσεις επικάλυψης μελετών επιλέχθηκε αυτή με το μεγαλύτερο δείγμα.

## 2.3 Εξαγωγή δεδομένων

Από κάθε κατάλληλη για εισαγωγή μελέτη εξήχθησαν τα εξής στοιχεία: λεπτομέρειες έκδοσης (πρώτος συγγραφέας, έτος δημοσίευσης, τίτλος), το πλήθος των ασθενών και μαρτύρων που γονοτυπώθηκαν, όλα τα διαθέσιμα SNPs από το δημοσιευμένο άρθρο, τα παραρτήματα του ή τις διαθέσιμες βάσεις δεδομένων με τα αντίστοιχα p-values και την θέση τους. Στην ανάλυση χρησιμοποιήθηκαν μόνο τα δεδομένα της αρχικής γονοτύπωσης. Δεδομένα αναπαραγωγής και επιβεβαίωσης καθώς και επικαλυπτόμενα datasets απορρίφθηκαν.

## 2.4 Διαχωρισμός των bins

Με τρόπο ίδιο με τις μελέτες γενετικής σύνδεσης όλα τα χρωμοσώματα χωρίσθηκαν σε 120 bins περίπου ίδιου μεγέθους και ονομάσθηκαν επίσης με τον ίδιο τρόπο δηλαδή ως x.γ όπου x το χρωμόσωμα και γ το Bin κατά σειρά από την αρχή του χρωμοσώματος (Πίνακας 6). Επιπλέον η φυσική τοποθεσία του κάθε Bin προσδιορίσθηκε αναζητώντας το ζεύγος βάσεων έναρξης και λήξης του bin από τον πλησιέστερο πολυμορφικό δείκτη ενός χάρτη Marshfield και την ανεύρεση του στον UCSC Genome Browser on Human (GRch38/hg38 Assembly).

## 2.5 Ανεύρεση σημαντικών SNP

Από κάθε μελέτη αναζητήθηκε το SNP με την μεγαλύτερη στατιστική σημαντικότητα με όρους τιμής p-value σε κάθε bin. Λόγω του πολύ μεγάλου αριθμού SNPs για να γίνει ο

εντοπισμός τους δημιουργήθηκε κώδικας στη γλώσσα προγραμματισμού Python με την χρήση Jupiter notebook στον οποίο πραγματοποιήθηκε μια βήμα-βήμα αλγοριθμική προσέγγιση κατά την οποία καταγράφηκαν τα SNPs που αντιστοιχούν σε κάθε bin από κάθε μελέτη και στη συνέχεια διατηρήθηκε αυτό με το πιο σημαντικό στατιστικό (53, 54). Ο κώδικας είναι διαθέσιμος στην ιστοσελίδα [https://github.com/andreoua/basepair\\_to\\_bin](https://github.com/andreoua/basepair_to_bin).

## 2.6 Μετα ανάλυση μελετών

Για κάθε μελέτη τα bins έλαβαν βαθμολογίες κατάταξης (ranks) με τιμές από 1 έως 120 σύμφωνα με τη στατιστική τους σημαντικότητα. Τα bins με το μικρότερο p-value έλαβε τον μεγαλύτερο rank (120). Τα Bins στα οποία δεν αντιστοιχούσε κάποιο SNP θεωρήθηκαν ως missing values και έλαβαν την τιμή -99 ώστε να αναγνωρισθούν από το λογισμικό ως τέτοιες. Στις περιπτώσεις όπου παρατηρήθηκαν ίσες τιμές p-values τα bins έλαβαν rank τον μέσο όρο των ranks τους. Οι ranks των bins αθροίστηκαν και βρέθηκε ο μέσος όρος τους. Ο μέσος rank (R) του κάθε bin είναι ένας δείκτης της συσχέτισης του με τον φαινότυπο της νόσου (νόσος Parkinson). Για να μελετηθεί η συνέπεια των αποτελεσμάτων αξιολογήθηκε και η ετερογένεια τους μεταξύ των μελετών με τον υπολογισμό του στατιστικού Q το οποίο ορίζεται ως το άθροισμα των τετραγώνων των αποκλίσεων από τον μέσο όρο των ranks από κάθε μελέτη όπως περιεγράφηκε στο γενικό μέρος (20, 21).

Για την πραγματοποίηση της ανάλυσης χρησιμοποιήθηκε το λογισμικό METRADISC-XL το οποίο είναι διαθέσιμο στην ιστοσελίδα <http://biomath.med.uth.gr/> και έχει χρησιμοποιηθεί για την ανάλυση δεδομένων από μελέτες μικροσυστοιχιών. Σε αυτήν την περίπτωση ο βιολογικός δείκτης ενδιαφέροντος είναι το χρωμοσωμικό Bin. Όπως αναφέρθηκε κάθε bin έλαβε rank ανάλογα με την σημαντικότητά του. Εφόσον υπάρχουν ελλείψεις δεδομένων διαφορετικός αριθμός bins θα έχει λάβει rank σε κάθε μελέτη. Έτσι κάθε αρχικός rank σταθμίζεται με έναν παράγοντα ο οποίος προκύπτει από το γινόμενο του μέγιστου αριθμού bins με διαθέσιμα p-values (nmax) δια του πλήθους των Bins που ταξινομήθηκαν στη συγκεκριμένη μελέτη.

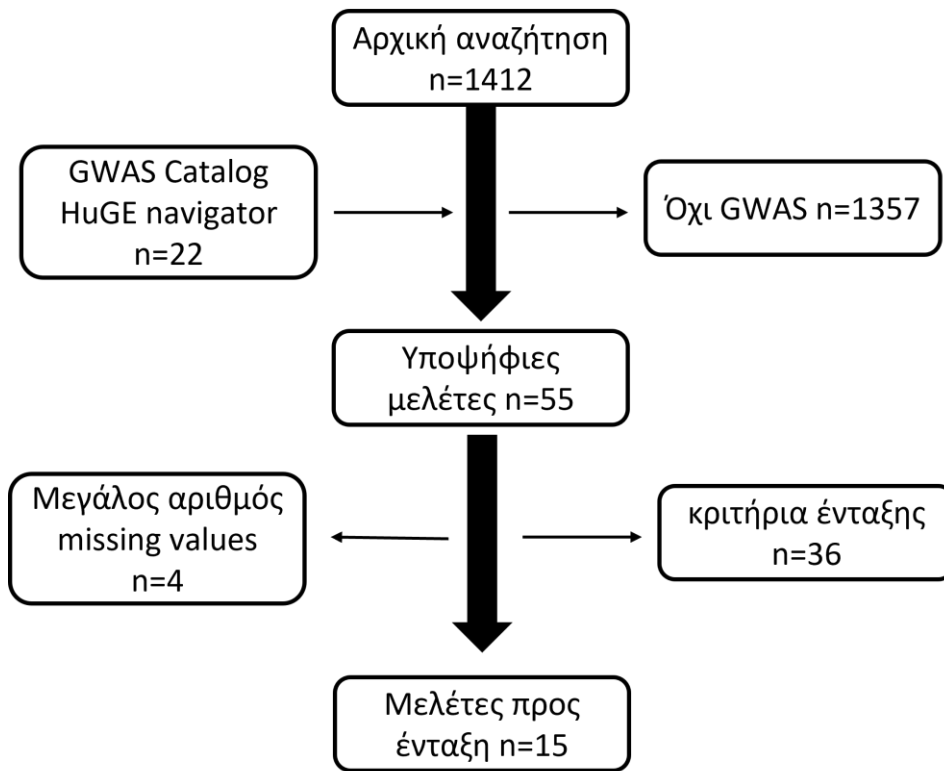
Η σημαντικότητα των στατιστικών R και Q αξιολογείται από το λογισμικό με την χρήση μεθοδολογίας Monte Carlo κατά την οποία οι ranks κάθε μελέτης προσομοιώνονται τυχαία για πολλές φορές (100000 φορές) και το λογισμικό υπολογίζει τα προσομοιωμένα στατιστικά και κατασκευάζει μηδενικές κατανομές αυτών. Από την στιγμή που υπάρχουν ελλείψεις δεδομένων δηλαδή δεν υπάρχει κατάταξη για όλα τα bins σε όλες τις μελέτες, το

κάθε bin συγκρίνεται με την μηδενική κατανομή που αντιστοιχεί στην ομάδα των bins που έχουν διαθέσιμη πληροφορία (rank) από τις ίδιες μελέτες. Αυτές οι ομάδες ονομάζονται τάξεις πληροφορίας (information classes) και ορίζονται ουσιαστικά από τα δεδομένα που λείπουν. Η σημαντικότητα των στατιστικών τελικά ορίζεται ως το ποσοστό των προσομοιωμένων στατιστικών που υπερβαίνει το παρατηρούμενο.

Το METRADISC-XL επιτρέπει την μη σταθμισμένη και τη σταθμισμένη ανάλυση. Πραγματοποιήθηκαν και οι δύο με την χρήση ως παράγοντα στάθμισης της ποσότητας  $(n_{1i} * n_{2i}) / (n_{1i} + n_{2i})$  όπου  $n_{1i}$  είναι ο αριθμός των ασθενών και  $n_{2i}$  το αριθμός των μαρτύρων στην μελέτη  $i$ .

### 3. ΑΠΟΤΕΛΕΣΜΑΤΑ

Η βιβλιογραφική αναζήτηση επέστρεψε 1412 καταγραφές από τις οποίες βρέθηκαν 55 άρθρα ως υποψήφια για ένταξη στη μελέτη. Η βάση δεδομένων GWAS catalog στο λήμμα «Parkinson disease» περιείχε 22 μελέτες οι οποίες ταυτίστηκαν με αυτές που περιέχονταν στην βάση GWAS integrator. Μετά την εφαρμογή των κριτηρίων ένταξης και αποκλεισμού 19 μελέτες επιλέχθηκαν για ανάλυση. Λαμβάνοντας υπόψη τον πολύ μεγάλο αριθμό απώλειας δεδομένων από τις μελέτες των Beecham (55), Davis (56), Satake (57) και Vacic (58) οι μελέτες αυτές επίσης αφαιρέθηκαν από την ανάλυση με αποτέλεσμα στην τελική ανάλυση να περιλαμβάνονται 15 GWASs (Εικόνα 11).



Εικόνα 11. Διάγραμμα ροής βιβλιογραφικής αναζήτησης μελετών GWAS

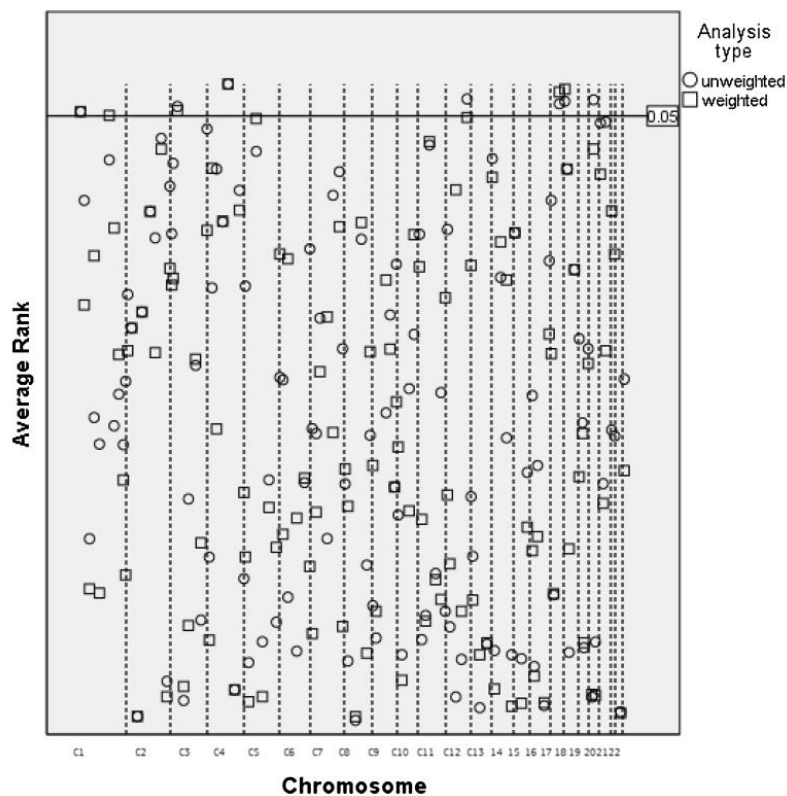
Η εκτέλεση του κώδικα ταυτοποίησης των πιο σημαντικών SNPs στα αντίστοιχα bins αναγνώρισε 543 SNPs και τα αντίστοιχα p-values στα bins που ανήκουν ενώ 1257 bins παρέμειναν κενά (Πίνακας 11). Με βάση την διαθεσιμότητα των δεδομένων τα 543 SNPs ανήκουν σε 92 διαφορετικές τάξεις πληροφορίας. Η εκτέλεση του λογισμικού METRADIS-ΧΙ για 15 μελέτες, 120 bins, 100000 προσομοιώσεις και 92 τάξεις πληροφορίας αποκάλυψε 7 σημαντικά bins για συσχέτιση με τον φαινότυπο της νόσου Parkinson (Πίνακας 12, Εικόνα 12, Εικόνα 13).

No.	Author	Year	Initial sample size (cases/controls)	Ethnicity	Extracted SNPs (n)	Matched bins (n)
1	Bandre-Ciga	2016	240/192	Caucasian	28	21
2	Beecham*	2013	484/1.145	Caucasian	1	1
3	Davis*	2013	31/767	Amish	3	3
4	Do	2011	3.426/29.624	Caucasian	390	65
5	Edwards	2010	604/619	Caucasian	72	33
6	Foo	2016	779/13.227	Asian	96	32
7	Fung	2006	267/270	Caucasian	26	17
8	Hamza	2010	2.000/1.986	Caucasian	89	16
9	Hu Y	2015	250/250	Chinese	22	21
10	Liu	2011	268/178	Ashkenazi	55	32
11	Pickrell	2016	9.619/324.522	Caucasian	25	20
12	Saad	2010	1.039/1.984	Caucasian	50	21
13	Satake*	2009	988/2.521	Japanese	20	4
14	Simon-Sanchez	2009	1.713/3.978	Caucasian	345	87
15	Simon-Sanchez	2011	772/2.024	Caucasian	30	8
16	Spencer	2010	1.705/5.175	Caucasian	55	24
17	Vacic*	2014	1.130/2.611	Ashkenazi	4	4
18	Maraganore	2005	381/363	Caucasian	190059	120
19	Chang	2017	6.476/302.042	Caucasian	27	26
Total			32208/693847		191397 ** (191369)	555 ** (543)

Πίνακας 11. Δημογραφικά χαρακτηριστικά μελετών GWAS. Οι μελέτες με \* αφαιρέθηκαν από την ανάλυση λόγω πολύ μικρού αριθμού SNPs (<5). \*\* Το πλήθος των SNPs και των αντίστοιχων bins που εισήχθησαν στην ανάλυση (39, 59-76).

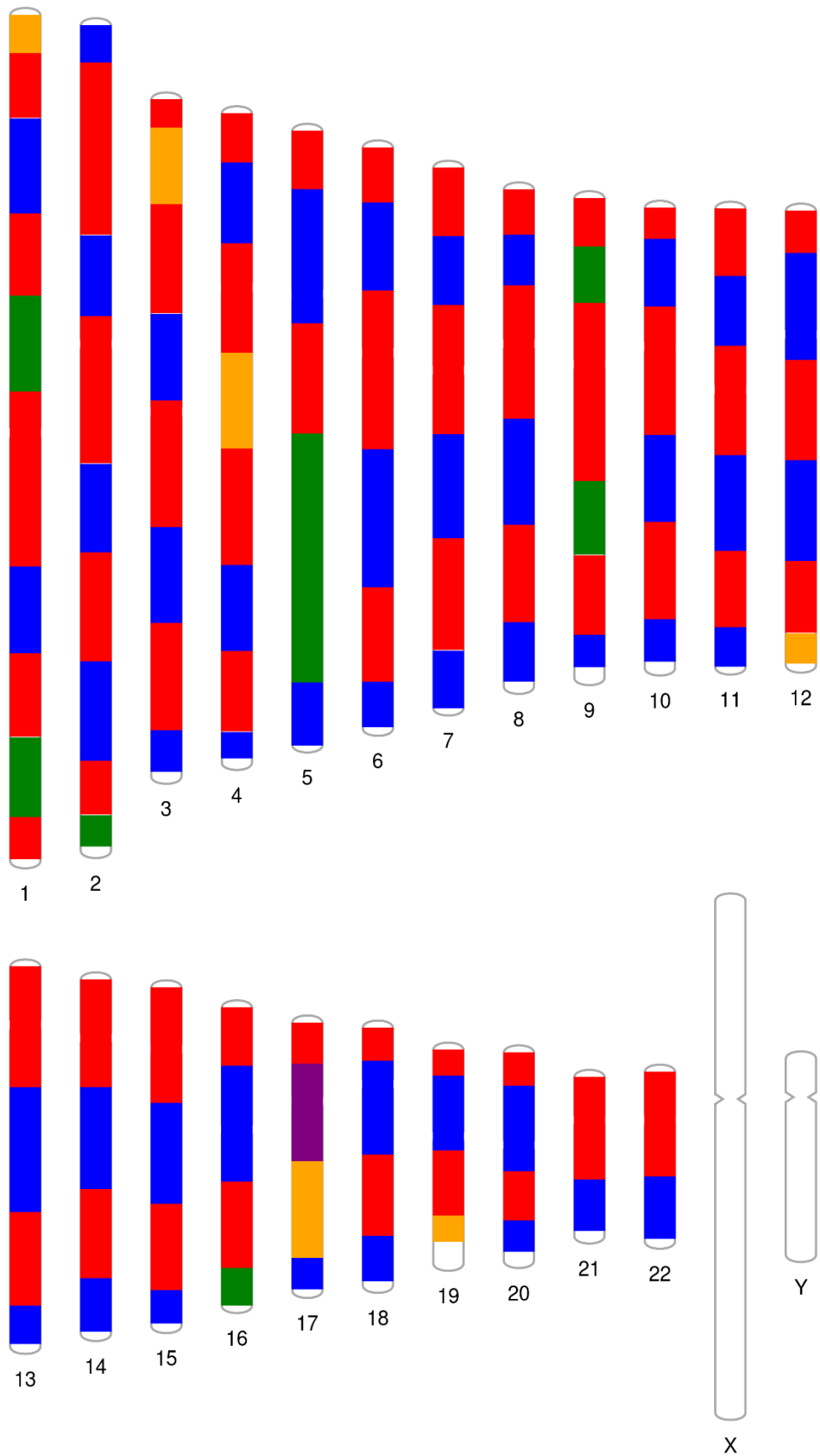
Bin	4.4	12.4	19.4	17.3	17.2	3.2	1.1
study 1	657,1	-99,0	-99,0	680,0	662,9	-99,0	-99,0
study 4	219,7	-99,0	-99,0	216,0	217,8	186,5	-99,0
study 5	436,4	-99,0	-99,0	367,3	-99,0	378,2	-99,0
study 6	450,0	-99,0	-99,0	-99,0	-99,0	435,0	-99,0
study 7	-99,0	-99,0	-99,0	-99,0	840,0	-99,0	-99,0
study 8	892,5	-99,0	-99,0	847,5	870,0	-99,0	-99,0
study 9	628,6	-99,0	-99,0	-99,0	-99,0	-99,0	628,6
study 10	-99,0	-99,0	-99,0	345,0	356,3	431,3	-99,0
study 11	720,0	-99,0	-99,0	708,0	-99,0	-99,0	-99,0
study 12	685,7	680,0	-99,0	-99,0	662,9	-99,0	-99,0
study 14	158,2	-99,0	139,8	143,7	156,9	127,9	-99,0
study 15	1800,0	-99,0	-99,0	-99,0	1755,0	-99,0	-99,0
study 16	590,0	-99,0	-99,0	575,0	585,0	580,0	-99,0
study 18	10,0	108,0	115,0	66,0	73,0	11,0	116,0
study 19	553,8	-99,0	-99,0	549,2	-99,0	-99,0	-99,0
Rmean	45141167	19968572	1496758	23183384	50969116	11496949	18539999
Rmean p	0,00	0,02	0,02	0,03	0,03	0,04	0,04
Rw/mean p	0,00	0,05	0,10	0,01	0,01	0,04	0,04
Q-mean p	0,12	0,14	0,50	0,26	0,50	0,12	0,97

Πίνακας 12. Τα σημαντικότερα bins και οι αντίστοιχοι ranks σε μη σταθμισμένη (Rmean) και σταθμισμένη (Rw/mean) ανάλυση με τα στατιστικά σημαντικότητας και ετερογένειας.



Εικόνα 12. Σταθμισμένη (τετράγωνα) και μη σταθμισμένη (κύκλοι) σημαντικότητα των μέσων ranks από 120 bins σε προσαρμοσμένα με βάση το μέγεθος χρωμοσώματα.





Εικόνα 13. Σημαντικά bins της μετα-ανάλυσης των GWLS (πράσινο), των GWAS (κίτρινο) και κοινά bins (μωβ)

## ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ: ΓΕΝΩΜΙΚΗ ΣΥΓΚΛΕΙΣΗ

Η έννοια της γενωμικής σύγκλισης πρωτοπαρουσιάστηκε από τον Hauser το 2003 σε μία μελέτη σύγκλισης γενετικών δεδομένων από μελέτη γενετικής έκφρασης και γενετικής σύνδεσης στη νόσο του Parkinson (77). Η βασική ιδέα πίσω αυτή την μελέτη είναι το γεγονός ότι ενώ οι διαφορετικές προσεγγίσεις ξεχωριστά (έκφραση και σύνδεση) έχουν δύναμη να αναγνωρίσουν υποψήφιους γενετικούς τόπους που σχετίζονται με σύνθετες νόσους, αυτοί παραμένουν μεγάλοι σε αριθμό. Έτσι η εστίαση και μείωση του αριθμού των υποψήφιων γενετικών τόπων είναι σημαντική. Με βάση το ίδιο σκεπτικό μπορούν να συνδυαστούν με την χρησιμοποίηση βιοπληροφορικών μεθόδων και άλλου είδους προσεγγίσεις όπως οι μελέτες γενετικής συσχέτισης, ολογενωμικές μελέτες γενετικής σύνδεσης και γενετικής συσχέτισης και επιγενετικές μελέτες.

Σε αυτή τη βάση το 2009 από το εργαστήριο Βιομαθηματικών του Π.Θ. έγινε η πρώτη σύγκλιση γενετικών δεδομένων από μελέτες ελεύθερες υπόθεσης (GWLS και GWAS) για πολλές σύνθετες παθήσεις (78). Η σύγκλιση αυτή των δύο μεθόδων εξέτασε το κατά πόσο τα δεδομένα από τις GWAS εκφρασμένα σε στατιστικά σημαντικά SNPs συμπίπτουν με τις περιοχές σύνδεσης που προκύπτουν από τις GWLS. Η σημαντικότητα αυτής της σύγκλισης αξιολογήθηκε σε σχέση με την πιθανότητα να είναι τυχαία, η οποία ισούται με το ποσοστό κάλυψης του γονιδιώματος από περιοχές σύνδεσης.

Προς την ίδια κατεύθυνση έχει αναφερθεί και διαφορετική προσέγγιση. Σε αυτήν συλλέχθηκαν μελέτες GWLS, GWAS και GWE (Genome Wide Expression) και οι περιοχές σύνδεσης, συσχέτισης και έκφρασης έλαβαν βαθμολογία (ranks) ανάλογα με τη σημαντικότητά τους και την παρουσία τους ή όχι και στις τρεις κατηγορίες μελετών (79) υπολογίζοντας ένα συνολικό σκορ με σκοπό την προτεραιοποίηση γενετικών τόπων.

Στην παρούσα μελέτη, λόγω των πολλαπλών ελλείψεων δεδομένων και του τεράστιου όγκου των SNPs που έχουν μελετηθεί σε GAS η εφαρμογή των μεθόδων γενετικής σύγκλισης δεν ήταν δυνατόν να εφαρμοστεί παρά με την οπτικοποίηση των δεδομένων όπως αναφέρεται και στη συζήτηση.

## ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ: ΣΥΖΗΤΗΣΗ

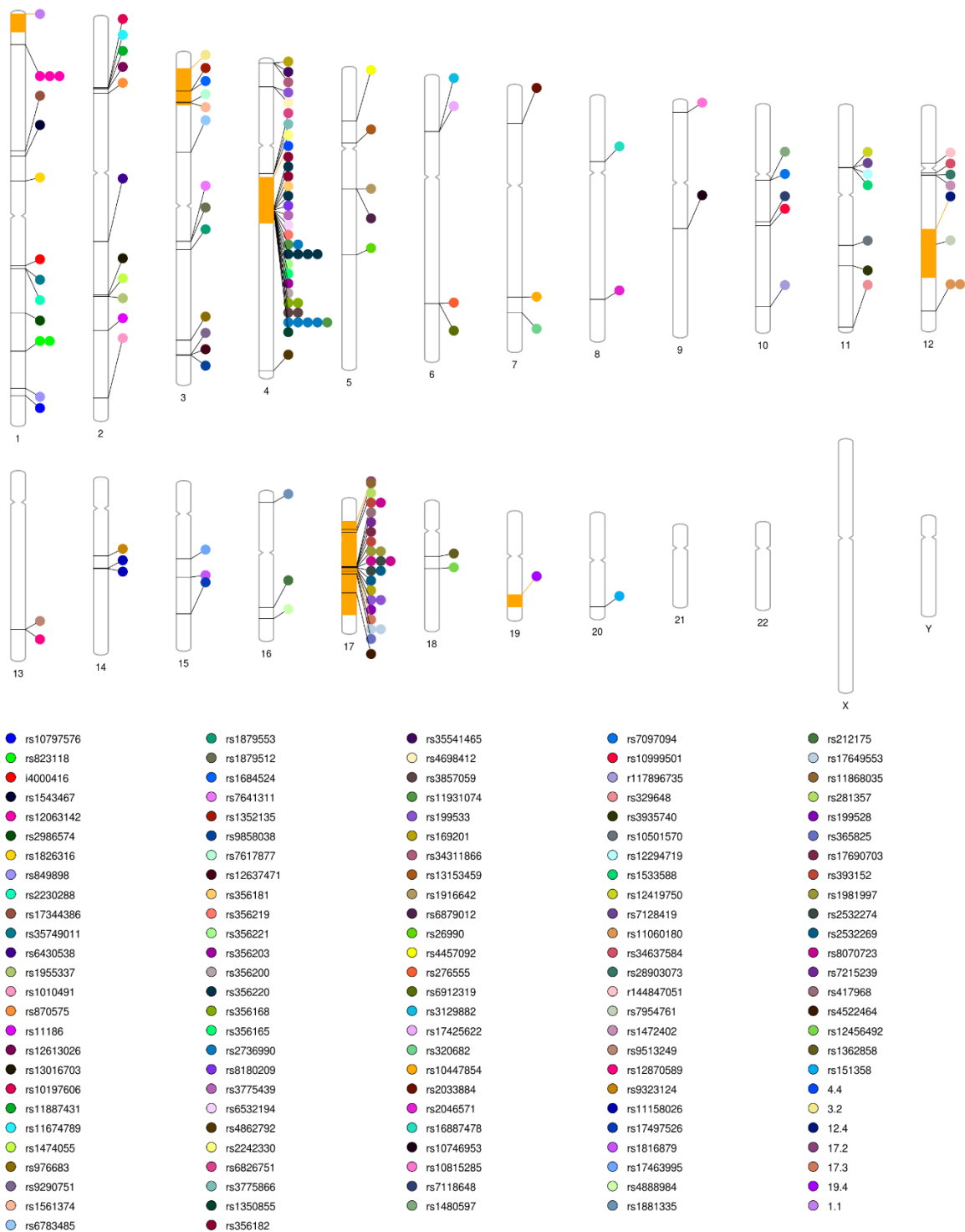
Η αναζήτηση της κληρονομικότητας στη νόσο του Parkinson υπήρξε ένας μακρής και συναρπαστικός δρόμος με πολυάριθμες επιτυχίες αλλά και σημαντικές δυσκολίες. Η ανάπτυξη της τεχνολογίας υπήρξε αρωγός σε αυτή τη προσπάθεια ενώ η εγγενής δυσκολία του θέματος η τροχοπέδη της. Έως τώρα 20 γονίδια έχουν ενοχοποιηθεί για τη νόσο ενώ πάνω από 90 SNPs έχουν αναγνωρισθεί να σχετίζονται με το ρίσκο εμφάνισης της (40). Οι GWLS και στη συνέχεια οι GWAS πρόσθεσαν τεράστιο όφελος σε αυτή την αναζήτηση παρόλα αυτά απαιτώντας σημαντικό κόπο και μέσα. Οι μέθοδοι αυτές έχουν βοηθήσει στην ανακάλυψη ενός σημαντικού μέρους της κληρονομισημότητας της νόσου με ένα ακόμη μεγαλύτερο μέρος να παραμένει άγνωστη. Ο διαθέσιμος τρόπος για να προσεγγισθεί αυτό το πρόβλημα είναι τα όλο και μεγαλύτερα δείγματα ασθενών και μαρτύρων αυξάνοντας ταυτόχρονα και το υπολογιστικό βάρος.

Η παρούσα διατριβή είχε σαν σκοπό την ανεύρεση περιοχών με ενδείξεις συσχέτισης με το ρίσκο εμφάνισης νόσου Parkinson από διαφορετικά είδη μελετών (GWLS και GWAS) με μια νέα προσέγγιση στην μετα ανάλυση τους που να απαιτεί λιγότερα μέσα και να είναι πιο κατανοητή από τους κλινικούς ιατρούς. Αυτές οι μη παραμετρικές προσεγγίσεις έχουν επιπλέον το πλεονέκτημα να μην περιορίζονται από κάποια υπόθεση και να μπορούν να συνδυάσουν δεδομένα από μελέτες με διαφορετικές μεθοδολογίες. Ένας πρόσθετος στόχος ήταν ο συνδυασμός των ευρημάτων από τις μετα αναλύσεις των διαφορετικών μελετών με το σκεπτικό ότι περιοχές με ένδειξη σημαντικότητας και στις δύο προσεγγίσεις θα έχει μεγαλύτερες πιθανότητες να αποτελεί πραγματική περιοχή ενδιαφέροντος για περαιτέρω μελέτη.

Στην μετα ανάλυση των GWLS βρέθηκαν εννέα bins με σημαντικό mean rank (1.5, 1.9, 2.9, 5.4, 5.5, 9.2, 9.4, 16.4, 17.2) σε σταθμισμένη ή/και μη σταθμισμένη ανάλυση. Από αυτά το bin 1.9 (chr1: 212916368-236478654) εμφανίζει το μεγαλύτερο σημαντικό mean rank με ταυτόχρονα χαμηλή ετερογένεια (Εικόνα 10, Πίνακας 10).

Στην μετα ανάλυση των GWAS ο συνδυασμός 534 SNPs από 15 μελέτες είχε ως αποτέλεσμα ένα bin με σημαντικότητα σε γενωμικό επίπεδο και επιπλέον

έξι με μικρότερη συσχέτιση ( $p\text{-value} < 0,05$ ). Από τα 138 πιο σημαντικά αρχικά γονοτυπωμένα SNPs στις μελέτες GWAS τα 43 βρίσκονται εντός ενός από αυτά τα επτά σημαντικά bins.



Εικόνα 14. Τα σημαντικά bins όπως προέκυψαν από την ανάλυση (κίτρινες περιοχές) και τα σημαντικότερα SNPs όπως αναφέρονται στις ενταγμένες GWAS

Το πιο σημαντικό bin είναι το 4.4 (chr4: 70530434-9873681). Τουλάχιστον 22 SNPs από 10 διαφορετικές μελέτες εμφάνισαν στατιστική σημαντικότητα συσχέτισης με τη νόσο και βρίσκονται εντός του bin 4.4. Σε 7 μελέτες, το πιο σημαντικό SNP βρίσκεται στο Bin 4.4 με αποτέλεσμα να λαμβάνει και τον μεγαλύτερο rank (120) με αποτέλεσμα να εμφανίζει και

χαμηλή ετερογένεια. Επιπλέον τα 37 από τα 67 SNPs που αναφέρονται να σχετίζονται με την νόσο Parkinson στην βάση GWAS catalog (πρόσβαση δεδομένων 27 Ιουλίου 2020) βρίσκονται επίσης σε αυτό. Στην περιοχή του Bin 4.4 βρίσκεται το γονίδιο SNCA (Chr4: 89700345-89838315), ένα καλά αναγνωρισμένο γονίδιο με σημαντική πιθανότητα να αποτελεί ένα πραγματικό αιτιολογικό γονίδιο (80). Το bin 3.2 επίσης έδειξε κάποιου βαθμού συνέπεια εμφάνισης μεταξύ των μελετών με τον rank να βρίσκεται στο ανώτερο 25% στις 6 από τις 7 μελέτες για τις οποίες υπήρχαν δεδομένα αλλά με p-value του μέσου rank ίσο με 0.03. Τα bin 17.2 και 17.3 εμφάνισαν σημαντικότητα στο επίπεδο του 0,01 αλλά με σημαντική ετερογένεια. Το 17.3 είναι η περιοχή όπου εδράζεται το γονίδιο MAPT.

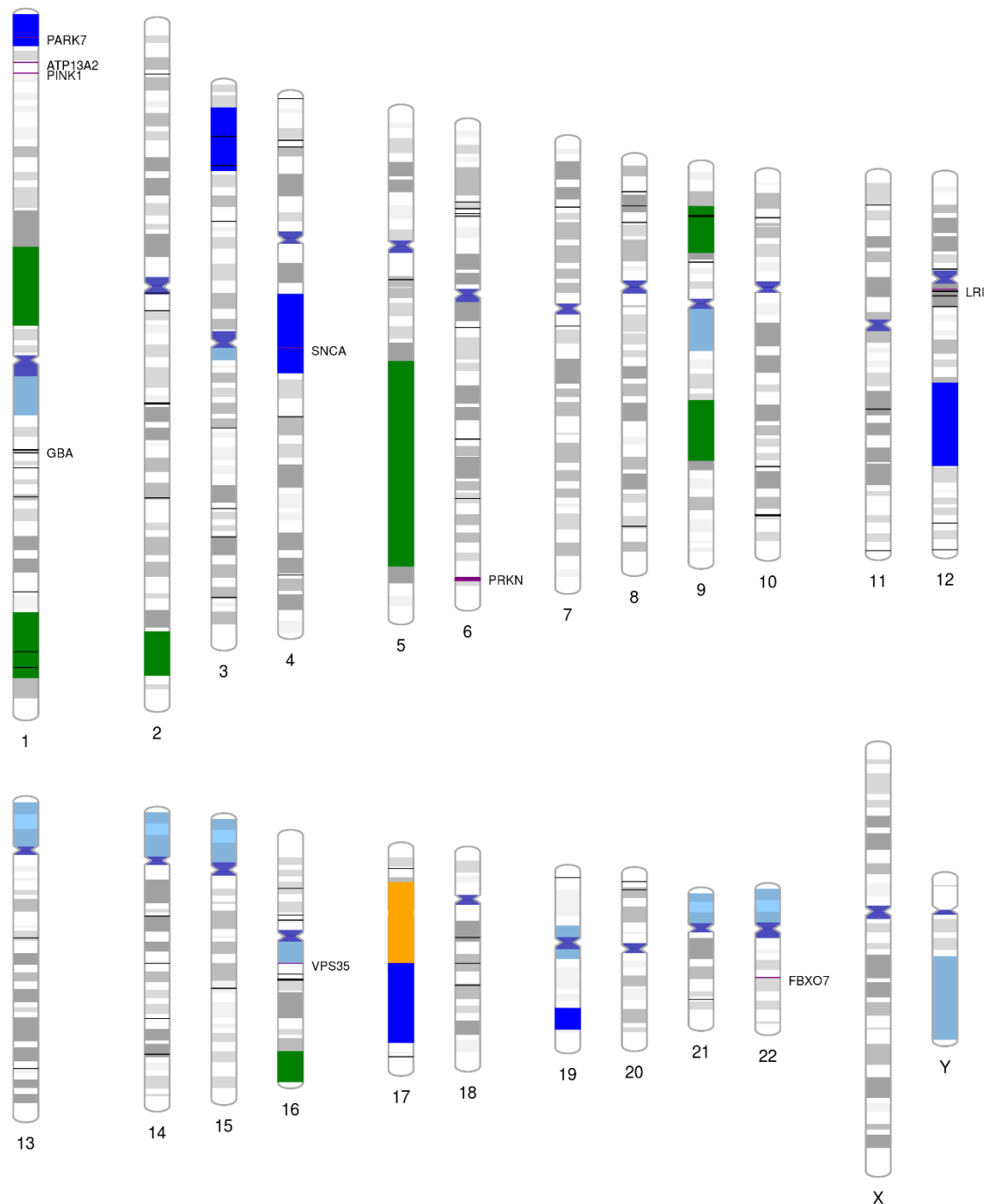
Παρά την δυνητική ικανότητα της ανωτέρω μεθοδολογίας με κάποιες προσθήκες να αναγνωρίσει και να συνδυάσει αξιόπιστα μελέτες διαφορετικών μεθοδολογιών υπήρξε ένας σημαντικός αριθμός δυσκολιών. Ο εντοπισμός και καταγραφή όλων των διαθέσιμων δεδομένων υπήρξε αδύνατος. Στην περίπτωση των GWLS υπήρξε σημαντική κάλυψη των δεδομένων με την χρήση του λογισμικού εξαγωγής από τα γραφήματα αλλά και πάλι υπήρξε απώλεια δεδομένων επιπλέον του ρίσκου Bias από την διαδικασία εξαγωγής. Στην περίπτωση των GWAS από την άλλη παρόλη την αναφερόμενη ελεύθερη διακίνηση των datasets είναι αδύνατο αυτά να ανακτηθούν από μεμονωμένους χρήστες ακόμη και μετά από επικοινωνία με τους συγγραφείς. Έτσι η απώλεια δεδομένων ήταν σημαντική και παρόλο που το METRADISC-XL μπορεί να την διαχειριστεί δεν παύει να μειώνει την ποιότητα των συμπερασμάτων και να αυξάνει την παρατηρούμενη ετερογένεια αν και οι πιθανότητες να προκύψουν ψευδώς θετικά αποτελέσματα είναι μικρές.

Ένα ακόμη πρόβλημα είναι αυτό της αντιστοίχισης ενός γενετικού χάρτη με έναν φυσικό χάρτη. Η λανθασμένη αναγνώριση της τοποθεσίας των markers η διαφορετική έκδοση του φυσικού χάρτη μπορεί να οδηγήσει σε συστηματικά σφάλματα και δυσχέρεια στην αναπαραγωγή των αποτελεσμάτων.

Ιδανικά σε αυτή τη διατριβή θα ήταν ενδιαφέρον να παρουσιαστούν και τα δεδομένα από τις μελέτες συσχέτισης γονιδίων που αφορούν τη νόσο Parkinson (Gene Association Studies – GAS) και να αναζητηθούν κοινά σημεία με τις μετα αναλύσεις των GWAS και των GWLS. Με δεδομένο ότι υπάρχουν πλέον δεκάδες χιλιάδες SNPs που έχουν μελετηθεί κάτι τέτοιο θα ήταν μη αποδοτικό.

Στην Εικόνα 15 φαίνονται συγκεντρωτικά τα σημαντικά bins της παρούσας μετα ανάλυσης των GWLS (μπλε) και των GWAS (πράσινο) καθώς και τα σημαντικότερα SNPs από

την μεγαλύτερη έως τώρα μετα ανάλυση GAS (40) και τα σημαντικότερα αναγνωρισμένα γονίδια της νόσου (80). Περιγράφεται έτσι μια γενική εικόνα της γενετικής αρχιτεκτονικής της νόσου Parkinson.



Εικόνα 15. Phenogramm. Σημαντικά bins μεταναλυσης GWLS (μπλε), GWAS(πράσινο) και κοινά (κίτρινο). Σημαντικά SNPs από τη μεγαλύτερη έως τώρα μεταναλυση GWAS (μαύρες γραμμές) (40). Πιο σημαντικά αιτιολογικά γονίδια της PD (80).

## BIBΛΙΟΓΡΑΦΙΑ

1. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263-5.
2. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. *PLoS computational biology*. 2012;8(12):e1002822.
3. Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB. *Campbell biology*: Pearson Boston; 2014.
4. Teare MD. *Genetic epidemiology*: Springer; 2011.
5. Jackson M, Marks L, May GH, Wilson JB. The genetic basis of disease. *Essays in biochemistry*. 2018;62(5):643-723.
6. Dausset J, Cann H, Cohen D, Lathrop M, Lalouel J-M, White R. Centre d'etude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. *Genomics*. 1990;6(3):575-7.
7. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *The American Journal of Human Genetics*. 1998;63(3):861-9.
8. Teare MD, Barrett JH. Genetic linkage studies. *The Lancet*. 2005;366(9490):1036-44.
9. Lipner EM, Greenberg DA. The rise and fall and rise of linkage analysis as a technique for finding and characterizing inherited influences on disease expression. *Disease gene identification*. 2018:381-97.
10. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296(5576):2225-9.
11. Hansel NN, Gao L, Rafaels N, Mathias R, Neptune E, Tankersley C, et al. Leptin receptor polymorphisms and lung function decline in COPD. *European Respiratory Journal*. 2009;34(1):103-10.
12. Lewis CM, Knight J. Introduction to genetic association studies. *Cold Spring Harbor Protocols*. 2012;2012(3):pdb.top068163.
13. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996;273(5281):1516-7.
14. Consortium IH. The international HapMap project. *Nature*. 2003;426(6968):789-96.
15. Consortium GP. A global reference for human genetic variation. *Nature*. 2015;526(7571):68.
16. Allison DB, Heo M. Meta-analysis of linkage data under worst-case conditions: a demonstration using the human OB region. *Genetics*. 1998;148(2):859-65.
17. Edwards AW. RA Fischer, statistical methods for research workers, (1925). *Landmark Writings in Western Mathematics 1640-1940*: Elsevier; 2005. p. 856-70.
18. Wise L, Lanchbury J, Lewis C. Meta-analysis of genome searches. *Annals of human genetics*. 1999;63(3):263-72.
19. Levinson DF, Levinson MD, Segurado R, Lewis CM. Genome scan meta-analysis of schizophrenia and bipolar disorder, part I: Methods and power analysis. *The American Journal of Human Genetics*. 2003;73(1):17-33.



20. Zintzaras E, Ioannidis JP. Heterogeneity testing in meta-analysis of genome searches. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*. 2005;28(2):123-37.
21. Zintzaras E, Ioannidis JP. HEGESMA: genome search meta-analysis and heterogeneity testing. *Bioinformatics*. 2005;21(18):3672-3.
22. Zintzaras E, Ioannidis JP. Meta-analysis for ranked discovery datasets: theoretical framework and empirical demonstration for microarrays. *Computational biology and chemistry*. 2008;32(1):39-47.
23. Evangelou E, Ioannidis J. Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*. 2013;14(6):379-89.
24. Zeggini E, Ioannidis JP. Meta-analysis in genome-wide association studies. 2009.
25. Kavvoura FK, Ioannidis J. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Human genetics*. 2008;123(1):1-14.
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials*. 1986;7(3):177-88.
27. Zintzaras E, Ioannidis JP. METRADISC-XL: a program for meta-analysis of multidimensional ranked discovery oriented datasets including microarrays. *Computer methods and programs in biomedicine*. 2012;108(3):1243-6.
28. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, et al. Parkinson disease. *Nature reviews Disease primers*. 2017;3(1):1-21.
29. Radhakrishnan DM, Goyal V. Parkinson's disease: A review. *Neurology India*. 2018;66(7):26.
30. Childbirth L, Care B, Assistance F. Parkinson Disease.
31. Tysnes O-B, Storstein A. Epidemiology of Parkinson's disease. *Journal of neural transmission*. 2017;124(8):901-5.
32. MacMahon Copas AN, McComish SF, Fletcher JM, Caldwell MA. The pathogenesis of Parkinson's disease: a complex interplay between astrocytes, microglia, and T lymphocytes? *Frontiers in Neurology*. 2021:771.
33. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Movement disorders*. 2015;30(12):1591-601.
34. Goetz C, Poewe W, Dubois B, Schrag A, Stern M, Lang A. MDS-Unified Parkinson's Disease Rating Scale (MDS-UPDRS). Available from the International Parkinson and Movement Disorder Society website: <https://www.movementdisorders.org/MDS/MDS-Rating-Scales/MDS-Unified-Parkinsons-Disease-Rating-Scale-MDS-UPDRS.htm>. 2008.
35. Armstrong MJ, Okun MS. Diagnosis and treatment of Parkinson disease: a review. *Jama*. 2020;323(6):548-60.
36. Funayama M, Nishioka K, Li Y, Hattori N. Molecular genetics of Parkinson's disease: Contributions and global trends. *Journal of Human Genetics*. 2022:1-6.
37. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in the  $\alpha$ -synuclein gene identified in families with Parkinson's disease. *science*. 1997;276(5321):2045-7.
38. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide B-MM, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS genetics*. 2012;8(3):e1002548.
39. Simón-Sánchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nature genetics*. 2009;41(12):1308-12.

40. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *The Lancet Neurology*. 2019;18(12):1091-102.
41. Drevon D, Fursa SR, Malcolm AL. Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior modification*. 2017;41(2):323-39.
42. Rohatgi A. Webplotdigitizer: Version 4.4. URL <https://automeris.io/WebPlotDigitizer>. 2020.
43. Van Wijk JJ, editor *The value of visualization*. VIS 05 IEEE Visualization, 2005; 2005: IEEE.
44. DeStefano AL, Golbe LI, Mark MH, Lazzarini AM, Maher NE, Saint-Hilaire M, et al. Genome-wide scan for Parkinson's disease: the GenePD Study. *Neurology*. 2001;57(6):1124-6.
45. Hicks AA, Petursson H, Jonsson T, Stefansson H, Johannsdottir HS, Sainz J, et al. A susceptibility gene for late-onset idiopathic Parkinson's disease. *Annals of neurology*. 2002;52(5):549-55.
46. Pankratz N, Nichols WC, Uniacke SK, Halter C, Murrell J, Rudolph A, et al. Genome-wide linkage analysis and evidence of gene-by-gene interactions in a sample of 362 multiplex Parkinson disease families. *Human molecular genetics*. 2003;12(20):2599-608.
47. Martinez M, Brice A, Vaughan JR, Zimprich A, Breteler MM, Meco G, et al. Genome-wide scan linkage analysis for Parkinson's disease: the European genetic study of Parkinson's disease. *Journal of medical genetics*. 2004;41(12):900-7.
48. Gao X, Martin ER, Liu Y, Mayhew G, Vance JM, Scott WK. Genome-wide linkage screen in familial Parkinson disease identifies loci on chromosomes 3 and 18. *American journal of human genetics*. 2009;84(4):499-504.
49. Cummings AC, Lee SL, McCauley JL, Jiang L, Crunk A, McFarland LL, et al. A genome-wide linkage screen in the Amish with Parkinson disease points to chromosome 6. *Ann Hum Genet*. 2011;75(3):351-8.
50. Bruggemann N, Kulper W, Hagenah J, Bauer P, Pattaro C, Tadic V, et al. Autosomal dominant Parkinson's disease in a large German pedigree. *Acta neurologica Scandinavica*. 2012;126(2):129-37.
51. GWAS Integrator: a bioinformatics tool to explore human genetic associations reported in published genome-wide association studies. [Internet]. 2011. Available from: <https://phgkb.cdc.gov/PHGKB/gWAHitStartPage.action>  
<https://www.nature.com/articles/ejhg201191.pdf>.
52. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic acids research*. 2019;47(D1):D1005-d12.
53. Andreou A. A Jupyter notebook used to match a list of research papers, base pairs and an associated p-value to their corresponding bin and filter for the smallest value.: [andreoua/basepair\\_to\\_bin](https://github.com/andreoua/basepair_to_bin). 2019.
54. Andreou A, Rikos D. Base pair to Bin Matching. Harvard Dataverse, 2019 2019. Report No.
55. Beecham GWD, D. W. | Scott, W. K. | Martin, E. R. | Schellenberg, G. | Nuytemans, K. | Larson, E. B. | Buxbaum, J. D. | Trojanowski, J. Q. | Van Deerlin, V. M. | Hurtig, H. I. | Mash, D. C. | Beach, T. G. | Troncoso, J. C. | Pletnikova, O. | Frosch, M. P. | Ghetti, B. | Foroud, T. M. | Honig, L. S. | Marder, K. | Vonsattel, J. P. | Goldman, S. M. | Vinters, H. V. | Ross, O. A. | Wszolek, Z. K. | Wang, L. | Dykxhoorn, D. M. | Pericak-Vance, M. A. | Montine, T. J. | Leverenz, J. B. | Dawson, T. M. | Vance, J. M. PARK10 is a major locus for sporadic neuropathologically confirmed Parkinson disease. *Neurology*. 2015;84(10):972-80.

56. Davis MFC, A. C. | D'Aoust, L. N. | Jiang, L. | Velez Edwards, D. R. | Laux, R. | Reinhart-Mercer, L. | Fuzzell, D. | Scott, W. K. | Pericak-Vance, M. A. | Lee, S. L. | Haines, J. L. Parkinson disease loci in the mid-western Amish. *Hum Genet.* 2013;132(11):1213-21.
57. Satake WN, Y. | Mizuta, I. | Hirota, Y. | Ito, C. | Kubo, M. | Kawaguchi, T. | Tsunoda, T. | Watanabe, M. | Takeda, A. | Tomiyama, H. | Nakashima, K. | Hasegawa, K. | Obata, F. | Yoshikawa, T. | Kawakami, H. | Sakoda, S. | Yamamoto, M. | Hattori, N. | Murata, M. | Nakamura, Y. | Toda, T. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nature genetics.* 2009;41(12):1303-7.
58. Vacic VO, L. J. | Clark, L. N. | Bar-Shira, A. | Gana-Weisz, M. | Gurevich, T. | Gusev, A. | Kedmi, M. | Kenny, E. E. | Liu, X. | Mejia-Santana, H. | Mirelman, A. | Raymond, D. | Saunders-Pullman, R. | Desnick, R. J. | Atzmon, G. | Burns, E. R. | Ostrer, H. | Hakonarson, H. | Bergman, A. | Barzilay, N. | Darvasi, A. | Peter, I. | Guha, S. | Lencz, T. | Giladi, N. | Marder, K. | Pe'er, I. | Bressman, S. B. | Orr-Urtreger, A. Genome-wide mapping of IBD segments in an Ashkenazi PD cohort identifies associated haplotypes. *Human molecular genetics.* 2014;23(17):4693-702.
59. Vacic V, Ozelius LJ, Clark LN, Bar-Shira A, Gana-Weisz M, Gurevich T, et al. Genome-wide mapping of IBD segments in an Ashkenazi PD cohort identifies associated haplotypes. *Human molecular genetics.* 2014;23(17):4693-702.
60. Spencer CC, Plagnol V, Strange A, Gardner M, Paisan-Ruiz C, Band G, et al. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Human molecular genetics.* 2011;20(2):345-53.
61. Simón-Sánchez J, van Hilten JJ, van de Warrenburg B, Post B, Berendse HW, Arepalli S, et al. Genome-wide association study confirms extant PD risk loci among the Dutch. *European journal of human genetics : EJHG.* 2011;19(6):655-61.
62. Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nature genetics.* 2009;41(12):1303-7.
63. Saad M, Lesage S, Saint-Pierre A, Corvol JC, Zelenika D, Lambert JC, et al. Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population. *Human molecular genetics.* 2011;20(3):615-27.
64. Pickrell JK, Berisa T, Liu JZ. Detection and interpretation of shared genetic influences on 42 human traits. 2016;48(7):709-17.
65. Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA, et al. High-resolution whole-genome association study of Parkinson disease. *American journal of human genetics.* 2005;77(5):685-93.
66. Liu X, Cheng R, Verbitsky M, Kisselev S, Browne A, Mejia-Sanatana H, et al. Genome-wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. *BMC medical genetics.* 2011;12:104.
67. Hu Y, Deng L, Zhang J, Fang X, Mei P, Cao X, et al. A Pooling Genome-Wide Association Study Combining a Pathway Analysis for Typical Sporadic Parkinson's Disease in the Han Population of Chinese Mainland. *Molecular neurobiology.* 2016;53(7):4302-18.
68. Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, Yearout D, et al. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nature genetics.* 2010;42(9):781-5.
69. Fung HC, Scholz S, Matarin M, Simón-Sánchez J, Hernandez D, Britton A, et al. Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *The Lancet Neurology.* 2006;5(11):911-6.
70. Foo JN, Tan LC, Irwan ID, Au WL, Low HQ, Prakash KM, et al. Genome-wide association study of Parkinson's disease in East Asians. *Human molecular genetics.* 2017;26(1):226-32.

71. Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, et al. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet.* 2010;74(2):97-109.
72. Do CB, Tung JY, Dorfman E, Kiefer AK, Drabant EM, Francke U, et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* 2011;7(6):e1002141.
73. Davis MF, Cummings AC, D'Aoust LN, Jiang L, Velez Edwards DR, Laux R, et al. Parkinson disease loci in the mid-western Amish. *Hum Genet.* 2013;132(11):1213-21.
74. Chang D, Nalls MA, Hallgrímsson IB, Hunkapiller J, van der Brug M, Cai F, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *2017;49(10):1511-6.*
75. Beecham GW, Dickson DW, Scott WK, Martin ER, Schellenberg G, Nuytemans K, et al. PARK10 is a major locus for sporadic neuropathologically confirmed Parkinson disease. *Neurology.* 2015;84(10):972-80.
76. Bandrés-Ciga S, Price TR, Barrero FJ, Escamilla-Sevilla F, Pelegrina J, Arepalli S, et al. Genome-wide assessment of Parkinson's disease in a Southern Spanish population. *Neurobiology of aging.* 2016;45:213.e3-e9.
77. Hauser MA, Li Y-J, Takeuchi S, Walters R, Nouredine M, Maready M, et al. Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. *Human molecular genetics.* 2003;12(6):671-7.
78. Kitsios GD, Zintzaras E. Genomic convergence of genome-wide investigations for complex traits. *Annals of human genetics.* 2009;73(5):514-9.
79. Talwar P, Silla Y, Grover S, Gupta M, Agarwal R, Kushwaha S, et al. Genomic convergence and network analysis approach to identify candidate genes in Alzheimer's disease. *BMC genomics.* 2014;15(1):1-16.
80. Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. *The Lancet Neurology.* 2020;19(2):170-8.

## **Γ ΜΕΡΟΣ: ΠΑΡΑΡΤΗΜΑ**

	observed mean val	observed Q values	observed Ha value	observed B value	right-sided P-value	right-sided P ordere	left-sided	left-sided	left-sided P-val
1.9	97	2816	122	518	0,00114	0,12494	0,10056	0,13022	0,08468
5.4	87,5	5621,5	178	761	0,0128	0,4593	0,4489	0,55116	0,46202
1.5	87,285713	3943,428467	137,714294	646	0,01354	0,20946	0,21286	0,21912	0,23336
16.4	86,785713	3987,928711	137,714294	643	0,01406	0,09294	0,21612	0,21658	0,22566
5.5	83,928574	7332,213867	192,571426	863	0,02688	0,19648	0,71686	0,67034	0,69736
1.8	79,071426	7412,214844	199,428574	889	0,06148	0,80736	0,72836	0,72352	0,75402
10.1	76,214287	7945,928711	202,285706	905	0,09608	0,978	0,79614	0,7485	0,7887
21.2	75,85714	8882,357422	229,142853	977	0,09814	0,96414	0,88296	0,90172	0,9086
17.2	75,85714	7648,357422	221,142883	871	0,09902	0,8382	0,75864	0,86384	0,71684
4.1	75,785713	6506,928711	177,714294	831	0,09922	0,74302	0,59438	0,54772	0,62678
10.4	75,85714	6846,857422	195,142853	866	0,10152	0,91896	0,64918	0,69242	0,70758
7.6	75,35714	4593,356934	163,142853	699	0,10622	0,70812	0,2962	0,42048	0,32872
9.4	75,285713	6292,928711	173,714294	811	0,10896	0,5834	0,55934	0,5135	0,57958
16.3	75,14286	8660,856445	238,857147	924	0,1115	0,47018	0,8666	0,93656	0,82598
2.9	74,14286	8080,356934	210,857147	941	0,12668	0,60618	0,8121	0,80802	0,85554
1.10	73,64286	9341,356445	244,142853	980	0,13676	0,6134	0,9158	0,95144	0,91276
4.5	73,64286	3005,357178	114,85714	559	0,1368	0,47194	0,11952	0,09888	0,1232
9.1	72,64286	7856,356934	212,857147	915	0,15218	0,63712	0,78468	0,81692	0,8076
10.3	71,85714	2372,857178	110,85714	490	0,16886	0,736	0,06968	0,08208	0,06626
7.5	71,214287	5836,928711	192,285706	758	0,18382	0,6932	0,48438	0,6666	0,45374
15.2	71,214287	2604,928467	118,285713	535	0,1852	0,8011	0,08352	0,11098	0,09716
13.3	70,85714	8859,356445	222,857147	973	0,19288	0,67844	0,87996	0,87272	0,9028
9.2	70,714287	8350,928711	234,285736	912	0,19388	0,59856	0,83666	0,92214	0,80304
3.7	70,64286	5448,356934	176,857147	770	0,19536	0,49802	0,42114	0,54146	0,48232
15.3	70,35714	2240,357178	103,85714	497	0,20022	0,46798	0,06202	0,05986	0,0706
17.1	70,214287	8307,928711	222,285706	937	0,20396	0,39152	0,8338	0,87128	0,8477
4.8	69,928574	3548,2146	146,428574	611	0,21106	0,36426	0,1673	0,27924	0,17882
7.4	69,64286	7177,356934	212,142853	866	0,21872	0,34386	0,70002	0,8157	0,70854
4.7	69,071426	7170,214355	211,428574	845	0,23466	0,42452	0,69342	0,80752	0,6573
1.6	68,928574	9181,213867	222,571426	1001	0,23532	0,35294	0,90298	0,87118	0,93408
4.4	68,785713	2393,928711	111,285713	511	0,24222	0,2869	0,07282	0,0853	0,08154
5.1	68,5	5470,5	172	763	0,2432	0,27276	0,42534	0,50142	0,4688
21.1	68,071426	8171,214355	217,428574	927	0,25714	0,30968	0,82028	0,84516	0,83228
12.6	68,071426	5639,214355	173,428574	792	0,25758	0,20552	0,45264	0,51316	0,3558
2.6	67,64286	3353,357178	107,428581	539	0,27256	0,23856	0,1511	0,07342	0,10312
7.2	66,5	8187,5	210	937	0,29608	0,43764	0,8221	0,80134	0,84878
7.3	66,5	7936,5	219	915	0,30048	0,56342	0,7915	0,85196	0,8059
16.2	66	8138	212	938	0,31572	0,5197	0,8115	0,8081	0,8454
11.5	65,714287	8526,928711	231,714294	950	0,32244	0,39458	0,85382	0,91246	0,86946
8.2	65,64286	10016,85742	252,142853	1016	0,32274	0,20886	0,94766	0,96634	0,94896
1.2	65,64286	7877,356934	220,142853	899	0,32438	0,3105	0,7915	0,8616	0,78048
15.4	65,785713	7988,928711	217,285736	929	0,32584	0,49054	0,80158	0,84484	0,83556
2.5	65,214287	6112,928711	174,714294	806	0,33618	0,25812	0,53036	0,52346	0,56926
10.2	65,214287	5360,928223	163,142883	736	0,33658	0,1671	0,4087	0,42126	0,40694
8.6	65,071426	5196,214844	151,571426	736	0,34104	0,1299	0,3851	0,32122	0,4068
22.2	64,928574	7591,214355	198,428574	915	0,34514	0,1008	0,75306	0,71998	0,80842
2.10	64,785713	4907,928711	153,714294	732	0,35462	0,07684	0,34154	0,34108	0,39844
1.3	64,35714	4890,357422	154,571411	647	0,3648	0,1065	0,33394	0,34294	0,23222
20.1	64,071426	10805,21387	261,571411	1055	0,3725	0,10926	0,97352	0,98094	0,97628
1.4	63,785713	9176,928711	223,285706	981	0,38004	0,11208	0,90456	0,87444	0,914
2.8	63,5	1691,5	99	427	0,39006	0,11492	0,0334	0,04752	0,03538
8.3	63,357143	6663,356934	177,857147	845	0,39504	0,0889	0,61702	0,54758	0,65724
19.3	62,714287	10247,92871	241,714294	1052	0,41446	0,185	0,9568	0,94322	0,9732
14.1	62,357143	4466,356934	137,857147	668	0,42414	0,21606	0,27618	0,21704	0,26696
19.4	62	11656,5	268	1115	0,4389	0,2505	0,98958	0,98796	0,99518
3.8	61,857143	4150,857422	132,857147	640	0,44102	0,13158	0,23944	0,18836	0,22352
5.2	61,857143	4460,356934	140,857147	691	0,44422	0,20852	0,27598	0,23964	0,31008
10.6	61,285713	7608,928223	204,285706	906	0,45696	0,23012	0,75128	0,75834	0,78754
1.7	61,214287	1694,928467	82,285713	409	0,4593	0,16898	0,03302	0,01782	0,02852

Εικόνα 16. Πίνακας αποτελεσμάτων μη σταθμισμένης ανάλυσης GWLS με HEGESMA (συνέχεια στην επόμενη σελίδα)

4.2	61,142857	8106,356934	215,142853	934	0,46464	0,06784	0,81618	0,83428	0,84552
14.4	61,142857	5476,357422	163,428558	749	0,46614	0,11928	0,42466	0,42026	0,43522
9.3	61	9373,5	250	953	0,46882	0,05096	0,91694	0,96566	0,87412
9.6	60,714287	5868,928711	174,285706	790	0,47536	0,02852	0,4887	0,5182	0,53064
20.2	60,714287	7264,928711	199,714294	888	0,47852	0,05592	0,71	0,728	0,75392
6.5	60,57143	9002,214844	212,571426	982	0,48596	0,0208	0,89322	0,8175	0,9145
13.2	60,285713	5810,928223	188,285706	790	0,49204	0,02262	0,48082	0,63784	0,5309
8.1	59,92857	6917,714355	206,428574	842	0,50396	0,01446	0,65754	0,77674	0,65024
14.2	60	8733,5	214	918	0,5051	0,02494	0,8719	0,82472	0,81474
16.1	59,642857	7737,356934	216,142853	903	0,5131	0,0157	0,76856	0,83538	0,78164
18.2	59	10015,5	234	1011	0,53412	0,04732	0,94768	0,92006	0,94494
13.4	58,642857	10010,85742	223,428558	1004	0,54608	0,06082	0,9495	0,87656	0,93952
2.1	58	8484,5	204	952	0,56518	0,13854	0,8481	0,75864	0,8702
10.5	57,714287	5654,928711	175,714294	790	0,5755	0,14558	0,4537	0,53094	0,52934
7.1	57,42857	7079,713867	184,285706	838	0,57976	0,05102	0,68306	0,6062	0,64238
18.3	57,5	5845	187	788	0,58054	0,07738	0,48728	0,62658	0,52558
2.4	57,5	5591,5	165	769	0,58376	0,13294	0,44466	0,43748	0,48154
3.6	55,714287	5710,928223	151,714294	742	0,63912	0,52196	0,4656	0,32438	0,42026
11.3	55,642857	6609,356934	181,428574	809	0,64014	0,42982	0,61206	0,58098	0,57652
19.1	55,142857	5490,356934	167,142853	766	0,65646	0,41424	0,43066	0,45748	0,47648
12.2	55,142857	8869,356445	213,142853	979	0,65908	0,538	0,88372	0,81978	0,91322
2.7	54,642857	5261,856934	164,142853	758	0,67152	0,51478	0,3953	0,42992	0,4586
5.3	54,5	6719	177	848	0,67354	0,45348	0,62868	0,54238	0,66578
6.3	54,42857	6788,214355	163,428574	826	0,67762	0,36576	0,63982	0,42306	0,6164
14.3	54,357143	6202,356934	170,571442	791	0,67962	0,28108	0,5445	0,48592	0,53228
11.4	54,285713	6196,929199	158,857147	750	0,68222	0,21204	0,54766	0,38652	0,44276
4.3	53,07143	5260,714355	165,571426	766	0,71596	0,57818	0,39548	0,44536	0,47684
20.3	51,714287	5338,928223	163,714294	737	0,75324	0,83922	0,40522	0,42756	0,4102
18.1	51,785713	1166,928589	67,714287	350	0,7537	0,88932	0,01386	0,00446	0,01376
2.2	51,57143	5491,214355	144,571426	728	0,75934	0,69406	0,42782	0,2671	0,38874
12.3	51,57143	6683,214355	182,571426	821	0,76142	0,7932	0,62362	0,5917	0,60362
3.3	51,285713	7658,928711	190,285706	857	0,76698	0,68368	0,7616	0,6534	0,68578
6.6	51	7296,5	192	867	0,7738	0,67184	0,71384	0,66854	0,7077
9.5	50,714287	7324,928711	183,142883	801	0,78054	0,658	0,71608	0,59244	0,55524
17.3	50,357143	6713,857422	185,857147	819	0,78846	0,5427	0,62864	0,61848	0,6014
6.2	50,5	2454,5	112	521	0,7893	0,61276	0,07448	0,08736	0,087
6.1	50,357143	5373,356934	169,142853	769	0,79096	0,42052	0,41124	0,47544	0,48212
1.1	50,07143	8229,714844	205,571426	917	0,79734	0,39846	0,82444	0,76866	0,81122
6.4	49,214287	4043,928467	130,285706	641	0,81578	0,58342	0,22472	0,17332	0,22338
3.1	49,142857	6083,356934	177,142853	795	0,82114	0,48458	0,52196	0,54022	0,53902
22.1	47,07143	2938,214111	127,571426	550	0,86408	0,91988	0,11088	0,15708	0,1118
15.1	45	3131,5	130	583	0,89784	0,99652	0,12638	0,16982	0,1441
18.4	44,785713	4754,428711	151,285706	711	0,90272	0,99396	0,3168	0,31756	0,35118
12.4	44,642857	5092,856934	162,142853	739	0,90532	0,9875	0,3654	0,41042	0,40978
8.4	44,57143	1889,214233	89,428558	438	0,90618	0,97318	0,0418	0,02746	0,039
11.2	44,285713	4137,928711	133,714294	639	0,90894	0,96188	0,23848	0,19424	0,22314
11.6	44	4027,5	124	582	0,91392	0,9474	0,22164	0,13684	0,14288
2.3	43,57143	2677,214355	118,571426	539	0,92024	0,93794	0,0911	0,11452	0,10174
12.5	43,142857	4386,357422	131,428558	594	0,92436	0,85546	0,27018	0,18068	0,16052
4.6	43,142857	2390,357178	115,142853	513	0,92704	0,92588	0,0687	0,09668	0,0795
13.1	42,642857	2249,857178	112,142853	497	0,93118	0,8406	0,06238	0,08776	0,07092
11.1	42,57143	5620,214355	160,571426	717	0,93476	0,73664	0,44768	0,39798	0,36622
5.6	41,785713	4381,428711	139,285706	661	0,94262	0,7662	0,2663	0,22994	0,25628
17.4	40,285713	2266,928711	101,285713	462	0,95708	0,88158	0,06202	0,05428	0,05026
19.2	39,357143	2347,857178	112,142853	507	0,96366	0,59262	0,06882	0,08984	0,07858
3.2	39,57143	1161,214233	82,428574	355	0,96434	0,86898	0,01354	0,01786	0,01464
12.1	39,42857	3784,214355	135,428574	607	0,96472	0,76016	0,19466	0,2043	0,1754
3.4	36,07143	2263,714355	108,571426	488	0,98326	0,88866	0,06546	0,07698	0,06736
3.5	35,57143	2457,714355	113,571426	511	0,98548	0,7716	0,07604	0,09342	0,08128
20.4	33,07143	859,714294	70,571426	301	0,99366	0,45814	0,0038	0,0053	0,0039
8.5	33,214287	876,428528	71,714287	304	0,9937	0,81204	0,00444	0,00624	0,00462

	observed mean v	observed	observed	observed	observed	right-sided	right-sided	left-sided	left-sided	left-sided
1.9	97	99,41876	402,6515	18,40595	10,43335	0,002	0,21746	0,11588	0,17206	0,13268
5.4	87,5	95,21817	573,555	21,60981	11,95535	0,0066	0,18258	0,24226	0,30512	0,2226
17.2	75,85714	91,70737	903,1815	29,03559	11,67172	0,01484	0,24862	0,52842	0,67972	0,205
2.9	74,14286	91,42691	1138,679	29,80296	15,46523	0,01562	0,10188	0,71692	0,71632	0,5141
5.5	83,928574	90,87207	752,618	23,06596	14,06763	0,01666	0,045	0,3942	0,37724	0,38566
16.4	86,785713	86,85727	274,4704	13,17026	8,675068	0,03412	0,24168	0,04786	0,04046	0,06062
9.4	75,285713	86,14124	558,0552	18,93848	10,19606	0,04018	0,1825	0,23034	0,1901	0,12198
9.2	70,714287	85,30233	1016,442	30,81792	13,2652	0,04576	0,154	0,61942	0,75702	0,31594
4.5	73,64286	83,87029	452,7742	16,97193	9,903738	0,05696	0,2051	0,15042	0,12482	0,10674
1.5	87,285713	83,76212	493,8607	18,07409	11,99851	0,05706	0,11556	0,18086	0,16036	0,22796
11.5	65,714287	83,29818	1330,535	34,89581	16,88419	0,0627	0,0808	0,82922	0,89368	0,63528
10.1	76,214287	83,07701	877,8301	24,77034	15,21545	0,0633	0,04322	0,5072	0,46724	0,48798
3.7	70,64286	80,38223	774,6044	24,42381	14,22344	0,08996	0,24384	0,41438	0,44814	0,3995
15.2	71,214287	79,04112	298,2234	14,70719	8,356269	0,1049	0,37468	0,05832	0,06716	0,05218
7.4	69,64286	78,40493	1143,7	31,56534	18,11404	0,11384	0,37222	0,71672	0,7896	0,74026
10.2	65,214287	78,0606	981,8962	26,63479	15,30927	0,11888	0,31422	0,59824	0,56322	0,49618
21.2	75,85714	77,75971	1390,549	34,35473	19,80289	0,12186	0,15462	0,86508	0,88406	0,86454
4.8	69,928574	77,86704	545,8822	21,78024	12,09834	0,12324	0,23338	0,22012	0,3144	0,23222
1.10	73,64286	77,42472	1520,978	37,04646	21,17407	0,12474	0,06508	0,91582	0,943	0,93836
7.6	75,35714	77,64465	555,1896	21,59126	12,89445	0,12598	0,0966	0,22974	0,30844	0,2912
13.2	60,285713	75,04186	920,1993	28,731	13,89722	0,16304	0,38324	0,54078	0,66442	0,37084
8.2	65,64286	74,88628	1653,979	39,3768	20,89619	0,166	0,29788	0,95306	0,97296	0,92554
2.6	67,64286	74,59639	583,1558	15,6122	11,20472	0,17206	0,25178	0,25056	0,0855	0,1737
7.5	71,214287	73,44861	872,8681	27,58984	15,00985	0,18856	0,42802	0,50386	0,61106	0,47016
5.2	61,857143	73,27623	673,4817	19,93662	12,8166	0,19482	0,23586	0,32884	0,23182	0,2861
4.4	68,785713	73,37531	249,7214	13,26347	8,447488	0,19498	0,32286	0,03664	0,03966	0,05228
8.6	65,071426	73,20692	900,0162	21,76249	15,61278	0,19704	0,1592	0,53028	0,31412	0,5271
15.3	70,35714	72,88303	185,3128	10,3545	7,348648	0,2028	0,1399	0,01894	0,01092	0,0285
13.3	70,85714	72,62533	1670,774	37,94378	21,9484	0,20586	0,06184	0,95652	0,95596	0,96604
9.1	72,64286	72,66745	702,774	23,09316	14,64987	0,20612	0,10602	0,35132	0,37758	0,43688
2.8	63,5	72,24338	236,3925	14,29764	6,4768	0,2157	0,05868	0,03264	0,05844	0,01502
2.10	64,785713	71,79481	1007,219	27,9544	17,3405	0,22388	0,06026	0,6171	0,62864	0,67702
10.4	75,85714	71,38981	949,3804	28,65856	16,97573	0,22988	0,06076	0,57226	0,6657	0,65006
9.3	61	71,30682	1484,624	38,16106	19,18256	0,23306	0,03616	0,90316	0,95978	0,82402
4.1	75,785713	70,90004	1059,898	27,89858	17,204	0,2404	0,03556	0,6572	0,6297	0,66534
17.1	70,214287	70,84942	1075,879	29,56912	18,0741	0,24414	0,01838	0,67012	0,70836	0,7396
12.6	68,071426	70,2351	921,6758	25,36309	16,95996	0,25374	0,02848	0,54666	0,4997	0,64532
3.6	55,714287	69,11565	1114,918	24,56606	16,57976	0,27994	0,05408	0,69632	0,45362	0,60986
8.3	63,357143	69,13028	906,6329	22,59043	16,03554	0,2803	0,0998	0,52782	0,35014	0,5575
9.6	60,714287	68,59019	616,3578	20,17913	12,18941	0,28838	0,03666	0,27724	0,24106	0,23538
10.3	71,85714	68,73986	351,0989	16,71026	9,32042	0,28938	0,05444	0,08682	0,1171	0,08426
14.1	62,357143	68,57786	718,4183	19,23472	13,82376	0,29346	0,01796	0,36688	0,20562	0,36834
2.5	65,214287	67,16783	1130,741	27,93509	18,49429	0,32066	0,0714	0,70784	0,6264	0,76838
1.3	64,35714	67,31421	611,3574	20,75435	12,82294	0,3211	0,09892	0,27132	0,26518	0,28388
8.1	59,92857	67,05752	1142,058	32,45021	17,059	0,3261	0,04552	0,71496	0,82336	0,64994
4.7	69,071426	67,00138	964,9305	29,35653	16,11979	0,32776	0,02588	0,5816	0,69664	0,56942
5.1	68,5	66,65433	470,4727	19,82746	11,50755	0,33926	0,02608	0,16244	0,22572	0,19486
14.4	61,142857	66,17372	660,9018	20,29267	13,61456	0,35026	0,03516	0,31454	0,2449	0,34702
1.7	61,214287	64,93453	157,4664	8,15291	6,003652	0,37936	0,16242	0,00988	0,00266	0,0108
11.4	54,285713	64,49206	1144,564	25,27185	17,24253	0,38648	0,18756	0,72304	0,49252	0,67372
10.6	61,285713	64,36614	1311,016	32,07937	20,23727	0,39256	0,1395	0,82174	0,80942	0,8885
1.8	79,071426	63,36861	1246,78	30,49611	17,56758	0,4144	0,3274	0,78758	0,74976	0,69558
14.2	60	63,23658	1389,244	32,22786	19,79365	0,42162	0,2623	0,8662	0,81758	0,86684
18.3	57,5	63,21684	932,7004	28,99295	15,81034	0,42332	0,1747	0,55416	0,6788	0,5424
2.7	54,642857	62,64085	1021,354	28,86804	17,06459	0,4378	0,24526	0,62506	0,66908	0,65172
2.4	57,5	61,00811	997,9175	25,66891	17,38022	0,48002	0,72054	0,60852	0,51078	0,67846
20.2	60,714287	60,34596	1009,61	28,08886	17,90307	0,49198	0,81908	0,61924	0,63614	0,72098
13.4	58,642857	60,31562	1688,534	34,92521	22,38826	0,4935	0,73452	0,96086	0,89868	0,97676
5.3	54,5	58,96676	1307,158	30,96486	20,06797	0,52826	0,94878	0,82448	0,77038	0,88344

Εικόνα 17. Πίνακας αποτελεσμάτων σταθμισμένης ανάλυσης GWLS με HEGESMA (συνέχεια στην επόμενη σελίδα)



6.3	54,42857	58,40001	1483,822	32,2133	21,41984	0,54504	0,96998	0,9041	0,81448	0,9488
21.1	68,071426	57,70942	1306,828	33,61626	18,83424	0,56436	0,98786	0,81818	0,86084	0,79496
10.5	57,714287	57,48953	782,0981	25,38821	15,61837	0,56736	0,98322	0,42268	0,4974	0,52284
7.3	66,5	57,28967	1065,16	29,48298	17,21783	0,57174	0,97648	0,65868	0,69932	0,66378
6.5	60,57143	56,92047	1319,666	30,79137	20,2903	0,57966	0,91984	0,828	0,76042	0,89412
1.2	65,64286	57,13078	1106,996	31,27308	16,64751	0,57994	0,96572	0,69092	0,77722	0,61668
19.3	62,714287	56,9125	1796,064	39,56575	23,02243	0,58354	0,86344	0,97836	0,97598	0,9884
20.1	64,071426	56,94791	1741,592	40,1826	21,4695	0,58658	0,95362	0,96912	0,97944	0,94818
1.4	63,785713	56,75221	947,86	25,98903	16,14256	0,58964	0,82592	0,56528	0,52832	0,5724
11.3	55,642857	56,5254	1012,066	25,86532	16,34568	0,5958	0,80176	0,61718	0,52042	0,58782
16.3	75,14286	56,34006	1350,676	35,64939	16,26613	0,6002	0,76056	0,8439	0,91584	0,5836
7.2	66,5	55,81875	1113,263	28,57752	17,06344	0,61154	0,81196	0,69624	0,65946	0,65106
19.1	55,142857	55,6441	1184,573	30,99003	18,90718	0,6139	0,77116	0,74802	0,76892	0,80412
4.3	53,07143	55,57906	964,4832	27,71513	17,24139	0,61526	0,57044	0,57826	0,61422	0,66452
15.4	65,785713	55,62787	1007,93	28,55335	16,52228	0,6153	0,6718	0,61554	0,65784	0,60408
18.2	59	55,54276	1468,516	34,86891	20,39487	0,61724	0,45996	0,8974	0,8974	0,89962
7.1	57,42857	55,24297	1306,493	31,12054	18,29999	0,62914	0,44518	0,82098	0,77228	0,75466
6.2	50,5	54,89168	497,3344	20,50165	12,00712	0,6347	0,45022	0,1835	0,25586	0,22804
1.6	68,928574	54,56154	1297,717	30,54782	18,13529	0,63802	0,44766	0,8167	0,74996	0,74318
4.2	61,142857	54,40168	1135,49	30,58392	18,1734	0,64824	0,3842	0,71282	0,75004	0,74616
17.3	50,357143	54,14003	1288,016	32,40658	19,10259	0,64974	0,35672	0,80944	0,82238	0,8168
22.2	64,928574	53,9638	847,408	23,44472	14,79854	0,65794	0,30246	0,47964	0,39874	0,4488
16.2	66	53,40199	1026,928	27,47195	15,73069	0,67008	0,37226	0,63304	0,60422	0,53542
3.8	61,857143	53,38556	433,2321	15,64969	9,986753	0,67204	0,26664	0,1355	0,0872	0,11056
18.1	51,785713	52,86917	109,2991	7,296021	5,36403	0,68146	0,31952	0,0036	0,00072	0,00368
9.5	50,714287	51,71175	904,6757	24,49124	15,63106	0,71046	0,59114	0,5263	0,44544	0,51894
11.6	44	48,79939	805,4854	21,45813	14,31958	0,77212	0,99162	0,4424	0,30014	0,40942
19.4	62	48,60736	1820,292	40,80161	21,25044	0,77702	0,98642	0,9814	0,98528	0,94206
12.3	51,57143	48,36714	954,0506	28,18458	16,39313	0,78446	0,98152	0,57046	0,63872	0,59196
14.3	54,357143	47,40434	709,5972	20,91013	12,68529	0,79942	0,99408	0,35996	0,27368	0,2739
6.6	51	47,34354	1309,436	32,61289	19,17566	0,80172	0,98774	0,82346	0,82932	0,8221
20.3	51,714287	47,18389	751,1072	24,57441	14,42831	0,80316	0,91666	0,39558	0,45442	0,41778
12.2	55,142857	47,21162	1660,921	36,55314	20,85275	0,80574	0,97832	0,95542	0,93394	0,9241
12.5	43,142857	47,1941	898,6201	23,69191	14,62787	0,80678	0,95568	0,51964	0,40608	0,4318
2.1	58	47,15044	1190,051	28,85756	17,85164	0,80716	0,85958	0,74964	0,67204	0,7194
13.1	42,642857	46,61838	422,7263	19,67464	10,99886	0,81304	0,87686	0,13032	0,22074	0,16386
16.1	59,642857	45,6807	1083,456	29,99178	15,21862	0,83284	0,9348	0,67152	0,71988	0,48718
8.4	44,57143	44,9268	233,9286	11,61116	8,047862	0,8466	0,9567	0,03164	0,02234	0,0415
22.1	47,07143	44,44922	474,8788	20,42293	11,35707	0,85392	0,91434	0,16438	0,25186	0,18404
3.1	49,142857	44,56342	1033,627	29,45282	17,09733	0,85558	0,94838	0,63602	0,69954	0,65522
3.3	51,285713	43,75209	643,2977	22,37269	12,79039	0,86354	0,9331	0,30142	0,3444	0,28208
17.4	40,285713	42,72188	435,5397	17,31969	10,78197	0,88148	0,93106	0,13826	0,13418	0,15188
2.2	51,57143	42,72279	621,0518	18,9895	11,96055	0,88268	0,96582	0,2807	0,19226	0,22172
3.2	39,57143	42,0483	188,9351	13,06557	7,433531	0,889	0,94028	0,0201	0,03892	0,03
15.1	45	41,85787	480,1776	20,73373	11,82592	0,89098	0,90514	0,16998	0,26682	0,21512
6.1	50,357143	41,72276	833,8959	26,44687	15,11848	0,89594	0,8521	0,47022	0,55342	0,48038
18.4	44,785713	41,22034	971,3609	28,00851	16,51066	0,90078	0,83624	0,58566	0,62896	0,60374
2.3	43,57143	39,67923	380,6279	18,35041	10,43127	0,91894	0,93148	0,10222	0,16728	0,13234
6.4	49,214287	39,18593	477,3831	17,38587	10,61915	0,92564	0,91512	0,1661	0,1345	0,13928
11.2	44,285713	37,48219	383,5315	15,27751	9,65049	0,93838	0,96884	0,1028	0,07714	0,0951
1.1	50,07143	37,00931	1186,48	31,41313	15,51872	0,94616	0,95422	0,74624	0,78596	0,5157
3.4	36,07143	35,40653	461,6662	19,3838	11,68793	0,95568	0,98124	0,15668	0,20988	0,20478
4.6	43,142857	35,38054	352,0935	17,78473	8,958564	0,95822	0,95086	0,08602	0,15038	0,06936
3.5	35,57143	33,78315	485,8582	20,07915	11,79618	0,96874	0,9765	0,17392	0,23458	0,21232
11.1	42,57143	33,21604	464,1425	18,99953	10,1098	0,97074	0,96068	0,15754	0,19312	0,11638
19.2	39,357143	32,69777	385,313	18,2795	10,10228	0,97302	0,62658	0,10668	0,16812	0,11728
5.6	41,785713	32,84061	700,8243	23,66508	12,30134	0,9736	0,80038	0,34986	0,40744	0,2479
12.4	44,642857	32,9281	717,3524	24,98683	12,22481	0,97368	0,91074	0,36596	0,47692	0,24394
8.5	33,214287	31,77727	127,8253	10,5172	6,146792	0,9776	0,5142	0,00604	0,01134	0,0126
12.1	39,42857	31,4823	344,9272	16,39245	9,008284	0,9795	0,07142	0,08368	0,10772	0,07348
20.4	33,07143	31,51994	123,0973	10,22893	6,018833	0,97962	0,27322	0,00546	0,0104	0,01112

BIN	gene	R mean	right p for R mean	left p for R mean	right p for R weighted mean	left p for R weighted mean	Q mean	right p for Q mean	left p for Q mean
4.4	32	45141167	0,0000	1E+16	0,0000	1E+16	5907285120000	0,1177	0,8823
12.4	82	19968572	0,0234	0,9770	0,0525	0,9480	71355630000	0,1420	0,8584
19.4	112	1496758	0,0247	0,9754	0,1019	0,8981	24048259	0,5040	0,4961
17.3	103	23183384	0,0271	0,9729	0,0080	0,9920	381469440000	0,2562	0,7438
17.2	102	50969116	0,0311	0,9689	0,0125	0,9875	5464047360000	0,4977	0,5023
3.2	22	11496949	0,0353	0,9647	0,0417	0,9583	64727785000	0,1155	0,8845
1.1	1	18539999	0,0431	0,9658	0,0431	0,9658	60412870000	0,9658	0,0431
20.4	116	6875513	0,0596	0,9404	0,4182	0,5818	15298963750	0,0496	0,9504
20.2	114	5928418	0,0610	0,9390	0,1415	0,8585	12547641250	0,2948	0,7052
3.7	27	19289648	0,0705	0,9295	0,2296	0,7704	192869900000	0,6735	0,3265
2.7	17	17164484	0,0852	0,9148	0,1020	0,8980	181901060000	0,1048	0,8952
11.2	74	14187698	0,0957	0,9043	0,0903	0,9097	94686080000	0,1175	0,8825
5.2	38	14094430	0,1056	0,8944	0,0542	0,9458	112697420000	0,4213	0,5787
13.3	87	5804133	0,1170	0,8830	0,1463	0,8537	14084941250	0,1086	0,8914
1.6	6	14172159	0,1189	0,8811	0,0492	0,9508	132682440000	0,1981	0,8019
3.1	21	14257026	0,1244	0,8756	0,3047	0,6953	79542525000	0,2829	0,7171
17.4	104	1050000	0,1334	0,8749	0,1334	0,8749	0	1E+16	1E+16
4.2	30	52525957	0,1335	0,8665	0,5410	0,4590	5601566720000	0,1356	0,8644
7.5	53	10407399	0,1375	0,8625	0,2238	0,7762	56129580000	0,0342	0,9658
2.9	19	5315205	0,1603	0,8397	0,2892	0,7108	17811152500	0,0290	0,9710
4.6	34	7905454	0,1666	0,8336	0,1979	0,8023	9537059375	0,3489	0,6513
7.4	52	8511105	0,1745	0,8255	0,5465	0,4535	61671950000	0,0104	0,9896
16.4	100	1234231	0,1825	0,8176	0,4228	0,5772	2172812	0,8480	0,1521
1.2	2	17979297	0,1825	0,8175	0,3464	0,6536	123350520000	0,5433	0,4567
2.5	15	970000	0,2000	0,8084	0,2000	0,8084	0	1E+16	1E+16
4.3	31	23405974	0,2161	0,7839	0,2151	0,7849	147907920000	0,4239	0,5762
11.6	78	13173921	0,2282	0,7718	0,6445	0,3555	108655180000	0,2894	0,7106
14.4	92	930000	0,2333	0,7749	0,2333	0,7749	0	1E+16	1E+16
2.10	20	6179897	0,2347	0,7653	0,3153	0,6847	10134494375	0,4549	0,5451
10.5	71	8762894	0,2357	0,7644	0,2867	0,7133	35205252500	0,3196	0,6804
2.6	16	19262880	0,2414	0,7586	0,4213	0,5787	121913380000	0,4669	0,5331
8.4	58	8719944	0,2434	0,7566	0,2173	0,7827	21582652500	0,1946	0,8054
6.5	47	18822003	0,2586	0,7415	0,7567	0,2433	202085860000	0,0835	0,9165
16.3	99	20865852	0,2774	0,7226	0,3921	0,6079	198615520000	0,2258	0,7742
9.6	66	1132143	0,2826	0,7175	0,4985	0,5016	2086632	0,8505	0,1496
18.2	106	860000	0,2911	0,7171	0,2911	0,7171	0	1E+16	1E+16
14.1	89	12716959	0,3033	0,6968	0,2476	0,7525	80830465000	0,6936	0,3065
4.8	36	18069777	0,3170	0,6830	0,7417	0,2583	190544640000	0,2024	0,7976
4.1	29	48208667	0,3197	0,6803	0,1323	0,8677	5594917760000	0,3598	0,6402
2.1	11	20557668	0,3301	0,6699	0,4182	0,5818	370811320000	0,2353	0,7647
2.4	14	780000	0,3575	0,6509	0,3575	0,6509	0	1E+16	1E+16
9.4	64	11230391	0,3620	0,6380	0,4157	0,5843	91820040000	0,1997	0,8003
7.2	50	18845793	0,3672	0,6328	0,4512	0,5488	169441440000	0,5687	0,4313
2.2	12	750000	0,3826	0,6256	0,3826	0,6256	0	1E+16	1E+16
10.4	70	8123543	0,3926	0,6074	0,2362	0,7638	23188060000	0,7352	0,2648
18.3	107	17399159	0,3995	0,6005	0,6161	0,3839	159966720000	0,5976	0,4024
19.2	110	7838750	0,4151	0,5855	0,4383	0,5620	10800825000	0,3660	0,6346
7.6	54	6224468	0,4152	0,5849	0,8507	0,1494	13250547500	0,2687	0,7314
3.5	25	1680542	0,4414	0,5586	0,4315	0,5685	459444648	0,2946	0,7054
5.6	42	5279304	0,4601	0,5399	0,2667	0,7333	8876189375	0,7940	0,2060
22.2	120	14294841	0,4629	0,5371	0,6064	0,3936	192403220000	0,5311	0,4689
6.1	43	16964050	0,4639	0,5361	0,7061	0,2939	215196300000	0,3724	0,6276
1.10	10	980989	0,4663	0,5338	0,7700	0,2301	7921076	0,7097	0,2904
10.3	69	21462039	0,4777	0,5223	0,6694	0,3306	389661400000	0,6187	0,3813
11.4	76	21329895	0,4840	0,5160	0,8083	0,1917	237544960000	0,0241	0,9759
1.8	8	54187471	0,4865	0,5135	0,4243	0,5757	5127086400000	0,5891	0,4109
15.3	95	6682830	0,4886	0,5114	0,7320	0,2680	22178392500	0,2026	0,7974
9.3	63	13112542	0,5157	0,4843	0,3074	0,6927	80742325000	0,7943	0,2057
1.4	4	18068668	0,5230	0,4770	0,2693	0,7307	170387260000	0,6151	0,3850
18.4	108	22509724	0,5312	0,4688	0,5482	0,4518	131651510000	0,7527	0,2473
1.7	7	17189047	0,5360	0,4640	0,2257	0,7743	177516320000	0,5086	0,4914
6.6	48	10878135	0,5405	0,4595	0,8620	0,1380	90589790000	0,2896	0,7104
21.1	117	4854428	0,5426	0,4574	0,1992	0,8008	11281203750	0,7191	0,2809
7.1	49	6808647	0,5482	0,4518	0,6715	0,3284	23569042500	0,2122	0,7878
8.6	60	19009960	0,5509	0,4491	0,4198	0,5802	304296820000	0,6995	0,3006
21.2	118	6804567	0,5522	0,4478	0,2669	0,7331	20310980000	0,5967	0,4033

14.2	90	16430923	0,5550	0,4450	0,3074	0,6926	156030040000	0,7445	0,2555
1.5	5	16009086	0,5647	0,4353	0,7983	0,2017	306291980000	0,2750	0,7250
1.9	9	17611077	0,5658	0,4342	0,6211	0,3789	158670310000	0,4510	0,5490
16.1	97	10118498	0,5983	0,4017	0,7098	0,2902	515105550000	0,6959	0,3041
15.2	94	12766842	0,6091	0,3909	0,6951	0,3049	929088500000	0,7919	0,2081
5.4	40	16025411	0,6208	0,3792	0,6641	0,3359	102135790000	0,3425	0,6575
6.4	46	13438474	0,6252	0,3748	0,6182	0,3818	111283560000	0,5072	0,4928
20.3	115	22723372	0,6268	0,3732	0,6576	0,3424	186374300000	0,6462	0,3538
8.1	55	18134020	0,6272	0,3728	0,6037	0,3963	298584060000	0,4618	0,5382
9.5	65	450000	0,6321	0,3763	0,6321	0,3763	0	1E+16	1E+16
12.5	83	64010571	0,6472	0,3528	0,2845	0,7155	47269392000000	0,8561	0,1439
3.4	24	9413696	0,6508	0,3492	0,8493	0,1507	306754800000	0,7738	0,2262
10.1	67	20958066	0,6759	0,3241	0,5694	0,4306	12744329000000	0,7869	0,2130
7.3	51	13899236	0,7129	0,2871	0,3653	0,6347	17290900000000	0,8392	0,1608
1.3	3	4908339	0,7132	0,2868	0,7918	0,2082	9402606250	0,7147	0,2853
12.6	84	5374611	0,7406	0,2594	0,8093	0,1907	12453296250	0,6833	0,3167
3.8	28	753901	0,7421	0,2580	0,8722	0,1279	1158208	0,8887	0,1113
8.5	59	741758	0,7544	0,2457	0,8927	0,1074	1930289	0,8567	0,1435
11.3	75	21178887	0,7676	0,2324	0,7775	0,2225	19720500000000	0,8680	0,1320
4.7	35	720055	0,7758	0,2243	0,6406	0,3594	33629009	0,4235	0,5766
17.1	101	24516201	0,8013	0,1987	0,7993	0,2007	25123574000000	0,8512	0,1488
6.2	44	50106494	0,8049	0,1951	0,2742	0,7258	52648739200000	0,6655	0,3345
9.1	61	11230499	0,8182	0,1818	0,5982	0,4018	918870800000	0,3205	0,6795
11.5	77	6280472	0,8270	0,1730	0,3352	0,6648	17129428750	0,7946	0,2054
11.1	73	19517618	0,8334	0,1666	0,8424	0,1576	26201362000000	0,6295	0,3706
3.6	26	4484896	0,8407	0,1593	0,7196	0,2804	9979286875	0,8830	0,1170
5.5	41	5012894	0,8439	0,1561	0,7266	0,2734	10823783750	0,7300	0,2700
12.1	79	19120490	0,8515	0,1485	0,7524	0,2476	19306526000000	0,7361	0,2639
9.2	62	10487115	0,8688	0,1312	0,8270	0,1730	971973500000	0,8047	0,1953
10.6	72	607912	0,8714	0,1286	0,6827	0,3173	47611636	0,3276	0,6725
5.3	39	16867650	0,8748	0,1252	0,9609	0,0391	15094978000000	0,9402	0,0598
20.1	113	14789094	0,8750	0,1250	0,9590	0,0410	772268300000	0,7913	0,2087
13.2	86	9947071	0,8783	0,1217	0,8765	0,1235	901123700000	0,1335	0,8665
19.1	109	6385396	0,8839	0,1160	0,8764	0,1236	27751675000	0,8899	0,1101
13.4	88	16920715	0,8886	0,1119	0,9488	0,0517	494910800000	0,9995	0,0010
6.3	45	19975479	0,8894	0,1106	0,6809	0,3191	19449936000000	0,7253	0,2747
18.1	105	12860614	0,8914	0,1086	0,7291	0,2710	800439300000	0,0981	0,9020
10.2	68	17584507	0,8951	0,1049	0,9351	0,0649	29102896000000	0,5565	0,4435
14.3	91	10019813	0,8952	0,1049	0,9760	0,0240	31587005000	0,8865	0,1135
15.1	93	15425850	0,9014	0,0986	0,9713	0,0287	943574500000	0,4241	0,5759
12.3	81	10816713	0,9023	0,0977	0,8269	0,1731	410725200000	0,9114	0,0886
8.2	56	12015631	0,9049	0,0951	0,6623	0,3377	916837500000	0,8226	0,1774
5.1	37	9464555	0,9073	0,0927	0,9688	0,0312	844621200000	0,6123	0,3877
15.4	96	4966809	0,9134	0,0866	0,9286	0,0714	8875757500	0,9823	0,0177
2.8	18	14694031	0,9367	0,0633	0,9609	0,0391	13272039000000	0,8220	0,1780
4.5	33	70000	0,9500	0,0584	0,9500	0,0584	0	1E+16	1E+16
19.3	111	7161289	0,9613	0,0387	0,9572	0,0428	12539858750	0,9233	0,0768
12.2	80	47862676	0,9613	0,0387	0,1658	0,8342	51228499200000	0,9449	0,0551
3.3	23	21452310	0,9669	0,0331	0,9449	0,0551	27706612000000	0,7284	0,2716
16.2	98	15877069	0,9750	0,0250	0,9705	0,0295	14254375000000	0,9572	0,0428
13.1	85	1077026	0,9783	0,0218	0,8952	0,1048	277534512	0,5765	0,4235
22.1	119	8925330	0,9847	0,0153	0,9865	0,0135	925818400000	0,9756	0,0244
2.3	13	20000	0,9917	0,0166	0,9917	0,0166	0	1E+16	1E+16
8.3	57	12417289	0,9980	0,0020	0,9923	0,0077	860209600000	0,9044	0,0957

Εικόνα 18. Συνολικά αποτελέσματα ανάλυσης GWAS με το METRADISC-XL