



**ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**  
**ΤΜΗΜΑ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ**



**ΑΚΑΔΗΜΙΑ ΑΘΗΝΩΝ**  
**ΙΔΡΥΜΑ ΙΑΤΡΟΒΙΟΛΟΓΙΚΩΝ ΕΡΕΥΝΩΝ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ**  
**“ΒΙΟΛΟΓΙΑ ΣΥΣΤΗΜΑΤΩΝ”**

**Μεταπτυχιακή Διπλωματική Εργασία**

Μετα-Ανάλυση δεδομένων RNA-Seq της νόσου του Πάρκινσον



**Κατίνα Μ. Σπάχο**

Επιβλέπων καθηγητής:  
Πολυδεύκης Χατζόπουλος, Καθηγητής ΓΠΑ

**ΑΘΗΝΑ**  
**2023**

ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ  
ΤΜΗΜΑ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΑΚΑΔΗΜΙΑ ΑΘΗΝΩΝ  
ΙΔΡΥΜΑ ΙΑΤΡΟΒΙΟΛΟΓΙΚΩΝ ΕΡΕΥΝΩΝ

**Μεταπτυχιακή Διπλωματική Εργασία**

Μετα-Ανάλυση δεδομένων RNA-Seq της νόσου του Πάρκινσον

“Meta-Analysis RNA-Seq data of Parkinson’s disease”

**Κατίνα Μ. Σπάχο**

Εξεταστική Επιτροπή:

Πολυδεύκης Χατζόπουλος, Καθηγητής ΓΠΑ (Επιβλέπων)

Ιωάννης Μιχαλόπουλος, Ειδικός Λειτουργικός Επιστήμονας Β', ΙΙΒΕΑΑ

Γεράσιμος Δάρας, Επίκουρος Καθηγητής ΓΠΑ

### Περίληψη

Η δεύτερη σε συχνότητα νευροεκφυλιστική νόσος του Πάρκινσον οφείλεται στη σταδιακή εκφύλιση των ντοπαμινεργικών νευρώνων της μέλαινας ουσίας των βασικών γαγγλίων και τη δημιουργία των σωματίων Lewy (Lewy Bodies), μετατρέποντας τους φυσιολογικούς ιστούς του εγκεφάλου σε παθολογικούς. Σκοπός της παρούσας εργασίας είναι η εύρεση γονιδίων που εκφράζονται διαφορετικά σε εγκεφάλους ασθενών με τη νόσο του Πάρκινσον και υγιών, με χρήση δεδομένων αλληλούχησης επόμενης γενιάς (RNA-Seq), προκειμένου να ταυτοποιηθούν οι βασικές διεργασίες που σχετίζονται με την εμφάνιση του φαινοτύπου της νόσου. Μέσω των βάσεων δεδομένων BioProject και ENA, με τη χρήση εργαλείων Βιοπληροφορικής, ανακτήθηκαν 4 έρευνες που αποτέλεσαν το υλικό της μελέτης. Ο υπολογισμός της διαφορικής έκφρασης των γονιδίων σε κάθε μελέτη διενεργήθηκε με το εργαλείο DESeq2. Εφαρμόστηκε η μέθοδος της μετα-ανάλυσης, με την οποία ταυτοποιήθηκαν 339 διαφορετικά εκφρασμένα γονίδια που σχετίζονται με τη νόσο του Πάρκινσον, εκ των οποίων τα 70 ήταν υποεκφρασμένα και τα 269 υπερεκφρασμένα. Οι προεξάρχουσες βιολογικές διαδικασίες στις οποίες συμμετέχουν τα γονίδια αυτά, βρέθηκαν μέσω του εργαλείου WebGestalt και της βάση δεδομένων STRING. Οι βιολογικές διεργασίες που σχετίζονται με τους υπερεκπροσωπημένους όρους των υποεκφρασμένων γονιδίων αφορούν τη μνήμη, τις νοητικές διαδικασίες, τις συνάψεις και γενικότερα τη μετάδοση του σήματος. Με στατιστική σημαντικότητα  $<0,05$ , τα κυριότερα εκ των υποεκφρασμένων γονιδίων ήταν το *BDNF*, η υποέκφραση του οποίου οδηγεί στην απώλεια νευροπροστασίας, το *TAC1* που σχετίζεται με τη νευροφλεγμονή, το *CRH* που μειώνει την απόκριση στο στρες επί Πάρκινσον, το *SST* που σχετίζεται με την παρουσία κινητικών και μη συμπτωμάτων και το *CALB1*, η υποέκφραση του οποίου προάγει το οξειδωτικό στρες. Οι βιολογικές διεργασίες που βρέθηκαν να σχετίζονται με τα υπερεκφρασμένα γονίδια σχετίζονταν με νευρολογικούς και νευροχημικούς όρους 5 εκ των 26 μοριακών συνοδών πρωτεϊνών – *HSPB1*, *DNAJB1*, *HSPA6*, *HSPA1B*, *HSPA1A*, εμπλέκονται με τις μοριακές λειτουργίες. Τα υπερεκφρασμένα γονίδια απαρτίζουν ένα πυκνό δίκτυο αλληλεπίδρασης πρωτεϊνών. Οι κυριότεροι υπερεκπροσωπημένοι όροι τους σχετίζονται με τη φλεγμονή. Τα κυριότερα υπερεκφρασμένα γονίδια που βρέθηκαν ( $FDR < 0,05$ ), με σημαντικές σχέσεις μεταξύ τους ήταν τα: *CSF1*, *IL-18*, *TLR5*, *CXCR4*, *HSPB1*, που σχετίζονται με την παρουσία νευροφλεγμονής και οξειδωτικού στρες και το *HSPA1B* που όταν υπερεκφράζεται αυξάνει την κινητική νευροπάθεια. Συμπερασματικά, η ταυτοποίηση των γονιδίων που εμπλέκονται με τη μετάδοση του νευρικού σήματος καθώς και με νευρολογικές, νευροχημικές και φλεγμονώδεις διεργασίες, μέσω της γενετικής έρευνας, μπορούν να οδηγήσουν στην αποτελεσματική εξατομικευμένη θεραπευτική προσέγγιση των ασθενών, αλλά και στην εύρεση νέων θεραπευτικών σχημάτων, προάγοντας τη νευροπροστασία και προλαμβάνοντας τη νευροφλεγμονή στη νόσο του Πάρκινσον.

**Επιστημονική περιοχή:** Νευροεκφυλιστικές διαταραχές

**Λέξεις κλειδιά:** Νόσος Πάρκινσον, γονιδιακό υπόβαθρο, εξατομικευμένη θεραπεία, ασυνουκλείνη, νευροεκφυλισμός, οξειδωτικό στρες, αλληλούχηση RNA-Seq, δίκτυα, μετα-ανάλυση, διαφορική έκφραση

## Meta-analysis RNA-Seq data of Parkinson's disease

MSc Systems Biology  
Department of Biotechnology AUA  
Institute for Biomedical Research AA

### Abstract

The second most common neurodegenerative of Parkinson's is due to the gradual degeneration of dopaminergic neurons of the substantia nigra of the basal ganglia and the formation of Lewy Bodies, turning normal brain tissues into abnormal ones. The purpose of the present study is to find genes that are differentially expressed in the brains of patients with Parkinson's disease and healthy ones, using next-generation sequencing (RNA-Seq) data, in order to identify basic processes related to the appearance of the disease phenotype. Through the Bioproject and ENA databases, with the use of Bioinformatics tools, 4 studies were retrieved, which constituted the material of the present research. Calculation of differential gene expression in each study was performed with DESeq2 tool. The method of meta-analysis was applied, which identified 339 differentially expressed genes associated with Parkinson's disease, of which 70 were underexpressed and 269 were overexpressed. The prominent biological processes in which these genes participate were found through the WebGestalt tool and the STRING database. The biological processes related to the over-represented terms of the underexpressed genes concern memory, metal processes, synapses and in general signal transmission. With statistical significance  $<0.05$ , the main underexpressed genes were *BDNF*, whose underexpression lead to the loss of neuroprotection, *TAC1* associated with neuroinflammation, *CRH* which reduces the stress response in Parkinson's, *SST* associated in the presence of motor and non-motor symptoms, and *CALB1*, the underexpression of which promotes oxidative stress. The biological processes found to be associated with the overexpressed genes were related to neurological and neurochemical terms. 5 of the 26 molecular chaperone proteins – HSPB1, DNAJB1, HSPA6, HSPA1B, HSPA1A, are involved in molecular functions. Overexpressed genes constitute a dense protein interaction network. Their main overrepresented terms are related to inflammation. The main overexpressed genes found ( $FDR < 0.05$ ), with significant relationships between them were: *CSF1*, *IL-18*, *TLR5*, *CXCR4*, *HSPB1*, related to the presence of neuroinflammation and oxidative stress, and *HSPAB* which, when overexpressed, increases moto neuropathy. In conclusion, the identification of genes involved in nerve signal transmission as well as neurological, neurochemical and inflammatory processes, through genetic research, can lead to the effective personalized therapeutic approach of patients, but also to finding of new therapeutic regimens, promoting neuroprotection and preventing neuroinflammation in Parkinson's disease.

**Scientific area:** Neurodegenerative disorders

**Key-words:** Parkinson's disease, genetic background, individualized therapy, a-synuclein, neurodegeneration, oxidative stress, RNA-seq sequencing, networks, meta-analysis, differential expression

## Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών, κατά το ακαδημαϊκό έτος 2021-2023, υπό την επίβλεψη του Ειδικού Λειτουργικού Επιστήμονα Β' κ. Ιωάννη Μιχαλόπουλο. Για τη διεκπεραίωση της εργασίας θα ήθελα να ευχαριστήσω θερμά πρώτα από όλα τον επιβλέποντα της εργασίας αυτής κ. Μιχαλόπουλο για την επιστημονική και ηθική υποστήριξη που μου παρείχε καθ' όλη τη διάρκεια εκπόνησης της εργασίας. Ιδιαίτερα, θα ήθελα να τον ευχαριστήσω για τις ανεκτίμητες επιστημονικές συμβουλές και τον αποτελεσματικό τρόπο αναφορικά με την επιλογή του θέματος, τις συμβουλές για τον τρόπο συγγραφής του, τη συνεχή καθοδήγηση και τις πολλές ώρες διδασκαλίας που μου αφιέρωσε στην προσπάθεια επίλυσης οποιασδήποτε απορίας. Αφενός η εμπιστοσύνη που μου έδειξε και αφετέρου η κατανόηση, η επιμονή και το αμείωτο ενδιαφέρον του αποτέλεσαν το κίνητρο μου για συνεχή προσπάθεια και γνώση.

Επιπλέον, θα ήθελα να ευχαριστήσω και τα υπόλοιπα μέλη της της εξεταστικής επιτροπής τον Καθηγητή κ. Πολυδεύκη Χατζόπουλο και τον Επίκουρο Καθηγητή κ. Γεράσιμο Δάρα για την αμέριστη εμπιστοσύνη και ευκαιρία που μου έδωσαν να παρακολουθήσω το μεταπτυχιακό αυτό πρόγραμμα.

Ένα βαθύ και ολόψυχο ευχαριστώ, οφείλω στον Κωνσταντίνο Κυριακίδη όπου η πολύπλευρη στήριξη και γνώση του, συνέβαλε σημαντικά στην ολοκλήρωση της εργασίας. Ακόμη, θα ήθελα να ευχαριστήσω εκ βάθους καρδιάς τον φίλο και συνάδελφο Θεοδωρή Φώσκολο για την συνεχή υποστήριξη και καθοδήγηση σε κάθε στάδιο ολοκλήρωσης της εργασίας. Ένα μεγάλο ευχαριστώ και στον συμφοιτητή μου Βασίλειο Ζωγόπουλο για την προθυμία, την καλοσύνη, τις συμβουλές και την ανεκτίμητη βοήθεια που μου προσφέρει όλα αυτά τα χρόνια.

Τέλος, θα ήθελα να εκφράσω ευγνωμοσύνη για την οικογένειά μου -το στήριγμα μου- που αποδέχτηκαν όλες τις επιλογές μου και με στήριξαν με αγάπη, δύναμη και περηφάνεια όλα αυτά τα χρόνια της ζωής και των ακαδημαϊκών σπουδών μου. Σε αυτούς οφείλω τα πάντα.

*“When you feel like quitting, think about why you started”*

# Πίνακας Περιεχομένων

<b>1. Εισαγωγή.....</b>	<b>9</b>
1.1 Νευροεκφυλισμός.....	9
1.2 Σύντομη ιστορική αναδρομή και παράγοντες κινδύνου της νόσου του Πάρκινσον	10
1.2.1 Συμπτώματα της νόσου .....	10
1.2.2 Ο ρόλος των νευροδιαβιβαστών στη νόσο του Πάρκινσον.....	13
1.2.3 Μοριακοί μηχανισμοί που εμπλέκονται στη νόσο του Πάρκινσον.....	15
1.2.3.1 Ο ρόλος της λανθασμένης αναδίπλωσης και συσσώρευσης πρωτεϊνών .....	16
1.2.3.2 Μονογονιδιακές μορφές νόσου Πάρκινσον .....	17
1.2.3.3 Προδιαθεσικοί παράγοντες κινδύνου για τη νόσο του Πάρκινσον .....	21
1.2.3.4 Παθογενετικοί μηχανισμοί στη νόσο του Πάρκινσον .....	22
1.2.3.4.1 Δυσλειτουργία των κυτταρικών μηχανισμών πρωτεϊνικής αποικοδόμησης.....	22
1.2.3.4.2 Οξειδωτικό στρες και μιτοχονδριακή δυσλειτουργία.....	23
1.2.3.4.3 Ο ρόλος της διεγερτικής τοξικότητας στη νόσο του Πάρκινσον.....	24
1.2.3.4.4 Ο ρόλος της νευροφλεγμονής στη νόσο του Πάρκινσον.....	24
1.2.3.5 Διάγνωση της νόσου.....	25
1.2.3.6 Θεραπεία της νόσου .....	26
1.3 Μέθοδοι αλληλούχησης .....	30
1.3.1 Μεταγραφωμική: εξέλιξη και μέθοδοι.....	31
1.3.2 Τεχνολογίες αλληλούχησης επόμενης γενιάς (NGS) .....	33
1.3.2.1 Κατασκευή βιβλιοθηκών .....	34
1.3.2.2 Τεχνολογία της Illumina.....	36
1.3.2.3 Αλληλούχηση RNA-seq .....	37
1.3.2.4 Ανάλυση Διαφορικής Έκφρασης Γονιδίων (Differential Gene Expression Analysis) 38	
1.3.2.5 Χρήση των τεχνολογιών NGS και εφαρμογές αυτών στην κλινική πράξη και έρευνα 38	
<b>2. Σκοπός .....</b>	<b>39</b>
<b>3. Μέθοδοι .....</b>	<b>39</b>
3.1 Λήψη πρωτογενών δεδομένων .....	39
3.2 Βιολογικές Βάσεις Δεδομένων.....	40
3.3 Συλλογή Δεδομένων.....	41
3.4 Paired-end και Single-end Δεδομένα .....	43

3.5	NGS Pipelines: Bcbio-nextgen.....	44
3.5.1	Workflow του Bcbio-nextgen.....	45
3.5.2	Διαφορική έκφραση γονιδίων.....	48
3.5.3	Μετα-ανάλυση δεδομένων γονιδιακής έκφρασης.....	52
3.5.4	Δίκτυο αλληλεπίδρασης σύμφωνα με τη βάση STRING και WebGestalt.....	52
<b>4.</b>	<b>Αποτελέσματα.....</b>	<b>53</b>
<b>5.</b>	<b>Συζήτηση.....</b>	<b>80</b>
<b>6.</b>	<b>Παράρτημα.....</b>	<b>85</b>
<b>7.</b>	<b>Βιβλιογραφία.....</b>	<b>92</b>



# 1. Εισαγωγή

## 1.1 Νευροεκφυλισμός

Οι νευροεκφυλιστικές ασθένειες επηρεάζουν εκατομμύρια ανθρώπους παγκοσμίως εξαιτίας της αύξησης του προσδόκιμου επιβίωσης καθώς και του υπερπληθυσμού παγκοσμίως. Ετυμολογικά, η λέξη αποτελείται από το πρόθεμα «νευρο-» που σημαίνει νευρώνες και «εκφυλισμός» το οποίο αναφέρεται στη διαδικασία απώλειας δομής ή λειτουργίας των ιστών ή ακόμη και των οργάνων. Πολλές από τις πιο συχνά εμφανιζόμενες νευροεκφυλιστικές παθήσεις είναι η νόσος του Alzheimer (AD), η νόσος του Parkinson (PD), η μετωποκροταφική άνοια (FTD), η προϊούσα υπερπυρηνική παράλυση (PSP), η άνοια με Σωματία Lewy (DLB), η νόσος Huntington (HD), η ατροφία πολλαπλών συστημάτων (MSA), η φλοιοβασική εκφύλιση (CBD) και τα νοσήματα Prion (Erkkinen et al., 2018). Στην πράξη οι νευροεκφυλιστικές παθήσεις επιφέρουν ετερογενή κλινικές και παθολογικές εκφράσεις τα οποία επηρεάζουν συγκεκριμένους νευρώνες και εμφανίζονται όταν τα νευρικά κύτταρα στον εγκέφαλο ή το περιφερικό νευρικό σύστημα δυσλειτουργούν με την πάροδο του χρόνου και τελικά πεθαίνουν. Η κατάσταση αυτή προκαλεί προβλήματα στην κίνηση και νοητική δυσλειτουργία (άνοια). Οι νευροεκφυλιστικές ασθένειες σχετίζονται με πολλαπλούς παράγοντες κινδύνου με την ηλικία να αποτελεί τον πιο σημαντικό παράγοντα. Η γήρανση από την άλλη επηρεάζει διαφορετικά τα νευροανατομικά μονοπάτια και ο εκφυλισμός αυτών των οδών οδηγεί στο σχηματισμό παθογόνων πρωτεϊνών οι οποίες και εξαπλώνονται σε ανατομικές οδούς. Το πιο κοινό χαρακτηριστικό μεταξύ των νευροεκφυλιστικών παθήσεων που συμβάλλουν στον νευροεκφυλισμό, εκτός από την απώλεια νευρωνικών κυττάρων είναι η φλεγμονή και το οξειδωτικό στρες (Singh et al., 2019; Stephenson et al., 2018). Αν και μια σειρά από μεταλλάξεις γονιδίων και περιβαλλοντικές τοξίνες σχετίζονται με νευροεκφυλιστικές διαταραχές, οι αιτιολογικοί παράγοντες παραμένουν ελάχιστα κατανοητοί. Οι περισσότερες περιπτώσεις νευροεκφυλιστικής νόσου είναι πολυπαραγοντικές στις οποίες αλληλοεπιδρούν μεταξύ τους εξωτερικοί περιβαλλοντικοί και εσωτερικοί γενετικοί παράγοντες κινδύνου οι οποίοι δρουν συνεργικά για να καθορίσουν το «αλλοστατικό φορτίο» ενός ατόμου. Το αλλοστατικό φορτίο καθορίζει το ρυθμό νευρικής γήρανσης και αποτέλεσμα αυτού είναι η διαφορική διάσπαση των νευροανατομικών οδών που επηρεάζονται από τη σχετική χρήση ή μη χρήση αυτών κατά τη διάρκεια ζωής. Συνέπεια αυτού, είναι ο σχηματισμός μιας ή περισσότερων παθογόνων πρωτεϊνών, μερικές από τις οποίες ενδέχεται να διαδίδονται μέσω του εγκεφάλου από τις αρχικές θέσεις σχηματισμού κατά μήκος νευροανατομικών οδών με σκοπό να επηρεάσουν τις συνδεδεμένες περιοχές του εγκεφάλου. Οι παθολογικές πρωτεΐνες και οι διάφορες συγκεντρώσεις είναι εν τέλει υπεύθυνες για την κλινική και παθολογική ποικιλία στους φαινοτύπους ασθενειών. Η ελαχιστοποίηση των διαφόρων αυτών παραγόντων σε συνδυασμό με τη μεγιστοποίηση της γνωστικής και της σωματικής άσκησης μπορεί να καθιερωθεί αναγκαία ούτως ώστε να μειωθεί η συχνότητα εμφάνισης των νευροεκφυλιστικών παθήσεων. Έτσι λοιπόν, είναι επιτακτική η ανάγκη για καλύτερη κατανόηση του τι προκαλεί νευροεκφυλιστικές ασθένειες και ποιες νέες καινοτόμες προσεγγίσεις για θεραπεία και πρόληψη θα συμβάλλουν σημαντικά σε αυτές (Armstrong, 2020).

## Νόσος του Πάρκινσον

### 1.2 Σύντομη ιστορική αναδρομή και παράγοντες κινδύνου της νόσου του Πάρκινσον

Η νόσος του Πάρκινσον είναι μια χρόνια νευροεκφυλιστική νόσος η οποία χαρακτηρίστηκε επισήμως για πρώτη φορά το 1817 από τον Βρετανό ιατρό **James Parkinson** την οποία και περιέγραψε ως «τρομώδη παράλυση» (Parkinson, 1817). Περιπτώσεις ανθρώπων που εμφάνιζαν τη νόσο είχαν σαν κύριο χαρακτηριστικό τον τρόμο ηρεμίας και τη διαταραχή στάσης και βάδισης. Μετά τη νόσο του Alzheimer η νόσος του Πάρκινσον αποτέλεσε τη δεύτερη πιο συχνή εκφυλιστική διαταραχή του κεντρικού νευρικού συστήματος με συχνότερη εμφάνιση στη τρίτη ηλικία και συγκεκριμένα ατόμων άνω των 65 ετών με συχνότητα 1% και ατόμων άνω των 80 ετών με συχνότητα 4% (Dexter and Jenner, 2013). Η αιτία πρόκλησης παραμένει ακόμη άγνωστη, αλλά χαρακτηρίζεται ως μια πολυπαραγοντική νόσος καθώς έχει βρεθεί ότι οφείλεται σε ορισμένους γενετικούς και άλλους περιβαλλοντικούς παράγοντες (Balestrino and Schapira, 2020; Wirdefeldt et al., 2011). Η ηλικία αποτελεί το βασικότερο παράγοντα κινδύνου καθώς όσο αυξάνεται η ηλικία αυξάνεται και η πιθανότητα νόσησης. Όλοι οι άνθρωποι ανεξαρτήτου εθνικότητας μπορούν να νοσήσουν, με το αρσενικό φύλο να είναι πιο επιρρεπές στη νόσο κατά δύο φορές περισσότερο σε σχέση με το θηλυκό φύλο αν και το ποσοστό θνησιμότητας των γυναικών είναι πιο υψηλό και η νόσος εξελίσσεται με ταχύτερους ρυθμούς από ότι στους άνδρες (Dahodwala et al., 2018). Ωστόσο υπάρχει διαφορά στα δύο φύλα τόσο στους μηχανισμούς που εμπλέκονται στη νόσο, όσο και στα συμπτώματα που αυτοί εμφανίζουν (Baldereschi et al., 2000; Dahodwala et al., 2018; Georgiev et al., 2017; Lee and Gilbert, 2016). Ορισμένοι περιβαλλοντικοί παράγοντες που συνδέονται με τον κίνδυνο εμφάνισης του Πάρκινσον είναι η έκθεση των ατόμων σε φυτοφάρμακα και η διαμονή σε αγροτική περιοχή. Επιπλέον, η έκθεση σε περιβαλλοντικές τοξίνες όπως τα επίπεδα μαγγανίου, διοξειδίου του άνθρακα καθώς και η εισπνοή του τριγλωροαιθυλενίου φαίνεται να σχετίζονται με την εμφάνιση κάποιου είδους παρκινσονισμού και να επηρεάζουν αρνητικά τη νόσο. Ωστόσο, αρνητική συσχέτιση με τη νόσο βρέθηκε να έχει η επίδραση του καπνίσματος, η κατανάλωση τσαγιού, καφέ και αλκοόλ (Breckenridge et al., 2016; Noyce et al., 2012). Τέλος, η ύπαρξη θετικού οικογενειακού ιστορικού διαφαίνεται να αποτελεί παράγοντα κινδύνου εμφάνισης της νόσου με τους συγγενείς πρώτου βαθμού να εμφανίζουν συχνότητα μεγαλύτερη από μία έως τρεις φορές σε σύγκριση με αυτούς που δεν έχουν κάποιο ιστορικό (Noyce et al., 2012). Σε παγκόσμιο επίπεδο ο αριθμός των ατόμων που πάσχουν από την χρόνια αυτή νόσο είναι πάνω από 6 εκατομμύρια (Armstrong and Okun, 2020b).

#### 1.2.1 Συμπτώματα της νόσου

Η επίδραση της νόσου του Πάρκινσον και ο τρόπος με τον οποίο αυτή εξελίσσεται διαφέρει σημαντικά μεταξύ των ατόμων. Υπάρχουν πολλές περιπτώσεις ανθρώπων όπου η κατάσταση τους απαιτεί διάστημα αρκετών χρόνων προκειμένου να επιδεινωθεί και να έχει σημαντικές επιπτώσεις στην καθημερινή τους ζωή.

Τα συμπτώματα της νόσου του Πάρκινσον ανάλογα με το στάδιο εμφάνισης τους μπορούν να κατηγοριοποιηθούν στις εξής ομάδες:

- Πρώιμα συμπτώματα
- Κινητικά συμπτώματα
- Μη κινητικά συμπτώματα

Σε πρώιμα στάδια της νόσου τα συμπτώματα είναι αδιόρατα και πολλές φορές μη συγκεκριμένα με αποτέλεσμα να μην είναι εμφανή στον πάσχοντα ή και να μην γνωρίζει ότι τα συμπτώματα αυτά σχετίζονται με τη νόσο. Περιλαμβάνουν αίσθημα κόπωσης, απώλεια δύναμης, απώλεια μυρωδιάς, απώλεια συγκέντρωσης, αργή γραφή, υπερβολική υπνηλία, διαταραχή συμπεριφοράς REM, κακή ποιότητα ύπνου αλλά και κατάθλιψη (Armstrong and Okun, 2020b; Olson et al., 2000). Ο συνδυασμός των συμπτωμάτων διαφέρει από άτομο σε άτομο και ο τρόπος που επηρεάζει τον καθένα μεταβάλλεται ακόμη και σε καθημερινή βάση. Όσο εξελίσσεται η νόσος δημιουργούνται όλο και περισσότερα προβλήματα τόσο στην κίνηση όσο και στην αυτονομία του πάσχοντα με αποτέλεσμα να χρήζει βοήθειας.

Τα πιο συχνά κινητικά συμπτώματα που εμφανίζονται στη νόσο του Πάρκινσον είναι ο **κινητικός τρόμος**, η **μυϊκή δυσκαμψία**, η **βραδυκινησία** και η **ορθοστατική αστάθεια**.

*Κινητικός τρόμος (resting tremor)*: Ο κινητικός τρόμος του χεριού, του ποδιού, της κεφαλής, της γνάθου, των χειλιών και της γλώσσας είναι τα πιο συχνά συμπτώματα στη νόσο του Πάρκινσον. Το πόδι και η γνάθος επηρεάζονται πρωτίστως παρόλο που στις περισσότερες περιπτώσεις ασθενών ο τρόμος παρατηρείται αρχικά στο χέρι. Στο 75% των ασθενών ο τρόμος επηρεάζει τη μια πλευρά του πάσχοντα και σε μεταγενέστερο στάδιο εξαπλώνονται εξ ολοκλήρου σε όλο το σώμα. Ο τρόμος αυτός γίνεται ευκολότερα αντιληπτός σε κατάσταση ηρεμίας των χεριών αλλά και σε κατάσταση πίεσης όπου υπάρχει ανεξέλεγκτη κίνηση (Goldenberg, 2008).

*Μυϊκή δυσκαμψία (Rigidity)*: Η μυϊκή δυσκαμψία αναφέρεται στην αύξηση του μυϊκού τόνου και αποτελεί το δεύτερο πιο συχνό σύμπτωμα που εμφανίζουν οι ασθενείς με την πάθηση. Πιο συγκεκριμένα, το άτομο αισθάνεται αδύναμο, άκαμπτο, πόνο και κράμπες στις αρθρώσεις καθώς οι μύες γίνονται πιο σφιχτοί και έτσι υπάρχει η αίσθηση του πόνου. Από τις πιο συνηθισμένες αρχικές εκδηλώσεις του Πάρκινσον είναι ο έντονος πόνος στον ώμο που αισθάνεται ο ασθενής και συνηθώς γίνεται λάθος διάγνωση του ως αρθρίτιδα. Αυτό συνεπάγεται την αγνόηση των συμπτωμάτων της νόσου και την εκδήλωση νέων συμπτωμάτων (Jankovic, 2008).

*Βραδυκινησία (Bradykinesia)*: Η βραδύτητα των κινήσεων ή αλλιώς βραδυκινησία αναφέρεται στις απρόβλεπτες κινήσεις, τη μείωση και την επιβράδυνση των αυτόνομων κινήσεων του ασθενούς. Διαταραχές στο βασικό σύστημα γαγγλίων που είναι υπεύθυνα για τη ρύθμιση της κινητικότητας με αυτόματο τρόπο έχουν σαν αποτέλεσμα τη βραδυκινησία. Είναι ένα από τα πιο σημαντικά συμπτώματα που μπορούν να επηρεάσουν σε μεγάλο βαθμό την καθημερινή ζωή των ατόμων καθώς οι

ασθενείς αδυνατούν να κάνουν γρήγορες κινήσεις ή ακόμη και πολλές κινήσεις ταυτόχρονα. Η βραδυκίνησια περιλαμβάνει την αργή γραφή, τη μειωμένη έκφραση προσώπου, την κίνηση του βραχίονα κατά τη διάρκεια του περπατήματος, τη μη δυνατότητα κατάποσης με αποτέλεσμα το σάλιασμα, αλλαγές του τόνου της φωνής και το μη φυσιολογικό ρυθμό στο ανοιγοκλείσιμο των ματιών (Berardelli et al., 2001).

Ορθοστατική αστάθεια (Postural instability): Η ορθοστατική αστάθεια και η δυσκολία στον συντονισμό εκδηλώνεται στα τελευταία στάδια της νόσου του Πάρκινσον, αφού πρώτα έχουν εμφανιστεί άλλα συμπτώματα. Η διαταραχή των αντανακλαστικών στήριξης και συντονισμού έχει σαν αποτέλεσμα την ανισορροπία του ασθενούς όταν αυτός στέκεται όρθιος γεγονός που οδηγεί σε συχνές πτώσεις. Η ορθοστατική αστάθεια μαζί με το «πάγωμα» της βάδισης δηλαδή τη ξαφνική ανικανότητα του ασθενούς να κινηθεί, αποτελεί ένα από τα σοβαρότερα προβλήματα αφού μπορεί να αποτελέσει συχνή αιτία πτώσεων και πρόκλησης διάφορων καταγμάτων. Σε πιο σοβαρές περιπτώσεις ασθενών που πάσχουν από τη νόσο του Πάρκινσον παρατηρείται ανικανότητα στην έναρξη μιας κίνησης και αδυναμία κίνησης των άκρων ακόμη και στις πιο απλές περιπτώσεις όπως για παράδειγμα στο πέρασμα ενός εμποδίου ή και τη μια στροφή γύρω από το σώμα τους (Bhidayasiri, 2005; Jankovic, 2008).

Τα πιο συχνά παρατηρούμενα μη κινητικά συμπτώματα σε ασθενείς με Πάρκινσον περιλαμβάνουν **κατάθλιψη, αϋπνία και γνωστική δυσλειτουργία.**

Κατάθλιψη (Depression): Η κατάθλιψη είναι το πιο συχνό σύμπτωμα που παρουσιάζεται στους πάσχοντες από τη νόσο και μάλιστα σε ποσοστό περίπου 35% σε αντίθεση με τους υγιείς ή ατόμων που πάσχουν από κάποια άλλη ασθένεια. Συνήθως εμφανίζεται στα πρώιμα στάδια της νόσου και τα βασικά της χαρακτηριστικά είναι η απώλεια ευχαρίστησης, η έντονη θλίψη, η μελαγχολία και η αρνητική στάση στην αντιμετώπιση της καθημερινότητας γεγονός που οδηγεί σε αυτοκαταστροφικές τάσεις. Μέρος του συνδρόμου της κατάθλιψης αποτελούν και άλλα συμπτώματα όπως αυτά είναι: απώλεια όρεξης, ανώμαλη έκφραση προσώπου και διαταραχές ύπνου. Τα συμπτώματα όμως αυτά είναι δυνατόν να εμφανίζονται σε ασθενείς με Πάρκινσον χωρίς όμως απαραίτητα να έχουν κατάθλιψη. Η διαφοροποίηση αυτή καθιστά δύσκολο τον διαχωρισμό των ασθενών με τη νόσο του Πάρκινσον και των ασθενών που απλά εμφανίζουν κατάθλιψη και πολλές φορές μπορεί να οδηγήσει σε λάθος διάγνωση (Aarsland et al., 2011; Kano et al., 2011).

Αϋπνία (Insomnia): Οι σημαντικότερες και πιο συχνές διαταραχές που είναι εμφανής σε ασθενείς με τη νόσο του Πάρκινσον περιλαμβάνουν αϋπνία κατά τη διάρκεια της νύχτας, υπερβολική υπνηλία κατά τη διάρκεια της ημέρας, διαταραχή συμπεριφοράς κατά τον ύπνο REM (REM sleep behavior disorder -RBD), σύνδρομο ανήσυχων ποδιών RLS (restless legs-syndrome) και εφιάλτες με αποτέλεσμα τη διαταραχή του ύπνου σε ποσοστό περίπου 60-70% των ατόμων (Schrempf et al., 2014). Η διαταραχή συμπεριφοράς κατά τον ύπνο REM (ύπνος γρήγορων κινήσεων των ματιών) χαρακτηρίζεται από απώλεια μυϊκής ατονίας εξαιτίας ορισμένης δυσλειτουργίας στο σύστημα του εγκεφάλου και έτσι το άτομο μπορεί να κινεί ανεξέλεγκτα τα άκρα του και να εκδηλώνει μια βίαιη συμπεριφορά κατά τη διάρκεια του ύπνου, που μπορεί να

οδηγήσει στον αυτό-τραυματισμό (Suzuki et al., 2011). Επιπλέον το σύνδρομο ανήσυχων ποδιών RLS εξαιτίας της επιτακτικής επιθυμίας του ασθενούς να κινήσει τα άκρα του κατά τη διάρκεια του ύπνου, οδηγεί στην κίνηση αυτών με αποτέλεσμα τη διαταραχή του. Κατά τη διάρκεια του ύπνου ασθενείς με τη νόσο ενδέχεται να παρουσιάσουν συχνή ούρηση, τρόμο, μυϊκή δυσκαμψία, έντονα όνειρα και ψευδαισθήσεις και αυτό δεν οφείλεται μόνο στην παθολογία της νόσου και των παρενεργειών των φαρμάκων που λαμβάνουν αλλά και στην ενδεχομένως υπάρχουσας μορφής κατάθλιψη που αυτοί βιώνουν (Larsen and Tandberg, 2001).

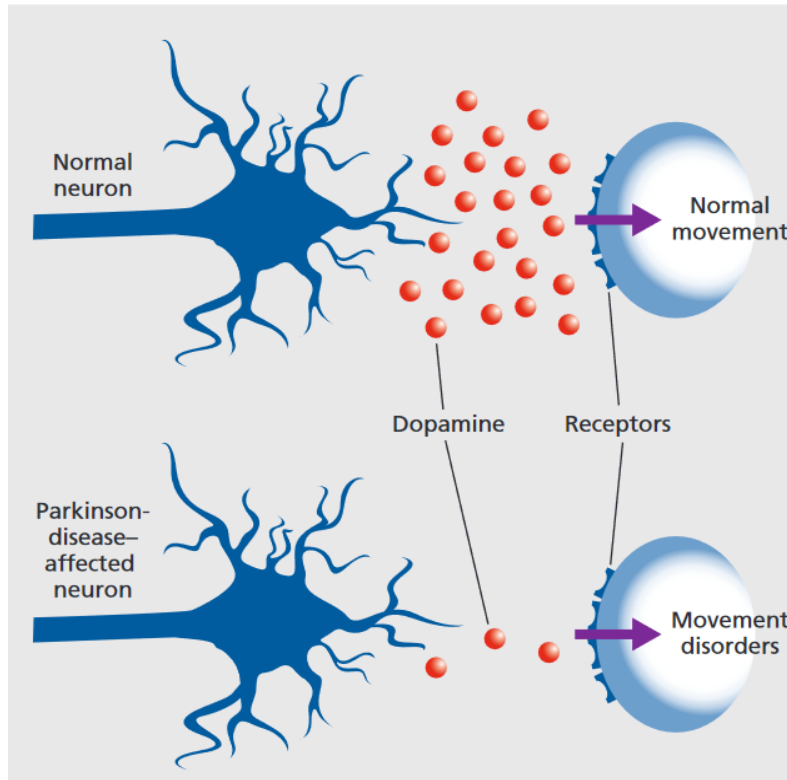
Γνωστική δυσλειτουργία και άνοια (Cognitive dysfunction and Dementia): Η διαταραχή των νοητικών λειτουργιών παραμένει ένα πολύ σημαντικό μη κινητικό σύμπτωμα που εμφανίζουν ασθενείς με τη νόσο του Πάρκινσον, καθώς η διάγνωση της είναι αρκετά περίπλοκη και απαιτητική. Η ήπια γνωστική διαταραχή (Mild Cognitive Impairment -MCI) είναι εμφανής κυρίως στα αρχικά στάδια της νόσου σε ποσοστό 20-30% (Aarsland et al., 2005) και η άνοια (Dementia) θα παρουσιαστεί μετά από 20 χρόνια ασθένειας σε ποσοστό περίπου 83% (Hely et al., 2008). Ο τρόμος, η κόπωση, η υπνηλία, η βραδυκινησία, ο πόνος, οι διάφορες διαταραχές στον ύπνο και η φαρμακευτική αγωγή είναι συμπτώματα ασθενών που μπορεί να επηρεάσουν σε σημαντικό βαθμό τη γνωστική λειτουργία. Η γνωστική δυσλειτουργία εκδηλώνεται με την βραδύτητα στην σκέψη, τη δυσκολία εύρεσης κατάλληλων λέξεων κατά την ομιλία τους, τη δυσκολία στην αυτό-παρακολούθηση και τη συγκέντρωση, την αδυναμία βούλησης, την αδεξιότητα στην ολοκλήρωση μιας εργασίας και πολλές φορές δεν υπάρχει ροή στον τρόπο σκέψης τους με αποτέλεσμα να μην μπορούν να γίνονται πλήρως κατανοητοί (Dirnberger and Jahanshahi, 2013; Meireles and Massano, 2012; Stuss, 2011).

### **1.2.2 Ο ρόλος των νευροδιαβιβαστών στη νόσο του Πάρκινσον**

Στον εγκέφαλο τα σήματα μεταφέρονται από ένα κύτταρο στο επόμενο μέσω των νευροδιαβιβαστών. Οι νευροδιαβιβαστές απελευθερώνονται στο μεσοδιάστημα μεταξύ του ενός κυττάρου με το άλλο όπου το δεύτερο περιλαμβάνει τους κατάλληλους υποδοχείς που θα μπορέσουν να συγκεντρώσουν αυτές τις χημικές ουσίες. Οι σημαντικότεροι νευροδιαβιβαστές είναι η ντοπαμίνη, η σεροτονίνη και η ακετυλοχολίνη.

- Ντοπαμίνη: Η κύρια περιοχή του εγκεφάλου που επηρεάζεται στη νόσο του Πάρκινσον είναι η μέλαινα ουσία (substantia nigra) η οποία είναι η βασική δομή γαγγλίων που βρίσκεται στο μεσεγκέφαλο και η παρουσία της είναι καθοριστική στην κίνηση. Στη μέλαινα ουσία υπάρχουν νευρώνες που εκκρίνουν μια βασική μονοαμίνη, την ντοπαμίνη που ανήκει στην οικογένεια κατεχολαμινών η οποία είναι ο πιο σημαντικός ρυθμιστικός νευροδιαβιβαστής. Ο ρόλος της είναι ουσιαστικός στη συμπεριφορά, τη γνωστική λειτουργία, τον ύπνο, τη συγκέντρωση και τη διάθεση. Τα βασικά παθολογικά χαρακτηριστικά της νόσου του Πάρκινσον είναι η απώλεια των ντοπαμινεργικών νευρώνων με αποχρωματισμό των **SNpc** (substantia nigra pars compacta) καθώς και η παρουσία των σωματίων **LBs** (σωμάτια Lewy) (Balestrino and Schapira, 2020).

Η εκφύλιση των ντοπαμινεργικών νευρώνων (νευρώνες με κύριο νευροδιαβιβαστή την ντοπαμίνη) στην περιοχή της μέλαινας ουσίας του μεσεγκεφάλου που συνδέονται μέσω της μελαινοραβδωτής οδού με το ραβδωτό σώμα έχει ως αποτέλεσμα το κινητικό τρόμο, τη μυϊκή δυσκαμψία, την ορθοστατική αστάθεια και τη βραδυκινησία. Έτσι λοιπόν η απώλεια των ντοπαμινεργικών νευρώνων οδηγεί στη μείωση της ντοπαμίνης του ραβδωτού, επομένως και κινητική δυσλειτουργία με τα πρώτα συμπτώματα να αρχίζουν να διαφαίνονται όταν τα ποσοστά της ντοπαμίνης έχουν μειωθεί πάνω από 80% (Masato et al., 2019) (Εικόνα 1).



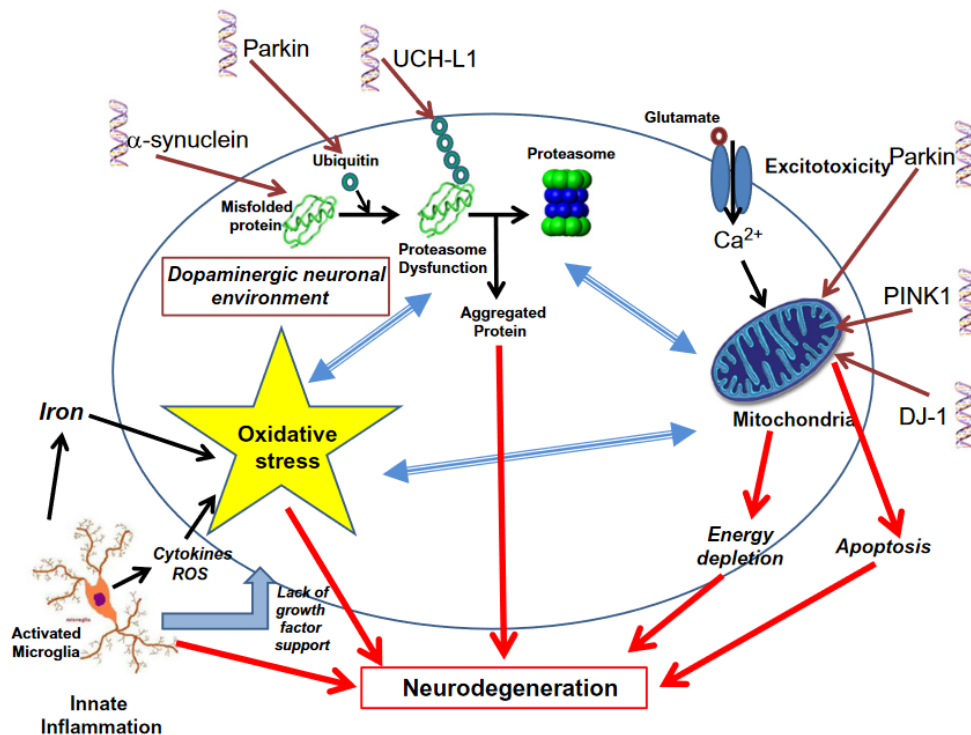
Εικόνα 1 Επίπεδα ντοπαμίνης σε φυσιολογικό και παθολογικό νευρώνα (Tiwari and Pal, 2017)

- Ακετυλοχολίνη: Η ακετυλοχολίνη (ACh) είναι επίσης ένας νευροδιαβιβαστής με καθοριστικό ρόλο τόσο στο κεντρικό, όσο και στο περιφερικό νευρικό σύστημα. Συντίθεται από ένα ένζυμο τη χολίνη ακετυλοτρανσφεράση (ChAT) η οποία καταλύει τη μεταφορά μιας ακετυλο-ομάδας από το συνένζυμο ακετυλο-CoA σε χολίνη για την παραγωγή της ακετυλοχολίνης. Οι χολινεργικοί νευρώνες χρησιμοποιούν τον νευροδιαβιβαστή ακετυλοχολίνη προκειμένου να μεταφέρουν μηνύματα μεταξύ των νευρικών κυττάρων. Στο κεντρικό νευρικό σύστημα ο ρόλος της ακετυλοχολίνης είναι εξίσου σημαντικός καθώς φαίνεται να σχετίζεται με τη μάθηση και τη μνήμη, αφού απελευθερώνεται στον εγκέφαλο κατά τη διάρκεια της μάθησης. Είναι απαραίτητη για τη σκέψη, τον ύπνο, τη μνήμη και τον έλεγχο των κινήσεων (Grossberg, 2017). Ο βασικός πρόσθιος εγκέφαλος είναι μια περιοχή με ιδιαίτερη λειτουργική σημασία λόγω του ρόλου που έχει στις γνωστικές

λειτουργίες της μνήμης και της μάθησης. Βασικές βλάβες στα κύτταρα που παράγουν ακετυλοχολίνη στο βασικό πρόσθιο εγκέφαλο και συγκεκριμένα στο βασικό πυρήνα Meynert (nbM) έχουν σαν αποτέλεσμα τη γνωστική εξασθένηση γεγονός που υποδηλώνει ότι το χολινεργικό σύστημα εμπλέκεται στη γνωστική δυσλειτουργία που παρατηρείται στη νόσο του Alzheimer (AD) και τη νόσο του Πάρκινσον (PD) (Lee et al., 2019). Η παραγωγή της ακετυλοχολίνης είναι άρρηκτα συνδεδεμένη με την ηλικία καθώς αυτή μειώνεται όσο αυξάνεται η ηλικία και έτσι οι ασθενείς σε μεγάλη ηλικία έχουν μειωμένη ικανότητα μάθησης, περιορισμένη ικανότητα μνήμης και πολλές φορές αναπτύσσουν σοβαρές γνωστικές διαταραχές που μπορούν να επηρεάσουν τις καθημερινές τους λειτουργικές ικανότητες (Murman, 2015).

### 1.2.3 Μοριακοί μηχανισμοί που εμπλέκονται στη νόσο του Πάρκινσον

Η νόσος του Πάρκινσον είναι μια περίπλοκη και πολυπαραγοντική νόσος στην οποία εμπλέκονται τόσο γενετικοί όσο και μη γενετικοί παράγοντες, όπως είναι οι περιβαλλοντικοί παράγοντες προκαλώντας κινητικά και μη κινητικά συμπτώματα λόγω της σταδιακής εκφύλισης των ντοπαμινεργικών νευρώνων της μέλαινας ουσίας των βασικών γαγγλίων. Οι μηχανισμοί που εμπλέκονται στη νόσο του Πάρκινσον περιλαμβάνουν τη λανθασμένη αναδίπλωση και συσσώρευση των πρωτεϊνών, τη βλάβη στις οδούς αποδόμησης των πρωτεϊνών, τη μιτοχονδριακή βλάβη, το οξειδωτικό στρες, τη νευροφλεγμονή, την διεγερτική τοξικότητα και τις διάφορες γονιδιακές μεταλλάξεις (Εικόνα 2).



Εικόνα 2 Βασικοί μοριακοί μηχανισμοί που εμπλέκονται στη νόσο του Πάρκινσον (Dexter and Jenner, 2013)

### 1.2.3.1 Ο ρόλος της λανθασμένης αναδίπλωσης και συσσώρευσης πρωτεϊνών

Συσσώρευση της α-συνουκλεΐνης (alpha- synuclein (SNCA)): Η α-συνουκλεΐνη, η βήτα- και η γάμμα-συνουκλεΐνη (140, 134, 127 αμινοξέα) ανήκουν στην οικογένεια των συνουκλεϊνικών πρωτεϊνών (Clayton and George, 1998). Είναι μικρές, διαλυτές πρωτεΐνες που εκφράζονται κυρίως σε νευρικό ιστό. Ιδιαίτερα η άλφα και βήτα-συνουκλεΐνη εντοπίζονται στον εγκεφαλικό ιστό στα προσυναπτικά άκρα των νευρωνικών συνάψεων, ενώ η γάμμα-συνουκλεΐνη κυρίως στο περιφερικό νευρικό σύστημα και στον αμφιβληστροειδή χιτώνα (George, 2002). Η α-συνουκλεΐνη κωδικοποιείται από το γονίδιο **SNCA** και οι πρώτες μεταλλάξεις που σχετίζονται με τη νόσο του Πάρκινσον με αυτοσωμικό επικρατές τρόπο κληρονόμησής του, αφορούσαν το γονίδιο που την κωδικοποιεί για αυτό το λόγο η πρωτεΐνη αυτή συνδέεται γενετικά και νευροπαθολογικά με τη νόσο (Polymeropoulos et al., 1997). Η α-συνουκλεΐνη είναι μια άφθονα εκφρασμένη νευρωνική πρωτεΐνη με καθοριστικό ρόλο στην κατανόηση των νευροεκφυλιστικών διαταραχών που καλούνται συνουκλεϊνοπάθειες στις οποίες περιλαμβάνεται η νόσος του Πάρκινσον (PD), η άνοια με σωματία Lewy (Dementia with Lewy bodies- DLB), η ατροφία πολλαπλών συστημάτων (Multiple system atrophy- MSA) και η νευροαξονική δυστροφία με κύριο χαρακτηριστικό τους τη συσσώρευση της α-συνουκλεΐνης στα ενδοκυτταρικά συσσωματώματα (Gai et al., 1998; Newell et al., 1999; Spillantini et al., 1997). Αυτές οι μη φυσιολογικές εναποθέσεις καλούνται σωματία Lewy bodies (LBs) και εντοπίζονται σε διάφορα σημεία του κεντρικού νευρικού συστήματος όπως είναι ο οσφρητικός βολβός, ο υποθάλαμος, ο πυρήνας Meynert, η μέλαινα ουσία, η παρεγκεφαλίδα και ο αμυγδαλοειδής πυρήνας (Wakabayashi et al., 2013). Η κύρια διαφορά μεταξύ των ασθενειών συνίσταται στη διασπορά των σωματίων Lewy και ανάλογα τις εμπλεκόμενες περιοχές του εγκεφάλου μπορεί να επηρεάσει την επεξεργασία πληροφοριών, τη μνήμη, τη γλώσσα, την σκέψη και την κινητικότητα. Αν και υπό φυσιολογικές συνθήκες η α-συνουκλεΐνη υπάρχει εντός του κυτταροπλάσματος ως ένα μονομερές και χωρίς να έχει κάποια συγκεκριμένη στερεοδιαμόρφωση, το κύριο χαρακτηριστικό του μορίου είναι ότι υφίσταται μια σταδιακή αλλαγή αρχικά σε μια αναδιπλωμένη α-ελικοειδή δευτεροταγή δομή η οποία είναι αρκετά επιρρεπής στο σχηματισμό oligομερών με αποδεδειγμένα νευροτοξική δράση και εν συνεχεία στη δημιουργία συμπλόκων ινιδίων με β-πυχωτή δομή και τελικώς τη δημιουργία αδιάλυτων μορφών όπως είναι τα σωματία Lewy στη νόσο του Πάρκινσον (Stefanis, 2012; Vekrellis et al., 2004). Ο ακριβής μηχανισμός για τη συσσώρευση της α-συνουκλεΐνης παραμένει ακόμη αδιευκρίνιστος αλλά υπάρχουν ευρήματα που υποστηρίζουν ότι διάφορες μετα-μεταφραστικές τροποποιήσεις της α-συνουκλεΐνης και ιδιαίτερα η φωσφορυλίωση στη θέση S129 (η πιο συχνή μετα-μεταφραστική τροποποίηση) εμπλέκεται στη διαδικασία συσσώρευσης της καθώς η αυξημένη έκφραση της στα κύτταρα μπορεί να προκαλέσει επίσης μιτοχονδριακό έλλειμμα, να προάγει το οξειδωτικό στρες και να οδηγήσει και σε κυτταρικό θάνατο (Anderson et al., 2006; Liu et al., 2011). Οι μετα-μεταφραστικές τροποποιήσεις των πρωτεϊνών αποτελούν συνέπεια του οξειδωτικού στρες με αποτέλεσμα αυτές να αποκτούν μη φυσιολογικές ιδιότητες και να συσσωρεύονται διαταράσσοντας έτσι και τους



μηχανισμούς αποικοδόμησης τους όπου συμβάλλουν στην παθολογία διάφορων νευροεκφυλιστικών καταστάσεων. Βασικό ρόλο στην αποικοδόμηση των πρωτεϊνών και τον έλεγχο του ολιγομερισμού της α-συνουκλεΐνης αποτελούν οι δυσλειτουργίες των οδών της Ουβικουιτίνης-πρωτεασώματος (UPS) και αυτοφαγίας-λυσosώματος (Rubinsztein, 2006). Μεταλλάξεις, διπλασιασμοί και τριπλασιασμοί του γονιδίου που κωδικοποιούν την α-συνουκλεΐνη (A53T, E46K, A30P) έχουν ταυτοποιηθεί ότι σχετίζονται με την εμφάνιση της νόσου του Πάρκινσον και είναι υπεύθυνες για τη μη σωστή αναδίπλωση των πρωτεϊνών και της συνάθροισης της α-συνουκλεΐνης στα σωματίδια Lewy. Η δημιουργία ολιγομερών μορφών θεωρείται τοξική για τα κύτταρα τα οποία προκύπτουν από την τάση που έχει η α-συνουκλεΐνη να πολυμερίζεται. Αναφορικά με την τοξικότητα της α-συνουκλεΐνης έχει προκύψει ότι η α-συνουκλεΐνη είτε συναθροίζεται με σκοπό να προστατεύσει το κύτταρο από άλλα πιο τοξικά ολιγομερή της είτε η συνάθροιση της είναι στοιχείο πρόβλεψης της τοξικότητας για τους νευρώνες (Goedert, 1999; McNaught et al., 2002).

**Ρόλος της πρωτεΐνης Ταυ (Tau):** Η πρωτεΐνη Ταυ η οποία κωδικοποιείται από το γονίδιο MAPT που βρίσκεται στο χρωμόσωμα 17q21 είναι μέλος της οικογένειας πρωτεϊνών που σχετίζονται με τους μικροσωληνίσκους (Microtubule-associated protein tau- MAPT) με καθοριστικό ρόλο τόσο στη σταθεροποίηση των μικροσωληνίσκων μέσω των οποίων πραγματοποιείται η μεταφορά των θρεπτικών συστατικών και άλλων βασικών ουσιών σε διαφορετικά μέρη των νευρώνων, όσο και στη σύνδεση στη μεμβράνη και στη ρύθμιση της αξονικής μεταφοράς (Brandt et al., 1995; Weingarten et al., 1975). Οι πρωτεΐνες Ταυ αν και είναι άφθονα κατανεμημένες σε νευρώνες του κεντρικού νευρικού συστήματος (ΚΝΣ), τα επίπεδα τους είναι αρκετά χαμηλά στα αστροκύτταρα και στα ολιγοδενδροκύτταρα του κεντρικού νευρικού συστήματος (Shin et al., 1991). Σε φυσιολογικές καταστάσεις η πρωτεΐνη Ταυ είναι διαλυτή και ξεδιπλωμένη, ωστόσο οι διάφορες αλλαγές στις μορφές ισομορφών ή υπερ- φωσφορυλίωση οδηγούν στην αναδίπλωση της και αυτή γίνεται αδιάλυτη και έτσι οι υπερ- φωσφορυλιωμένες Ταυ πρωτεΐνες συγκεντρώνονται μεταξύ τους για να σχηματίσουν ολιγομερή και να εξελιχθούν εν συνεχεία στη συσσώρευση των νευροϊνιδιακών μπερδεμάτων (Neurofibrillary tangles- NFTs) της πρωτεΐνης εντός των εγκεφαλικών κυττάρων με αποτέλεσμα την πρόκληση δυσλειτουργίας στους νευρώνες και στην αξονική μεταφορά (Alonso et al., 2001; King et al., 2000). Η συσσώρευση αυτή αποτελεί παθολογικό χαρακτηριστικό γνώρισμα πολλών νευροεκφυλιστικών διαταραχών συμπεριλαμβανομένου της νόσου του Αλτσχάιμερ και της νόσου του Πάρκινσον οι οποίες λόγω της συσσώρευσης καλούνται ταυροπαθήσεις (tauopathies). Μια εκτενής μελέτη σε επίπεδο γονιδιώματος (GWA) επιβεβαίωσε περιπτώσεις πως το γονίδιο MAPT συνδέεται στενά με το Πάρκινσον (Nalls et al., 2014).

### **1.2.3.2 Μονογονιδιακές μορφές νόσου Πάρκινσον**

Πολυετής έρευνα στη νόσο του Πάρκινσον οδήγησαν στην ανακάλυψη γονιδιακών μεταλλάξεων καθώς και στον εντοπισμό πολλών γενετικών παραγόντων κινδύνου που αυξάνουν τον κίνδυνο εμφάνισης της νόσου (**Πίνακας 1**). Οι μεταλλάξεις αυτές αντιπροσωπεύουν περίπου το 30% των περιπτώσεων που συνδέονται με οικογενείς μορφές της νόσου (Μενδελιανού τύπου κληρονομικότητας) και το 3-5% των

σποραδικών περιπτώσεων (Klein and Westenberger, 2012). Οι μεταλλάξεις στα γονίδια SNCA (PARK1/PARK4) και LRRK2 (PARK8) σχετίζονται με τη νόσο του Πάρκινσον με αυτοσωμικό επικρατές πρότυπο κληρονομικότητας ενώ μεταλλάξεις στα γονίδια Parkin (PARK2), PINK1 (PARK6), DJ-1 (PARK7) και ATP13A2 (PARK9) σχετίζονται με αυτοσωμικό υπολειπόμενο τρόπο κληρονομικότητας.

Locus	Map position	Gene/protein	Putative function	Inheritance pattern/ phenotype PD	Mutation in familial PD	Variants and risk of sporadic PD
PARK1/ PARK4	4q21	SNCA/ $\alpha$ -synuclein	Presynaptic protein, component of Lewy bodies, dopamine transmission	AD/EOPD with rapid progression and dementia, sporadic	PM and genomic duplication and triplication	Promoter polymorphism, 5' and 3' variants ↑ risk for PD
PARK2	6q25- q27	PARK2/E3 ubiquitin-protein ligase parkin	Ubiquitin E3 ligase, has neuroprotective function	AR/juvenile and EOPD with slow progression, dystonia; sporadic	Various mutations, exonic deletions, duplication, triplication	Promoter polymorphism and HM ↑ risk for PD
PARK5	4p14	UCH-L1/ubiquitin carboxyl-terminal hydrolase isozyme L1	Ubiquitin hydrolase	AD/LOPD	I93M	S18R variant may ↓ risk for PD
PARK6	1p35- p36	PINK1/serine, threonine-protein kinase PINK1	Mitochondrial protein kinase, has neuroprotective function	AR/EOPD with slow progression, tremor	PM and large deletions	HM may ↑ risk for PD
PARK7	1p36	PARK7/protein DJ-1	Chaperone, antioxidant	AR/EOPD, dystonia, psychiatric symptoms	PM and large deletions	HM may ↑ risk for PD
PARK8	12q12	LRRK2/leucine-rich repeat serine/threonine-protein kinase 2	Protein kinase, protect cells from stress-induced mitochondrial dysfunction	AD/LOPD, tremor	Different mutations and missense variants	HM may ↑ risk for PD
PARK9	1p36	ATP13A2/probable cation-transporting ATPase 13A2	Lysosomal ATPase	AR/juvenile Kufor-Rakeb syndrome, EOPD	PM	unknown

AD, autosomal-dominant; AR, autosomal-recessive; EO, early onset; LO, late onset; PM, point mutations; HM, heterozygous mutations.

Πίνακας 1 Συγκεντρωτικός πίνακας γονιδίων που εμπλέκονται στην παθογένεση της νόσου του Πάρκινσον (Shadrina *et al.*, 2010)

Παρακάτω περιγράφονται αναλυτικά τα γονίδια:

- SNCA (PARK1- PARK4):** Πρωτοφανής συσχέτιση με τη νόσο του Πάρκινσον αποτέλεσαν οι μεταλλάξεις του γονιδίου SNCA με αυτοσωμικό επικρατές πρότυπο κληρονομικότητας. Η πρωτεΐνη που κωδικοποιεί το γονίδιο αυτό αποτελεί θεμελιώδη συστατικό των σωματίων Lewy, των νευριτών Lewy καθώς και άλλων συνουκλεινοπαθειών (Spillantini *et al.*, 1997) (Spillantini *et al.*, 1998). Τρεις μεταλλάξεις του γονιδίου (A53T, A30P, E46K) που έχουν αναλυθεί, σχετίζονται με την εμφάνιση της νόσου του Πάρκινσον και είναι σπάνια αντιμετώπισιμες. Ιδιαίτερα η μετάλλαξη A53T είναι πιο συχνά εμφανιζόμενη και ασθενείς με τη μετάλλαξη αυτή εμφανίζουν πρόωμη ηλικία έναρξης ( $\leq 50$  ετών). Έχει ανακαλυφθεί σε τουλάχιστον 13 οικογένειες που αφορούν κυρίως ελληνικούς και ιταλικούς πληθυσμούς με ενδεχομένως έναν κοινό τους πρόγονο (Spira *et al.*, 2001). Οι ασθενείς με τη μετάλλαξη αυτή ανταποκρίνονταν θετικά στη λεβοντόπα αν και παρουσίαζαν προοδευτική γνωστική εξασθένηση. Τα χαρακτηριστικά που εμφάνιζαν ήταν τυπικά της ιδιοπαθούς νόσου του Πάρκινσον με επιπλέον βασικά χαρακτηριστικά όπως ορθοστατική υπόταση, ακράτεια ούρων και σοβαρό υποαερισμό (Klein and Westenberger, 2012) (Spira *et al.*, 2001). Διπλασιασμοί και τριπλασιασμοί του γονιδίου SNCA (PARK4) που έχουν βρεθεί επάγονται την αυξημένη έκφραση του γονιδίου η οποία θεωρείται τοξική. Προπαντός οι ασθενείς με τους τριπλασιασμούς είχαν πρόωμη έναρξη, ταχύτερη εξέλιξη της νόσου και έντονα συμπτώματα όπως άνοια και συχνή δυσαυτονομία (Ross *et al.*, 2008).

- LRRK2 (PARK8):** Μεταλλάξεις στο γονίδιο LRRK2 με ιδιαίτερα πιο συχνή την μετάλλαξη Gly2019Ser, έχουν κεντρίσει το επιστημονικό ενδιαφέρον καθώς αποτελούν τη συχνότερη αιτία που παρατηρείται τόσο σε ασθενείς με αυτοσωμικό επικρατή τύπο κληρονομικότητας όσο και σε ασθενείς με σποραδική μορφή της νόσου του Πάρκινσον. Αντιπροσωπεύουν το 10% των περιπτώσεων με αυτοσωμικό επικρατές τρόπο κληρονόμησης και το 3,6% των σποραδικών περιπτώσεων (Lesage and Brice, 2009). Κλινικά χαρακτηριστικά των ασθενών είναι η καθυστερημένη έναρξη της νόσου με θετική ανταπόκριση στη θεραπεία με λεβοντόπα και την άνοια να μην αποτελεί συχνό σύμπτωμα (Klein and Westenberger, 2012). Επιπλέον, έχει βρεθεί ότι το LRRK2 εντοπίζεται ενδοκυτταρικά όπως στα μιτοχόνδρια, στη συσκευή Golgi, στις λιπιδικές στοιβάδες (rafts), στο ενδοπλασματικό δίκτυο και τον κυτταροσκελετό που υποδηλώνουν έτσι την πιθανή συσχέτιση του LRRK2 στη μεταφορά του νευρικού σήματος καθώς έχει βρεθεί ότι διάφορες τροποποιήσεις στη λειτουργία του LRRK2 στις λιπιδικές στοιβάδες συμβάλλουν στη παθογένεση της νόσου του Πάρκινσον (Brown and London, 1998) (Hatano et al., 2007).
- Parkin (PARK2):** Καθοριστικής σημασίας στην εμφάνιση της νόσου του Πάρκινσον όπως έχει ήδη προαναφερθεί αποτελούν οι γενετικοί παράγοντες και ειδικότερα η ταυτοποίηση των διαφόρων γονιδιακών μεταλλάξεων. Ένα από τα γονίδια που σχετίζεται με την πρόωμη ανάπτυξη και μάλιστα με την έναρξη της νόσου του Πάρκινσον σε ηλικία ίση ή ακόμη και μικρότερη των 40 ετών με αυτοσωμικό υπολειπόμενο τρόπο κληρονομικότητας είναι το γονίδιο PARK2. Ωστόσο οι μεταλλάξεις του γονιδίου προκαλούν τη νεανική μορφή της νόσου του Πάρκινσον για το λόγο του ότι τα συμπτώματα των ασθενών έχουν ήδη ξεκινήσει όντας σε ηλικία  $\leq 21$  ετών (Klein and Westenberger, 2012). Το γονίδιο PARK2 κωδικοποιεί την πρωτεΐνη Parkin που είναι μια E3 λιγάση ουβικουιτίνης υπεύθυνη για τη διαδικασία της μιτοφαγίας στα νευρικά κύτταρα. Η πρωτεΐνη Parkin εμπλέκεται στη νόσο του Πάρκινσον καθώς προάγει μέσω της διαδικασίας της αυτοφαγίας τόσο την απάλειψη όσο και την καταστροφή των μιτοχονδρίων που έχουν υποστεί κάποια βλάβη. Η νόσος του Πάρκινσον παθολογικά έχει χαρακτηριστεί από τη συσσώρευση ενδοκυτταρικών πρωτεϊνών και βασικός ρόλος της πρωτεΐνης Parkin είναι να ενεργοποιεί το σύστημα πρωτεασώματος-ουβικουιτίνης για εκτενή αποδόμηση των πρωτεϊνών της εξωτερικής μεμβράνης. Έτσι λοιπόν η μετάλλαξη στο Parkin θα εξασθενίσει και θα προκαλέσει βλάβη στις οδούς αποδόμησης των πρωτεϊνών με αποτέλεσμα αυτό να οδηγήσει σε τοξική συσσώρευση λανθασμένων ή αναδιπλωμένων πρωτεϊνών και αυτά με τη σειρά τους να οδηγήσουν σε νευρωνικό θάνατο (Chan et al., 2011). Τέλος, όσο αφορά την κλινική εικόνα αυτών των ασθενών, τα συμπτώματα τους εμφανίζονται πολύ νωρίτερα, η δυστονία και η υπερρεφλεξία είναι πιο συχνές, η νόσος εξελίσσεται πιο αργά και η ανταπόκριση στη λεβοντόπα είναι συχνά μεγαλύτερη (Lohmann et al., 2003).

- PINK1 (PARK6):** Η δεύτερη πιο κοινή αιτία που σχετίζεται με το Πάρκινσον με αυτοσωμικό υπολειπόμενο τρόπο κληρονομικότητας και με πρόωμη έναρξη είναι οι μεταλλάξεις του γονιδίου της PINK1 (Klein and Westenberger, 2012). Οι ασθενείς με τις μεταλλάξεις του γονιδίου της PINK1 χαρακτηρίζονται από πρόωμη έναρξη (από 32 έως 48 ετών) παρκινσονισμού με αργή εξέλιξη, παρατεταμένη απόκριση στη λεβοντόπα με συχνές δυσκινησίες που προκαλούνται λόγω της αγωγής με τη λεβοντόπα χωρίς όμως να παρουσιάζουν δυστονικές διαταραχές (Valente et al., 2001). Έχει συνεργική δράση με το Parkin καθώς λειτουργούν σε μια κοινή οδό προκειμένου να διασφαλιστεί η ακεραιότητα των μιτοχονδρίων. Στις περιπτώσεις που τα μιτοχόνδρια έχουν υποστεί κάποια βλάβη το PINK1 σταθεροποιείται στα μιτοχόνδρια με χαμηλότερο δυναμικό μεμβράνης όπου στο σημείο αυτό προσλαμβάνεται το Parkin το οποία αρχικά δεν είναι ενζυμικά ενεργό και στη συνέχεια αφού αποκτήσει ενζυμική ενεργότητα ξεκινάει την αυτοφαγική κάθαρση των μιτοχονδρίων (Youle and Narendra, 2011). Οι διάφορες τροποποιήσεις στο μηχανισμό παρακολούθησης και διάσπασης των μη λειτουργικών μιτοχονδρίων μπορεί να συμβάλλει στη συσσώρευση εκφυλιστικών μιτοχονδρίων (Chan et al., 2011). Όπως έχει ήδη αναφερθεί παραπάνω, ο ρόλος των μετα-μεταφραστικών τροποποιήσεων είναι καθοριστικός στη διαδικασία συσσώρευσης πρωτεϊνών και κατά συνέπεια στην αύξηση του οξειδωτικού στρες. Το PINK1 μπορεί να φωσφορυλιώσει μιτοχονδριακές πρωτεΐνες σε απόκριση στο οξειδωτικό στρες αποτρέποντας έτσι τη μιτοχονδριακή δυσλειτουργία η οποία έχει κεντρικό ρόλο στη παθογένεια της νόσου του Πάρκινσον, άρα και η μείωση της δράσης της μπορεί να οδηγήσει σε αύξηση του οξειδωτικού στρες (Valente et al., 2004).
- DJ-1 (PARK7):** Το τρίτο γονίδιο που σχετίζεται σπάνια με ποσοστό 1% εν συγκρίσει με τις μεταλλάξεις του Parkin (PARK2) με αυτοσωμικό υπολειπόμενο παρκινσονισμό πρόωμης έναρξης αφορά το γονίδιο DJ-1 (PARK7) (Hedrich et al., 2004). Η ηλικία έναρξης κυμαίνεται από 27 έως 40 ετών με τους ασθενείς να εμφανίζουν αρκετά συχνά δυστονία, ήπια δυσκαμψία, θετική ανταπόκριση στη θεραπεία με λεβοντόπα ακόμη και αν προκαλούνται ορισμένες δυσκινησίες μετά τη χορήγησή της ενώ εμφανής ήταν και οι ψυχιατρικές εκδηλώσεις σε ορισμένους ασθενείς (Dekker et al., 2003). Η DJ-1 σχηματίζει μια διμερή δομή που αποτελείται από 189 αμινοξέα και εντοπίζεται τόσο στο κυτταρόπλασμα όσο και στον πυρήνα και τα μιτοχόνδρια. Ο ρόλος της πρωτεΐνης DJ-1 είναι πολλαπλός και καθοριστικός στη νόσο του Πάρκινσον αφού συμβάλλει στην προστασία των ντοπαμινεργικών νευρώνων από τον νευροεκφυλισμό και ο άγριος τύπος DJ-1 λειτουργεί ως αντιοξειδωτικό. Επιπλέον, εμφανίζει τις ιδιότητες ως μοριακή σύνοδος πρωτεΐνη (chaperone) και ως μεταγραφικός ρυθμιστής που προστατεύει τα μιτοχόνδρια από το οξειδωτικό στρες, Υπερ-οξείδωση της DJ-1 παρατηρήθηκε σε σποραδικές περιπτώσεις της νόσου του Πάρκινσον και της νόσου του Alzheimer καθιστώντας την πρωτεΐνη ανενεργή, υποδεικνύοντας έτσι την εμπλοκή της

στην παθογένεια της νόσου του Πάρκινσον (Ariga et al., 2013) (Dolgacheva et al., 2019).

- **ATP13A2 (PARK9):** Τέλος, μεταλλάξεις του γονιδίου της ATP13A2 (PARK9), έχει βρεθεί ότι προκαλούν τη νόσο του Πάρκινσον με αυτοσωμικό υπολειπόμενο τύπο κληρονομικότητας η οποία καλείται Kufor-Rakeb Syndrome (KRS) (Najim al-Din et al., 1994). Ορισμένα χαρακτηριστικά του συνδρόμου είναι η νεανική ηλικία έναρξης της νόσου (11-16 ετών), η θετική ανταπόκριση στη θεραπεία με λεβοντόπα, η εμφάνιση ταχείας προόδου της νόσου, η άνοια, η υπερπυρηνική παράλυση και η πυραμιδική εκφύλιση (Di Fonzo et al., 2007). Η έκφραση της εντοπίζεται κυρίως σε εγκεφαλικούς ιστούς και πιο συγκεκριμένα στους ντοπαμινεργικούς νευρώνες της μέλαινας ουσίας και η αγρίου τύπου ATP13A2 εντοπίζεται στη λυσοσωμική μεμβράνη. Κωδικοποιεί ένα μέλος της υπερ οικογένειας ATPases τύπου P (P-type ATPases) και οι μηχανισμοί που προκαλούν τη φαινοτυπική μεταβλητότητα παραμένουν ακόμη ασαφής. Όμως οι διάφορες επιδράσεις των μεταλλάξεων σε βαθμό που επηρεάζουν τη λειτουργία της ATP13A2 μπορούν να διαδραματίσουν καθοριστικό ρόλο. Έτσι, η κατανόηση της λειτουργίας της ATP13A2 στην ομοιόσταση  $Zn^{2+}$ , στη μιτοχονδριακή λειτουργία και στη λυσοσωμική λειτουργία παρέχουν τη δυνατότητα για περαιτέρω κατανόηση του ρόλου της ATP13A2 στο Πάρκινσον (Park et al., 2015).

### 1.2.3.3 Προδιαθεσικοί παράγοντες κινδύνου για τη νόσο του Πάρκινσον

Μολονότι οι μονογονιδιακές μορφές συμβάλλουν σημαντικά στη νόσο του Πάρκινσον όπως αναλύθηκαν λεπτομερώς παραπάνω, υπάρχουν και άλλα γονίδια που πιθανώς να σχετίζονται με τη νόσο του Πάρκινσον χωρίς όμως να έχει επιβεβαιωθεί πλήρως η συσχέτιση τους, αλλά να επηρεάζουν τον κίνδυνο εμφάνισης της νόσου. Ορισμένα από τα γονίδια αυτά είναι: *UCHL1 (PARK5)*, *GYGYF2 (PARK11)*, *HtrA2/Omi (PARK13)*, *PLA2G6 (PARK14)* και *Fbxo7 (PARK15)*. Μεταλλάξεις στα γονίδια: *synphilin-1*, *NR4A2*, *POLG*, *HSPA9* και *PARL* θεωρήθηκαν ωστόσο νοσογόνα με ελάχιστη όμως συμβολή στην παθογένεια της νόσου του Πάρκινσον (Klein and Westenberger, 2012). Μερικά άλλα γονίδια όπως αυτά είναι: *MAPT*, *HLA-DRA*, *GAK-DGKQ*, *PARK16*, *BST1*, *GBA* (Πίνακας 2), έχουν συσχετιστεί με αυξημένο κίνδυνο εμφάνισης του Πάρκινσον και εντοπίστηκαν κυρίως με βάση των Genome-Wide Association Studies (GWAS). Αξίζει επιπλέον να αναφερθεί ότι μεταλλάξεις του γονιδίου *VPS35* που σχετίζεται με την αυτοφαγία, προκαλούν με αυτοσωμικό επικρατές τύπο κληρονομικότητας παρκινσονισμό με καθυστερημένη έναρξη (Trinh and Farrer, 2013). Η β-γλυκοκερεβροσιδάση το οποίο είναι ένα λυσοσωμικό ένζυμο κωδικοποιείται από το γονίδιο *GBA* με σημαντικό ρόλο στον μεταβολισμό των γλυκολιπιδίων. Διαταραχές στη λειτουργία της β-γλυκοκερεβροσιδάση προκαλούν την συσσώρευση της στο δικτυοενδοθηλιακό σύστημα των κυττάρων του αίματος, του ήπατος, του μυελού των οστών, του σπλήνα, των πνευμόνων και του νευρικού συστήματος το οποίο είναι γνωστό και ως νόσος Gaucher (GD) (Hruska et al., 2008). Η νόσος Gaucher είναι μια αυτοσωμική υπολειπόμενη νόσος η οποία χαρακτηρίζεται

από την ανεπάρκεια της β-γλυκοκερεβροσιδάσης. Οι μεταλλάξεις του γονιδίου GBA σχετίζονται με τις μεταβολές που μπορεί να υποστούν τα επίπεδα των λιπιδίων τα οποία κατά συνέπεια οδηγούν σε νόσους λυσοσωμικής αποθήκευσης (Lysosomal Storage Diseases (LSDs))τα οποία μπορεί να προκαλέσουν αυτοφαγική-λυσοσωμική δυσλειτουργία και συνουκλειϊνοπάθειες (Velayati et al., 2010).

Gene	Chromosome	Population	References
SNCA	4q21	USA, UK, France, Japan	34,50,51,58,105–107,131,160,161
MAPT	17q21.1	USA, UK, France	34,50,51,58,105–107,131,161,162
LRRK2	12q12	USA, Japan	50,51
HLA-DRA	6q21.3	USA, UK	51,58,105,106
GAK-DGKQ	4p16	USA, UK	34,51,58,105–107
PARK16	1q32	USA, UK, Japan	50,107,131
BST1	4p15	France, USA	50,107,161

Πίνακας 2 Genome-Wide Association Studies (Trinh and Farrer, 2013)

### 1.2.3.4 Παθογενετικοί μηχανισμοί στη νόσο του Πάρκινσον

#### 1.2.3.4.1 Δυσλειτουργία των κυτταρικών μηχανισμών πρωτεϊνικής αποικοδόμησης

Η τοξική συσσώρευση των λανθασμένων και αναδιπλωμένων πρωτεϊνών και πιο συγκεκριμένα της α-συνουκλειϊνης εντός των ντοπαμινεργικών κυττάρων που αυτά οδηγούν σε νευρικό θάνατο ήταν το έναυσμα για την περαιτέρω διερεύνηση των ρόλων του συστήματος ουβικουιτίνης-πρωτεασώματος (Ubiquitin-Proteasome System-UPS) και της αυτοφαγίας-λυσοσώματος (Autophagy lysosomal pathway-ALP) στην παθογένεια της νόσου του Πάρκινσον. Το σύστημα ουβικουιτίνης-πρωτεασώματος το οποίο έχει κρίσιμο ρόλο στην κυτταρική ρύθμιση, είναι ένα πρωτεολυτικό σύστημα εξαρτώμενο από ATP το οποίο προκαλεί επιλεκτική αποικοδόμηση από το πρωτεάσωμα. Το 26S πρωτεάσωμα είναι το βασικό ενζυμικό σύμπλοκο που υπάρχει στον πυρήνα και τον κυτταρόπλασμα των ευκαρυωτικών κυττάρων και έχει σημαντικό ρόλο στο σύστημα ουβικουιτίνης-πρωτεασώματος. Το σύστημα ουβικουιτίνης-πρωτεασώματος είναι υπεύθυνο για την αποτελεσματική απομάκρυνση των αναδιπλωμένων ή κατεστραμμένων πρωτεϊνών και των ενδοκυτταρικών πρωτεϊνών με μικρή διάρκεια ζωής που δημιουργούνται σε διάφορα κυτταρικά διαμερίσματα συμπεριλαμβανομένου του κυτταροπλάσματος, του πυρήνα και του ενδοπλασματικού δικτύου. Η διαδικασία αυτή επιτυγχάνεται μέσω μιας σειράς πολλαπλών σταδίων και την ενεργοποίηση πολλών και διαφόρων ενζύμων (McNaught et al., 2001). Η δυσλειτουργία του UPS και η συνακόλουθη συσσώρευση των λανθασμένων πρωτεϊνών εμπλέκονται στη παθογένεια της νόσου του Πάρκινσον όπως αυτό υποστηρίζεται από τις γενετικές μελέτες των γονιδίων που εμπλέκονται στην παθογένεια της σποραδικής μορφής του Πάρκινσον. Πράγματι μεταλλάξεις των

γονιδίων που κωδικοποιούν την  $\alpha$ -συνουκλεΐνη, την Parkin και την UCH-L1 (μέλος της οικογενείας των υδρολασών ουβικουιτίνης C-άκρου (UCHs)) επηρεάζουν σε μεγάλο βαθμό τη λειτουργία του μηχανισμού της ουβικουιτίνης-πρωτεασώματος. Ιδιαίτερα μεταλλάξεις στο γονίδιο της Parkin στον αυτοσωμικό υπολειπόμενο νεανικό παρκινσονισμό έχουν αποδείξει τη μείωση της ενζυματικής δραστηριότητας της E3 λιγάσης ουβικουιτίνης στη μέλαινα ουσία γεγονός που υποστηρίζει επίσης την υπόθεση ότι δυσλειτουργία του συστήματος ουβικουιτίνης-πρωτεασώματος οδηγεί σε νευροεκφυλισμό (Shimura et al., 2000).

Ένα άλλο σημαντικό μονοπάτι στην αποικοδόμηση των πρωτεϊνών στα ευκαρυωτικά κύτταρα είναι το μονοπάτι αυτοφαγίας-λυσosώματος. Διαχωρίζεται σε τρία διαφορετικά μονοπάτια βάσει του τρόπου με τον οποίο τα υποστρώματα φτάνουν στο λυσοσωμικό αυλό: μακροαυτοφαγία (χαρακτηρίζεται ως αυτοφαγία), μικροαυτοφαγία και συνοδευτική αυτοφαγία (Chaperone-mediated autophagy-CMA)(Cuervo et al., 2004; Levine and Klionsky, 2004). Η αυτοφαγία είναι ο πρωταρχικός μηχανισμός που εμπλέκεται στην αποικοδόμηση μακροχρόνιων, σταθερών πρωτεϊνών και ο μοναδικός στην ανακύκλωση ολόκληρων οργανιδίων όπως τα μιτοχόνδρια. Μεγάλα πρωτεϊνικά υπολείμματα όπως τα ολιγομερή και τα ινίδια εξαιτίας του μεγέθους τους αποικοδομούνται από αυτοφαγία η οποία είναι επίσης μια διαδικασία πολλών σταδίων που περιλαμβάνει το σχηματισμό των αυτοφαγοσωμάτων. Η  $\alpha$ -συνουκλεΐνη μπορεί να αποδομηθεί τόσο από το σύστημα UPS όσο και από αυτοφαγία, γεγονός που υποδηλώνει ότι η διαταραχή της αυτοφαγικής αποδόμησης συμβάλλει σημαντικά στο νευροεκφυλισμό στη νόσου του Πάρκινσον (Webb et al., 2003). Υπάρχουν πολλά σημεία στα οποία μπορεί να προκληθεί δυσλειτουργία του ALP, τόσο από την αποτυχία σχηματισμού των αυτοφαγοσωμάτων μέχρι και την έλλειψη ενζύμων σε λυσοσώματα που με τη σειρά τους μπορεί να προκαλέσουν τη συσσώρευση ανεπιθύμητων πρωτεϊνών και να οδηγήσουν σε κυτταρικό θάνατο. Τα ευρήματα ότι η λυσοσωμική δυσλειτουργία και οι μεταλλάξεις στο *ATP3A2* οδηγούν σε συσσωρευμένη  $\alpha$ -συνουκλεΐνη εξαιτίας της αποτυχίας εκτέλεσης της αυτοφαγίας αποτελούν επιπλέον χαρακτηριστικά που υποστηρίζουν την υπόθεση ότι και η δυσλειτουργία του ALP είναι ένας σημαντικός μηχανισμός νευροεκφυλισμού (Pan et al., 2008). Τέλος, η γήρανση αποτελεί έναν σημαντικό παράγοντα που σχετίζεται με τη δυσλειτουργία των συστημάτων ελέγχου των πρωτεϊνών, καθώς όσο αυξάνεται η ηλικία η οποία είναι ένας από τους βασικούς παράγοντες κινδύνου στη νόσο του Πάρκινσον, τόσο πιο ευάλωτος είναι ο εγκέφαλος στις δυσλειτουργίες αυτές (Rajawat et al., 2009).

#### 1.2.3.4.2 Οξειδωτικό στρες και μιτοχονδριακή δυσλειτουργία

Το οξειδωτικό στρες (Oxidative stress-OS) το οποίο αντιπροσωπεύει την ανισορροπία μεταξύ της παραγωγής και της αποτοξίνωσης αντιδραστικών ειδών οξυγόνου και αζώτου, παίζει σημαντικό ρόλο στη γήρανση του εγκεφάλου, στις νευροεκφυλιστικές ασθένειες καθώς και σε άλλες ασθένειες όπως είναι η ισχαιμία. Υπό φυσιολογικές συνθήκες τα μόρια αυτά χρησιμεύουν ως μόρια σηματοδότησης, ενώ μια υπερβολική ποσότητα αυτών των ελεύθερων ριζών συνεπώς προκαλεί δυσλειτουργία των πρωτεϊνών, λιπιδίων και το οξειδωτικό στρες. Παραγωγή των ελεύθερων ριζών (Reactive Oxygen Species-ROS) αποτελούν τόσο εξωγενείς όσο και ενδογενείς πηγές

με τα μιτοχόνδρια να αποτελούν τις κύριες ενδογενείς πηγές (Sarsour et al., 2010; Wang and Michaelis, 2010). Η νευροτοξική δράση του MPTP που είναι υποπροϊόν από τη σύνθεση της MPPP αναστέλλει τη δραστηριότητα του μιτοχονδριακού συμπλόκου I όπου ανεπάρκεια αυτού, συνεπάγεται την εκφύλιση των ντοπαμινεργικών νευρώνων γεγονός που υποδηλώνει το ρόλο της μιτοχονδριακής δυσλειτουργίας στη παθογένεια της νόσου του Πάρκινσον. Ενδιαφέρον επίσης αποτελούν και οι διάφορες μεταλλάξεις των: α-συνουκλεΐνη, Parkin, PINK1, DJ-1 που έχουν εντοπιστεί σε μελέτες ασθενών με τη νόσο του Πάρκινσον και έχουν συσχετιστεί με τη μιτοχονδριακή δυσλειτουργία. Η μιτοχονδριακή δυσλειτουργία συνεπάγεται σε διαφοροποιήσεις στη μιτοχονδριακή λειτουργία και δομή, αλλοιωμένο εντοπισμό πρωτεϊνών εντός των μιτοχονδρίων καθώς και μείωση της δραστηριότητας του συμπλόκου I όπως αναφέρθηκε νωρίτερα. Οι μεταλλάξεις πρωτεϊνών, οι διάφορες αλλαγές στη συσσώρευση της α-συνουκλεΐνης και η αλλοιωμένη πρωτεόλυση είναι παραδείγματα πρόκλησης του οξειδωτικού στρες στη νόσο του Πάρκινσον χωρίς όμως να υπάρχει συγκεκριμένος ρόλος του οξειδωτικού στρες στη νόσο αυτή. Τέλος η απώλεια λειτουργίας των DJ-1, PINK1, Parkin συμβάλλει στη μείωση της μιτοχονδριακής προστασίας έναντι του οξειδωτικού στρες το οποίο με τη σειρά του συνδράμει στην αύξηση της μιτοχονδριακής δυσλειτουργίας (Dexter and Jenner, 2013).

#### 1.2.3.4.3 Ο ρόλος της διεγερτικής τοξικότητας στη νόσο του Πάρκινσον

Οι νευροεκφυλιστικές ασθένειες χαρακτηρίζονται από την προοδευτική δυσλειτουργία και την απώλεια των νευρώνων που προκαλούνται από διάφορα νευρολογικά ελλείμματα. Η διεγερτική τοξικότητα συμβάλλει παθολογικά στον κυτταρικό θάνατο στη νόσο του Πάρκινσον με κύριο συντελεστή την υπερδιέγερση των υποθαλαμικών πυρήνων (Subthalamic nuclei-STN) οι οποίοι παράγουν γλουταμινικό. Η ντοπαμίνη ως ένας ανασταλτικός νευροδιαβιβαστής κρατά σε σταθερά επίπεδα την κατάσταση των υποθαλαμικών πυρήνων. Εξαιτίας της απώλειας των ντοπαμινεργικών νευρώνων στη νόσο του Πάρκινσον το STN ενεργοποιείται υπερβολικά οδηγώντας έτσι σε υπερβολική παραγωγή του γλουταμινικού νευροδιαβιβαστή. Το υπερβολικό γλουταμινικό συνδέεται με τους ιονοτροπικούς υποδοχείς (NMDA, AMPA) και ανοίγει τα κανάλια ασβεστίου ( $Ca^{++}$ ) τα οποία προκαλούν διεγερτική τοξικότητα. Η διεγερτική τοξικότητα μπορεί επίσης να προκληθεί από δυσλειτουργία των μιτοχονδρίων που εμφανίζεται στη νόσο του Πάρκινσον. Επιπρόσθετα, η υπερβολική ενεργοποίηση των υποδοχέων γλουταμινικού βλάπτει την κυτταρική ομοιόσταση ασβεστίου και ενεργοποιεί τη δημιουργία ελεύθερων ριζών οδηγώντας σε οξειδωτικό στρες καθώς και τον προγραμματισμένο κυτταρικό θάνατο. Τέλος, η αυξημένη παραγωγή γλουταμινικού η οποία οδηγεί σε διεγερτική τοξικότητα ασβεστίου  $Ca^{++}$  που καθιστά τους ντοπαμινεργικούς νευρώνες ευάλωτους σε νευροεκφυλισμό μπορεί να προκληθεί και από την περιβαλλοντική έκθεση τοξινών (Bjorling-Poulsen et al., 2008; Dexter and Jenner, 2013; Dong et al., 2009).

#### 1.2.3.4.4 Ο ρόλος της νευροφλεγμονής στη νόσο του Πάρκινσον

Οι διάφορες τοξίνες όπως είναι το 6-OHDA, το MPTP και η ροτενόνη σε συνδυασμό με τις υψηλές συγκεντρώσεις κυτοκινών όπως TNF- $\alpha$ , IL-1 $\alpha$  και IL-1 $\beta$  έχουν βρεθεί σε ασθενείς με τη νόσο του Πάρκινσον οι οποίες προκαλούν την άμεση ενεργοποίηση



της μικρογλοίας του εγκεφάλου η οποία παίζει βασικό ρόλο στη βλάβη των ντοπαμινεργικών νευρών της μέλαινας ουσίας και συμβάλουν και στην πρόκληση της νευροφλεγμονής (Dexter and Jenner, 2013). Ωστόσο υπάρχουν ευρήματα που υποστηρίζουν την άμεση σχέση του εντερικού μικροβιώματος με τη νόσο του Πάρκινσον. Η νόσος μπορεί να ξεκινήσει από μια φλεγμονή που μπορεί να προκληθεί αρχικά στο εντερικό μικροβίωμα (φυσική μικροχλωρίδα των βακτηρίων), μέσω της συστηματικής απελευθέρωσης ενδοτοξινών όπως λιποπολυσακχαρίτη (LPS) και άλλων μεταβολικών προϊόντων, τα οποία μπορούν να διαπεράσουν μέσω των εντερικών κυττάρων στο αίμα και τελικώς να επηρεάσουν το κεντρικό νευρικό σύστημα και να προκαλέσουν φλεγμονή στον εγκέφαλο (Baizabal-Carvallo and Alonso-Juarez, 2020).

### **1.2.3.5 Διάγνωση της νόσου**

Η νόσος του Πάρκινσον είναι μια νευροεκφυλιστική διαταραχή η οποία εκδηλώνεται με μια σειρά χαρακτηριστικών κινητικών και μη συμπτωμάτων. Ορισμένα από τα συμπτώματα αυτά εμφανίζονται κατά τη διάρκεια της φυσιολογικής γήρανσης και έτσι αυξάνονται οι πιθανότητες μιας λανθασμένης διάγνωσης. Δεν υπάρχει μοναδικός τρόπος με τον οποίο μπορεί να γίνει σωστά η διάγνωση αλλά η ύπαρξη διαφόρων συμπτωμάτων σε συνδυασμό με τα διαγνωστικά τεστ μπορούν να βοηθήσουν στη διάγνωση αυτής. Οι ασθενείς διαγιγνώσκονται με βάση το ιστορικό και ορισμένες νευρολογικές εξετάσεις. Όσον αφορά το ιστορικό θα πρέπει να πραγματοποιηθεί διεξοδική ανάκριση του ασθενούς αλλά και της οικογένειας προκειμένου να προσδιοριστεί επακριβώς ποια είναι η σειρά με την οποία εμφανίστηκαν τα συμπτώματα, ένα υπάρχει ιστορικό λήψης φαρμάκων τόσο στο παρελθόν όσο και στο παρόν τα οποία μπορούν να προκαλέσουν συμπτώματα παρκινσονισμού καθώς και εάν υπάρχει πιθανή έκθεση σε τοξικολογικές ουσίες (Massano and Bhatia, 2012).

Προκειμένου να γίνει η διάγνωση για τη νόσο του Πάρκινσον θα πρέπει ο ασθενής να έχει εμφανίσει για κάποια χρονική περίοδο δύο από τα παρακάτω συμπτώματα όπως αυτά είναι: Κινητικό τρόμο, βραδυκινησία, μυϊκή δυσκαμψία και ορθοστατική αστάθεια (Janjovic, 2008). Η διάγνωση απαιτεί την παρουσία δύο νευροπαθολογικών καταστάσεων: 1) την εκφύλιση των ντοπαμινεργικών νευρώνων της μέλαινας ουσίας και 2) την ύπαρξη των σωματίων Lewy. Η ύπαρξη των σωματίων Lewy γίνεται με μικροσκοπική εξέταση σε εγκεφάλους ασθενών που έχουν ήδη αποβιώσει. Όμως από μόνη της δεν αποτελεί αξιόπιστο και μοναδικό δείκτη με σημαντική διαγνωστική αξία καθώς η ύπαρξη των σωματίων Lewy έχει βρεθεί και σε άλλες ασθένειες όπως είναι το Alzheimer, η άνοια με σωματία Lewy όπου είναι μια ευρύτερη κατηγορία διαταραχών που χαρακτηρίζονται ως συνουκλείνοπάθειες (Corti et al., 2011). Αναφορικά με την κλινική εξέταση οι γιατροί μπορούν να πραγματοποιήσουν μια σειρά «τυπικών δοκιμών» προκειμένου να επιβεβαιώσουν τα παραπάνω συμπτώματα τα οποία θα συμβάλλουν έτσι στη διάγνωση της νόσου. Για παράδειγμα η βραδυκινησία μπορεί να αξιολογηθεί ζητώντας από τον ασθενή να πραγματοποιήσει όσο το δυνατόν ταχύτερα ορισμένες επαναλαμβανόμενες κινήσεις όπως το ανοιγοκλείσιμο του χεριού, το άγγιγμα του αντίχειρα και των δεικτών, να ακουμπήσει το πόδι του στο έδαφος και γενικότερα να ελέγξει την συνολική εικόνα του ασθενούς παρατηρώντας τις διάφορες

αυθόρμητες κινήσεις του όταν στέκεται όρθιος ή περπατάει αλλά ακόμη και τις διάφορες εκφράσεις του προσώπου και των ματιών του. Ο τρόμος μπορεί να αξιολογηθεί με τη προσεκτική παρατήρηση των κινήσεων και της συνολικής εικόνας του ασθενούς, ενώ από την άλλη η μυϊκή δυσκαμψία γίνεται αισθητή από την εξέταση της κίνησης των άκρων ή του λαιμού του ασθενούς. Τέλος, η ορθοστατική αστάθεια αξιολογείται με το αρχικό «τράβηγμα» που εφαρμόζεται στους ώμους του ασθενούς όπου στη συνέχεια θα πρέπει ο ίδιος να ανακτήσει την ισορροπία του χωρίς να υπάρχει κάποια πτώση (Massano and Bhatia, 2012). Υπάρχουν πολλές κλίμακες αξιολόγησης που χρησιμοποιούνται συχνά για να περιγράψουν την εξέλιξη των συμπτωμάτων της νόσου του Πάρκινσον σε διάφορα στάδια δίνοντας τη δυνατότητα μέτρησης της σοβαρότητας της νόσου, τη διάγνωση και τελικά τον τρόπο διαχείρισης των ασθενών. Η κλίμακα **Hoehn & Yahr** είναι το πιο συχνό σύστημα το οποίο χρησιμοποιείται εκτενώς για την περιγραφή μεγάλων πληθυσμών ασθενών με τη νόσο του Πάρκινσον η οποία βασίζεται στις κινητικές διαταραχές των ασθενών. Η αρχική κλίμακα περιλαμβάνει συνολικά 5 στάδια αριθμημένα από το 1 έως το 5, ενώ στη συνέχεια προστέθηκε η τροποποιημένη κλίμακα με επιπλέον ενδιάμεσα στάδια (**Εικόνα 3**) (Goetz et al., 2004).

Hoehn and Yahr scale	Modified Hoehn and Yahr scale
1: Unilateral involvement only usually with minimal or no functional disability	1.0: Unilateral involvement only
2: Bilateral or midline involvement without impairment of balance	1.5: Unilateral and axial involvement
3: Bilateral disease: mild to moderate disability with impaired postural reflexes; physically independent*	2.0: Bilateral involvement without impairment of balance
4: Severely disabling disease; still able to walk or stand unassisted	2.5: Mild bilateral disease with recovery on pull test
5: Confinement to bed or wheelchair unless aided	3.0: Mild to moderate bilateral disease; some postural instability; physically independent
	4.0: Severe disability; still able to walk or stand unassisted
	5.0: Wheelchair bound or bedridden unless aided

Εικόνα 3 Τροποποιημένη κλίμακα Hoehn & Yahr

### 1.2.3.6 Θεραπεία της νόσου

Η θεραπεία των ασθενών με Πάρκινσον είναι συντηρητική βάση της κλινικής εικόνας των ασθενών, καθώς δεν υπάρχουν διαθέσιμα αποτελεσματικά και αποκλειστικά θεραπευτικά μέσα. Όμως η έγκαιρη διάγνωση και η κατάλληλη στήριξη των ατόμων μπορούν να προσφέρουν μια πιο παραγωγική και μεγαλύτερης διάρκειας ζωή. Οι θεραπείες που εφαρμόζονται περιλαμβάνουν κυρίως φαρμακευτική αγωγή και χειρουργική θεραπεία. Οι θεραπείες αυτές επιβραδύνουν κυρίως την εξέλιξη της νόσου χωρίς όμως κάποιο από αυτά να αντιστρέφει τις επιπτώσεις της νόσου. Μερικές ωστόσο βοηθητικές θεραπείες όπως αλλαγές του τρόπου ζωής, ο ρόλο της διατροφής τόσο στη νευροπροστασία όσο και στον νευροεκφυλισμό, η ξεκούραση και άσκηση συντείνουν στη βελτίωση της κατάστασης τους. Οι φαρμακολογικές θεραπείες διαχωρίζονται σε δυο βασικές κατηγορίες, στα φάρμακα με ντοπαμινεργική δράση και στα φάρμακα με αντιχολινεργική δράση. Στην πρώτη κατηγορία ανήκουν: η λεβοντόπα, οι αγωνιστές υποδοχέων ντοπαμίνης, οι αναστολείς μονοαμινοξειδάσης B (MAO-B) και οι αναστολείς COMT (**Πίνακας 3**) (Armstrong and Okun, 2020b).

Πιο συγκεκριμένα:

- **Η λεβοντόπα (L-3,4-δihυδροξυφαινυλαλανίνη- Levodopa (LD))**

Η λεβοντόπα είναι ο πρόδρομος της ντοπαμίνης και χρησιμοποιείται ως παράγοντας αντικατάστασης των επιπέδων της ντοπαμίνης για τη θεραπεία των πασχόντων από τη νόσο του Πάρκινσον. Συνταγογραφείται σε ασθενείς όταν τα συμπτώματα τους έχουν επιδεινωθεί και είναι δύσκολο να ελεγχθούν με τη χορήγηση άλλων θεραπειών όπως οι αναστολείς της μονοαμινοξειδάσης (MAO-B) και οι αγωνιστές υποδοχέων ντοπαμίνης. Χρησιμοποιείται αποτελεσματικότερα για τον έλεγχο των βραδυκινητικών συμπτωμάτων που είναι εμφανή στη νόσο και συμβάλλει επίσης αποτελεσματικά στη βελτίωση της ποιότητας ζωής των νοσούντων. Σε αντίθεση με την ντοπαμίνη, η λεβοντόπα μπορεί να διασχίσει τον αιματοεγκεφαλικό φραγμό η οποία χορηγείται σε συνδυασμό με αναστολείς αποκαρβοξυλάσης (όπως καρβιντόπα και βενσεραζίδη) προκειμένου να μειωθούν κατά το δυνατόν τρόπο οι παρενέργειες της. Η λεβοντόπα μετατρέπεται σε ντοπαμίνη τόσο στο κεντρικό νευρικό σύστημα όσο και στην περιφέρεια. Έτσι οι αναστολείς αποκαρβοξυλάσης παρεμποδίζουν την μετατροπή της λεβοντόπα σε ντοπαμίνη στην περιφέρεια αποτέλεσμα αυτό να επιτραπεί περισσότερη λεβοντόπα να διασχίσει το αιματοεγκεφαλικό φραγμό. Όπως όλες οι φαρμακευτικές ουσίες έτσι και η χρήση της λεβοντόπα παρουσιάζουν συχνές ανεπιθύμητες ενέργειες στους ασθενείς όπως είναι: η ναυτία, η ζάλη, η υπνηλία και η κεφαλαλγία. Οι παρενέργειες αυτές μπορούν να ελαχιστοποιηθούν με χαμηλή δόση έναρξης κατά την έναρξη της θεραπείας και σταδιακής αύξησης όταν αυτό κριθεί αναγκαίο. Ιδιαίτερη προσοχή θα πρέπει να αποδοθεί στους ηλικιωμένους ασθενείς που λαμβάνουν λεβοντόπα, καθώς παρουσιάζουν ευαισθησία στις επιδράσεις του κεντρικού νευρικού συστήματος και συχνά παρουσιάζουν ορισμένες ανεπιθύμητες ενέργειες όπως ψευδαισθήσεις, ψύχωση, και σύγχυση. Η αύξηση της ποσότητας της καρβιντόπα και ιδιαίτερα η δομπεριδόνη χρησιμοποιούνται κυρίως για την ανακούφιση της ναυτίας με τη δεύτερη στην περίπτωση που η καρβιντόπα είναι αναποτελεσματική (Gandhi and Saadabadi, 2022).

- **Αγωνιστές υποδοχέων ντοπαμίνης (Dopamine agonists)**

Οι αγωνιστές υποδοχέων ντοπαμίνης μπορούν να είναι πιο αποτελεσματικοί κατά τα πρώτα στάδια της νόσου του Πάρκινσον διότι μπορούν να αυξήσουν τα επίπεδα της ντοπαμίνης στον εγκέφαλο καθώς επίσης μπορούν να χορηγηθούν και σε ασθενείς που επιθυμούν λιγότερη συχνή δοσολογία φαρμάκων. Οι πιο σύνηθες αγωνιστές είναι ηπραμιπεξόλη και η ροπινιρόλη. Η χρήση τους η οποία γίνεται στα αρχικά στάδια της νόσου μπορούν να καθυστερήσουν την ανάγκη για χρήση της λεβοντόπα και να χρησιμοποιηθούν για τον έλεγχο της ακαμψίας και της δυσκινησίας που προκαλούνται από τη χρήση της λεβοντόπα. Ο ρόλος τους είναι κυρίως προσθετικός στη θεραπεία με τη λεβοντόπα και ορισμένες από τις παρενέργειες που προκαλούνται είναι

περιφερικό οίδημα, διαταραχές ελέγχου των παλμών, δερματικούς ερεθισμούς, ψύχωση, υπνηλία και ένα ενοχλητικό σύνδρομο στέρησης. Τέλος, ένας άλλος αγωνιστής της ντοπαμίνης είναι η απομορφίνη η οποία χρησιμοποιείται για τη θεραπεία στις διακυμάνσεις κινητικότητας (on/off) των ασθενών που πάσχουν από τη νόσο του Πάρκινσον (Reich and Savitt, 2019).

- **Αναστολείς μονοαμινοξειδάσης Β (Monoamine oxidase B inhibitors- MAO-B)**

Η μείωση των επιπέδων της ντοπαμίνης στη νόσο του Πάρκινσον μπορεί να οφείλεται στη γρήγορη διάσπαση του από το καταλυτικό ένζυμο μονοαμινοξειδάση-Β (MAO-B). Έτσι λοιπόν η αναστολή του MAO-B αποτελεί σημαντική διαδικασία η οποία συμβάλει στη διατήρηση των επιπέδων της ντοπαμίνης στον εγκέφαλο η οποία μπορεί να βελτιώσει πολλά κινητικά συμπτώματα της νόσου. Οι πιο συχνά χρησιμοποιούμενοι αναστολείς MAO-B είναι η σελεργιλίνη, η ρασαγιλίνη και η σαφιναμίδη οι οποίες όταν χορηγηθούν σε συνδυασμό με τη λεβοντόπα, μπορούν να αυξήσουν την απόκριση στη λεβοντόπα για μεγαλύτερο διάστημα. Οι αναστολείς αυτοί μπορούν να χρησιμοποιηθούν είτε ως μονοθεραπεία είτε συνδυαστικά με άλλα φάρμακα και να συμβάλλουν για τη θεραπεία στις διακυμάνσεις κινητικότητας των ασθενών με τη νόσο του Πάρκινσον (Reich and Savitt, 2019).

- **Αναστολείς COMT (Κατεχολ-ο-μεθυλοτρασφεράση- Catechol-O-methyl transferase inhibitors)**

Η Κατεχολ-ο-μεθυλοτρασφεράση ευθύνεται για την έμμεση διάσπαση της ντοπαμίνης. Επομένως οι αναστολείς COMT αναστέλλουν το μεταβολισμό τόσο της λεβοντόπα όσο και της ντοπαμίνης με αποτέλεσμα αυτές να παραμένουν δραστικές για περισσότερο διάστημα. Οι πιο συνηθισμένοι αναστολείς COMT είναι η εντακαπόνη και η τολκαπόνη με τη δεύτερη να προκαλεί ηπατική τοξικότητα και να απαιτεί συχνή παρακολούθηση. Η εντακαπόνη δρα σε συνδυασμό με την καρβιντόπα και τη λεβοντόπα και προκαλεί αποχρωματισμό των ούρων αλλά ενδέχεται ακόμη να προκαλέσει και διάρροια. Και οι δυο αυτές φαρμακευτικές ουσίες σχετίζονται με σημαντική κλινική βελτίωση και μπορούν να μειώσουν την ευαισθησία της λεβοντόπα σε ασθενείς με Πάρκινσον (Antonini et al., 2008).

- **Αντιχολινεργικά φάρμακα**

Τα αντιχολινεργικά φάρμακα συμπεριλαμβανομένου της τριεξυφαινιδύλη χρησιμοποιούνται ως επί το πλείστον για τη θεραπεία του τρόμου και της δυσκαμψίας ασθενών με τη νόσο του Πάρκινσον με αρκετές όμως παρενέργειες όπως αυτές είναι ξηροστομία, δυσκοιλιότητα, κατακράτηση ούρων, ταχυκαρδία, θολή όραση και οι διάφορες διαταραχές στο γαστρεντερικό σύστημα. Σε μεγάλης ηλικίας ασθενών έχει παρατηρηθεί μάλιστα η σύγχυση, ως παρενέργεια των αντιχολινεργικών φαρμάκων. Μολονότι οι παρενέργειες της φαρμακευτικής αυτής ουσίας είναι αρκετές η τριεξυφαινιδύλη

χρησιμοποιείται σε ασθενείς που εμφανίζουν κινητικό τρέμουλο το οποίο είναι ανθεκτικό στη λεβοντόπα και μπορεί να χρησιμοποιηθεί είτε ως μονοθεραπεία είτε συνδυαστικά με άλλους παράγοντες (Reich and Savitt, 2019).

- **Άλλα φάρμακα**

Τα κινητικά συμπτώματα των ασθενών με τη νόσο του Πάρκινσον μπορούν να περιοριστούν με τη χορήγηση των φαρμάκων που αναφέρθηκαν εκτενώς παραπάνω. Ωστόσο, υπάρχουν και τα μη κινητικά συμπτώματα όπως το άγχος και η κατάθλιψη τα οποία εμφανίζονται στους πάσχοντες από τη νόσο του Πάρκινσον. Μια φαρμακευτική ουσία που χρησιμοποιείται για τη θεραπεία του άγχους είναι η βενζοδιαζεπίνη η οποία εμφανίζει ορισμένες παρενέργειες (Chen and Marsh, 2014).

Υπάρχουν και άλλες φαρμακευτικές ουσίες που χρησιμοποιούνται σε ασθενείς με Πάρκινσον ανάλογα των συμπτωμάτων τους. Για παράδειγμα για τον έλεγχο της δυσκινησίας χρησιμοποιείται η κλοζαπίνη η οποία μπορεί να προκαλέσει μια σειρά παρενεργειών όπως ακοκκιοκυττάρωση, σιελόρροια και καρδιακή τοξικότητα (Durif et al., 2004). Μια άλλη ουσία είναι το istradefylline το οποίο χρησιμοποιείται συνδυαστικά με τη λεβοντόπα και την καρβιντόπα για τη θεραπεία στις διακυμάνσεις κινητικότητας των ασθενών με Πάρκινσον με την εμφάνιση βέβαια ορισμένων παρενεργειών όπως ναυτία και παραισθήσεις (Armstrong and Okun, 2020b).

- **Χειρουργικές θεραπείες**

Οι χειρουργικές θεραπείες χρησιμοποιούνται σε προχωρημένα στάδια της νόσου όταν έχει γίνει πλέον ανεξέλεγκτη η εμφάνιση των διαφόρων συμπτωμάτων στους ασθενείς και η φαρμακευτική αγωγή είναι πλέον αναποτελεσματική. Γι' αυτό το λόγο στην περίπτωση όπου τα κινητικά συμπτώματα επιβαρύνουν και περιορίζουν σε μεγάλο βαθμό τους ασθενείς τότε λαμβάνονται υπόψη η βαθιά εγκεφαλική διέγερση (deep brain stimulation) και άλλες χειρουργικές προσεγγίσεις. Η βαθιά εγκεφαλική διέγερση στοχεύει συγκεκριμένα μέρη του εγκεφάλου όπως είναι ο υποθαλαμικός πυρήνας και η ωχρά σφαίρα και μπορεί να βοηθήσει τόσο τη μία όσο και τις δύο πλευρές του σώματος. Μια άλλη επιλογή είναι ο εστιασμένος υπέρηχος ο οποίος στοχεύει το τμήμα του εγκεφάλου που καλείται θάλαμος και χρησιμοποιείται μόνο για τον τρόπο και βοηθά μόνο τη μία πλευρά του σώματος. Μια άλλη χειρουργική προσέγγιση είναι η τοποθέτηση ενός σωλήνα στο έντερο με σκοπό τη συνεχή έγχυση φαρμακευτικής αγωγής καθ' όλη τη διάρκεια της ημέρας ώστε ο ασθενής να λαμβάνει σταθερή ποσότητα λεβοντόπα. Τέλος, όλες οι χειρουργικές θεραπείες απαιτούν λεπτομερή έλεγχο από ειδικούς γιατρούς και θα πρέπει να λαμβάνονται υπόψη στην περίπτωση των κινητικών συμπτωμάτων των ασθενών με τη νόσο του Πάρκινσον, καθώς δεν βοηθούν στη θεραπεία των μη κινητικών συμπτωμάτων (Armstrong and Okun, 2020a).

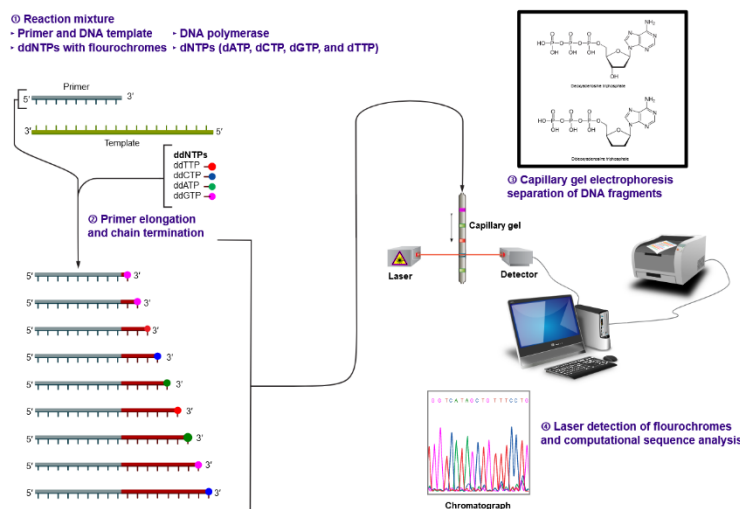
Category	Specific Agents and Typical Starting Dose	Therapeutic Uses				Most Common Adverse Effects Other Than Dyskinesia
		Early Symptomatic	Levodopa Adjunct	Wearing Off	Dyskinesia	
Levodopa preparations	Immediate-release <b>carbidopa-levodopa</b> (25/100 mg, 3 times/d)	●	○	●	●	Nausea
	Controlled-release <b>carbidopa-levodopa</b> (25/100 mg, 3 times/d)	●	○	●	●	Nausea
	Extended-release <b>carbidopa-levodopa</b> (23.75/95 mg, 3 times/d for 3 d; then 36.25/145 mg, 3 times/d for 3 d)	●	○	●	●	Nausea
	Enteral suspension <b>carbidopa-levodopa</b> (clinical titration)		○	●	●	Nausea
	Inhaled <b>levodopa</b> (as needed)		○	● <sup>b</sup>		Nausea, upper respiratory tract infection
Nonergot dopamine agonists <sup>c</sup>	Immediate-release <b>pramipexole</b> (0.125 mg, 3 times/d, increasing weekly) or extended-release <b>pramipexole</b> (0.375 mg, 1 time/d, increasing weekly)	●	●	●	●	Orthostatic hypotension, dizziness, nausea, sleepiness
	Immediate-release <b>ropinirole</b> (0.25 mg, 3 times/d, increasing weekly) or extended-release <b>ropinirole</b> (2 mg, 1 time/d, increasing weekly)	●	●	●	●	Orthostatic hypotension, dizziness, nausea, sleepiness
	Transdermal <b>rotigotine</b> (2 mg/24 h)	●	●	●	●	Site reactions, dizziness, orthostatic hypotension
	Injected <b>apomorphine</b> (as needed)			●		Site reactions, dizziness, orthostatic hypotension
Monoamine oxidase-B inhibitors	<b>Selegiline</b> (5 mg, 2 times/d)	●	●	●	●	Nausea, dizziness, insomnia
	<b>Rasagiline</b> (1 mg every morning)	●	●	●	●	Orthostatic hypotension, nausea
	<b>Safinamide</b> (50 mg/d)	●	●	●	●	Nausea
	<b>Zonisamide</b> (25 to 200 mg/d) <sup>d</sup>			●		Sleepiness, loss of appetite
Catechol-O-methyltransferase inhibitors	<b>Entacapone</b> (200 mg with each levodopa dose)			●	●	Nausea, diarrhea
	<b>Opicapone</b> (50 mg every night) <sup>e</sup>			●	●	Falls, insomnia, orthostatic hypotension
	<b>Tolcapone</b> (100 mg, 3 times/d) <sup>f</sup>			●	●	Gastrointestinal symptoms, orthostatic hypotension, sleep disorders
Other	<b>Anticholinergics</b> (eg, trihexyphenidyl, benzotropine, dose varies) <sup>g</sup>	●	●			Dizziness, anxiety
	<b>Amantadine</b> (dose varies by formulation) <sup>h</sup>	●	●		●	Orthostatic hypotension, hallucinations, edema, gastrointestinal symptoms
	<b>Istradefylline</b> (20 mg/d)			●		Nausea, hallucinations
	<b>Clozapine</b> (12.5–25 mg every night) <sup>f</sup>				●	Sleepiness, dizziness, tachycardia, constipation, orthostatic hypotension, sialorrhea

Πίνακας 3 Συγκεντρωτικός πίνακας φαρμακολογικών ουσιών που χρησιμοποιούνται στη νόσο του Πάρκινσον (Armstrong and Okun, 2020b)

### 1.3 Μέθοδοι αλληλούχησης

Η εφεύρεση των τεχνικών αλληλούχησης του DNA στη δεκαετία του 1970 σηματοδοτεί την αρχή μιας νέα εποχής στον κλάδο της Βιολογίας η οποία οδήγησε σε μια ταχεία επέκταση των γνώσεων αναφορικά με τα γονίδια, την ταυτοποίησή γονιδίων, τον εντοπισμό των διαφόρων μεταλλάξεων και την ανέλιξη δομών γονιδίων και άλλων ρυθμιστικών περιοχών. Το 1977 ο Sanger και οι συνεργάτες του ανέπτυξαν τη μέθοδο αλληλούχησης DNA κατά Sanger η οποία είναι γνωστή και ως μέθοδος τερματισμού αλυσίδας ή μέθοδος διδεοξυνουκλεοτιδίων η οποία βασίζεται στην αντίστροφη επισημασμένων νουκλεοτιδίων τερματισμού αλυσίδας τα οποία ενσωματώνονται από μια DNA πολυμεράση κατά τη διάρκεια αντιγραφής ενός εκμαγείου (Sanger et al., 1977). Η μέθοδος διδεοξυνουκλεοτιδίων χρησιμοποιήθηκε μάλιστα και στον προσδιορισμό της αλληλουχίας του πρώτου ανθρώπινου γονιδιώματος. Πιο συγκεκριμένα, σε αρχικό βήμα το δείγμα DNA που πρόκειται να προσδιοριστεί συνδυάζεται σε ένα σωλήνα με εκκινητή, DNA πολυμεράση, τα

τριφωσφορικά νουκλεοτίδια DNA dNTPs (dATP, dTTP, dGTP, dCTP) και τα αλλά 4 φθοριοσημασμένα διδεοξυνουκλεοτίδια (ddNTPs). Το μείγμα αυτό αρχικά θερμαίνεται με σκοπό τον διαχωρισμό των δυο κλώνων και στη συνέχεια ψύχεται ούτως ώστε ο εκκινητής να μπορεί να δεσμευτεί στο μονόκλωνο πρότυπο. Η DNA πολυμεράση στη συνέχεια ξεκινάει τη σύνθεση του DNA ξεκινώντας από τον εκκινητή. Τα ddNTPs τροποποιούνται χημικά με ένα φθορίζον σήμα και στερούνται μια χημική ομάδα που διακόπτει το σχηματισμό φωσφοδιαστερικού δεσμού προκαλώντας έτσι την DNA πολυμεράση να σταματήσει αφού συνδεθεί ένα διδεοξυνουκλεοτίδιο. Έτσι, μόλις εισαχθεί ένα ddNTP στο αναπτυσσόμενο κλώνο DNA η σύνθεση του κλώνου διακόπτεται και η διαδικασία αυτή έπειτα από επαναλαμβανόμενους κύκλους ενίσχυσης θα έχει ως αποτέλεσμα την αναπαράσταση όλων των θραυσμάτων με διαφορετικό μήκος το καθένα τα άκρα των οποίων περιέχουν μια ετικέτα φθορισμού. Αφού ολοκληρωθεί η διαδικασία τα θραύσματα μεταφέρονται μέσω ενός σωλήνα που περιέχει μια μήτρα πηκτώματος γέλης, σε μια διαδικασία που καλείται τριχοειδή ηλεκτροφόρηση. Τα θραύσματα τα οποία έχουν μικρότερο μήκος διαπερνούν γρηγορότερα τους πόρους του πηκτώματος σε αντίθεση με αυτά με μεγαλύτερο μήκος. Όταν το κάθε θραύσμα φτάσει στο τέλος του σωλήνα και διασχίσει τη γραμμή τερματισμού, τότε φωτίζεται με λέιζερ επιτρέποντας έτσι την ανίχνευση του χρώματος. Τα δεδομένα αυτά που καταγράφονται από τον ανιχνευτή μπορούν με διάφορες μεθόδους υπολογιστικής Βιολογίας να συγκεντρωθούν σε ένα χρωματογράφημα το οποίο αναπαριστά την καθορισμένη αλληλουχία. Η αλληλουχία διαβάζεται από τις κορυφές του χρωματογραφήματος (Εικόνα 4).



Εικόνα 4 Μέθοδος Sanger (<https://upload.wikimedia.org/wikipedia/commons/b/b2/Sanger-sequencing.svg>)

### 1.3.1 Μεταγραφωμική: εξέλιξη και μέθοδοι

Το κεντρικό δόγμα της μοριακής βιολογίας περιγράφει τη ροή γενετικών πληροφοριών που αποθηκεύονται σε γονίδια ως DNA τα οποία μεταγράφονται σε RNA και τελικά μεταφράζονται σε πρωτεΐνες (Crick, 1970; Crick, 1958). Η απόλυτη έκφραση των γενετικών πληροφοριών που τροποποιούνται από περιβαλλοντικούς παράγοντες χαρακτηρίζει το φαινότυπο του οργανισμού. Η μεταγραφή ενός υποσυνόλου γονιδίων

σε συμπληρωματικά μόρια RNA καθορίζει την ταυτότητα του κυττάρου και ρυθμίζει τις βιολογικές δραστηριότητες εντός του κυττάρου. Αυτά τα μόρια RNA αναφέρονται συλλογικά ως μεταγράφομα το οποίο περιλαμβάνει σε δεδομένη χρονική στιγμή την ολική γενετική έκφραση ενός κυττάρου. Για την εξήγηση των λειτουργικών στοιχείων του γονιδιώματος και την κατανόηση των μεταλλάξεων και των ασθενειών είναι απολύτως αναγκαία η μελέτη της ολικής γενετικής έκφρασης καθώς μπορούμε να γνωρίζουμε ποια γονίδια εκφράζονται και ποια όχι.

Το μεταγράφομα έχει υψηλό βαθμό πολυπλοκότητας και περιλαμβάνει ποικίλους τύπους κωδικοποιητικών και μη κωδικοποιητικών ειδών RNA. Τα μόρια του αγγελιαφόρου RNA (**mRNA**) είναι τα πιο συχνά μελετημένα είδη RNA επειδή κωδικοποιούν πρωτεΐνες. Υπάρχουν ωστόσο και άλλα είδη μη κωδικοποιητικών RNA τα οποία έχουν υποστηρικτικό ρόλο, ρυθμιστικό ή εμπλέκονται σε άλλες βιολογικές διεργασίες. Για παράδειγμα το ριβοσωμικό RNA (**rRNA**) και το μεταφορικό RNA (**tRNA**) παρέχουν υποστηρικτικό ρόλο στη μετάφραση. Μια άλλη κατηγορία είναι τα μικρά πυρηνικά μόρια RNA (**snRNA**) τα οποία εμπλέκονται στο μάτισμα καθώς και τα μικρά πυρηνισκικά μόρια RNA (**snoRNA**) τα οποία εμπλέκονται στην τροποποίηση των rRNAs. Πιο πρόσφατα, ανακαλύφθηκαν νέες κατηγορίες RNA τα οποία έχουν ρυθμιστικό ρόλο στην γονιδιακή έκφραση σε μετα-μεταγραφικό επίπεδο στα οποία περιλαμβάνονται τα μικρο-RNA (microRNA- **miRNA**) και τα **piRNA** τα οποία αποτελούν δυνητικούς τόπους ανθρώπινης νόσου (Stefani and Slack, 2008).

Οι μέθοδοι οι οποίες πρωτοχρησιμοποιήθηκαν στις μελέτες έκφρασης γονιδίων ήταν χαμηλής απόδοσης όπως το στύπωμα κατά Northern και η αλυσιδωτή αντίδραση πολυμεράσης (PCR) οι οποίες όμως περιορίζονταν στη μέτρηση ενός απλού μεταγράφου. Τις τελευταίες δυο δεκαετίες, έχουν αναπτυχθεί μέθοδοι για την ποσοτική μέτρηση της γονιδιακής έκφρασης σε ολόκληρο το γονιδίωμα. Οι πρώτες μεταγραφωμικές μελέτες πραγματοποιήθηκαν χρησιμοποιώντας τεχνολογίες μικροσυστοιχιών βασισμένες σε υβριδοποίηση οι οποίες όμως αν και παρέχουν μια επιλογή υψηλής απόδοσης ανάλυση, είχε αρκετούς περιορισμούς αναφορικά με την προγενέστερη γνώση των αλληλουχιών, τη δυσκολία στην ανάλυση παρόμοιων αλληλουχιών καθώς και την περιορισμένη ικανότητα για τον ακριβή προσδιορισμό των γονιδίων που υπερεκφράζονταν ή υποεκφράζονταν αντίστοιχα. Σε αντίθεση με τις μεθόδους οι οποίες βασίζονται σε υβριδοποίηση έχουν αναπτυχθεί και διάφορες προσεγγίσεις βάσει αλληλουχίας που καθορίζουν άμεσα την αλληλουχία cDNA. Σε μελέτες έκφρασης γονιδίων χρησιμοποιήθηκε η αλληλουχία Sanger των βιβλιοθηκών ετικέτας εκφρασμένης αλληλουχίας (EST) ή του συμπληρωματικού DNA (cDNA) η οποία όμως είναι χαμηλής απόδοσης, έχει υψηλό κόστος και δεν είναι ιδανική για την ποσοτικοποίηση των μεταγράφων. Για την εξάλειψη των περιορισμών αυτών αναπτύχθηκαν νέες προσεγγίσεις με βάση την ακολουθία για να διευκρινιστεί το μεταγράφομα όπως είναι η σειριακή ανάλυση γονιδιακής έκφρασης SAGE, η ανάλυση γονιδιακής έκφρασης CAGE, και της μαζικής παράλληλης αλληλούχησης MPSS οι οποίες όμως δεν απαιτούν την προγενέστερη γνώση της αλληλουχίας του οργανισμού. Οι προσεγγίσεις αυτές βάσει ετικετών είχαν υψηλή απόδοση και μπορούν να παρέχουν



ακριβή «ψηφιακά» επίπεδα έκφρασης γονιδίων αλλά έχουν υψηλό κόστος (Harbers and Carninci, 2005; Reinartz et al., 2002; Velculescu et al., 1995).

Έτσι, η εξέλιξη της τεχνολογίας με την ανάπτυξη πιο εξειδικευμένων τεχνολογιών υψηλής απόδοσης με λιγότερο κόστος έχει φέρει επανάσταση στον κλάδο της μεταγραφωμικής εν συγκρίσει με τις προηγούμενες μεθόδους προσδιορισμού αλληλούχησης. Η αλληλούχηση επόμενης γενιάς (**Next Generation Sequencing - NGS**) επιτρέπει την ανάλυση RNA μέσω συμπληρωματικού DNA (cDNA) προσφέροντας περαιτέρω κατανόηση της περίπλοκης και δυναμικής φύσης του μεταγραφώματος. Οι εξελίξεις στη ροή εργασιών της RNA-seq, από την προετοιμασία του δείγματος και την πλατφόρμα αλληλουχίας έως την ανάλυση δεδομένων βιοπληροφορικής συνέβαλε σημαντικά στην αποσαφήνιση διαφόρων παθολογικών και φυσιολογικών καταστάσεων (Wang et al., 2009). Η αλληλούχηση RNA-seq χρησιμοποιεί τις δυνατότητες μεθόδων προσδιορισμού αλληλουχίας υψηλής απόδοσης για την παροχή υψηλών πληροφοριών σχετικά με το μεταγράφομα ενός κυττάρου. Η RNA-seq δεν χρησιμοποιείται αποκλειστικά για την ταυτοποίηση της έκφρασης κοινών ή σπάνιων μεταγράφων αλλά και για την αναγνώριση και άλλων μη φυσιολογικών γεγονότων. Η αλληλούχηση αυτή μπορεί να χρησιμοποιηθεί από τις περισσότερες πλατφόρμες επόμενης γενιάς (NGS) με τη μόνη διαφορά που έχουν οι βιβλιοθήκες οι οποίες διαφέρουν ανάλογα με τις ειδικές απαιτήσεις που έχει η κάθε μια ξεχωριστά.

### 1.3.2 Τεχνολογίες αλληλούχησης επόμενης γενιάς (NGS)

Βασικό χαρακτηριστικό της αλληλούχησης επόμενης γενιάς (NGS) είναι η χρήση των μαζικά παράλληλων μεθόδων. Με άλλα λόγια, αυτό σημαίνει πως μεγάλος αριθμός δειγμάτων DNA αλληλουχούνται δίπλα δίπλα ταυτόχρονα στην ίδια συσκευή. Υπάρχει μια ποικιλία τεχνικών αλληλούχησης επόμενης γενιάς που χρησιμοποιούν διαφορετικές τεχνολογίες. Ωστόσο οι περισσότερες από αυτές μοιράζονται ένα κοινό σύνολο χαρακτηριστικών που τους διακρίνουν από την αλληλούχηση κατά Sanger όπως αυτά είναι: πολλές αντιδράσεις εκτελούνται συγχρόνως, οι αντιδράσεις είναι μικρές και πολλές μπορούν να γίνουν ταυτόχρονα, λόγω το ότι γίνονται παράλληλα τα αποτελέσματα είναι έτοιμα σε λιγότερο χρόνο και η αλληλουχία ενός γονιδιώματος είναι φθηνότερη από ό,τι με την αλληλούχηση κατά Sanger. Οι τεχνολογίες NGS μπορούν να κατηγοριοποιηθούν σε δυο βασικές κατηγορίες. Η πρώτη ομάδα αναφέρεται στις τεχνολογίες που βασίζονται σε μεθόδους PCR και στις οποίες ανήκουν οι πλατφόρμες **Roche 454 Genome Sequencer**, **HiSeq 2000 (Illumina)**, **AB SOLID™ System** και **Ion Personal Genome Machine (PGM)** (Εικόνα 5). Η άλλη ομάδα που ονομάζεται τεχνολογίες αλληλουχίας «single molecule» (SMS), δεν βασίζεται σε μεθόδους PCR και στην ομάδα αυτή ανήκουν δυο μοριακά συστήματα αλληλουχίας, **HeliScope** και **PacBio RS** (Shokralla et al., 2012).

Category	Platform	Read length (bp)	Max. number of reads/run	Sequencing output/run	Run time
PCR-based NGS technologies	Roche 454 GS FLX	400–500	$1 \times 10^6$	≤500 Mb	10 h
	Roche 454 GS FLX+	600–800	$1 \times 10^6$	≤700 Mb	23 h
	Roche 454 GS Junior	400–450	$1 \times 10^5$	~35 Mb	10 h
	Illumina HiSeq 2000	100–200	$6 \times 10^9$	≤540–600 Gb	11 d
	Illumina HiSeq 1000	100–200	$3 \times 10^9$	≤270–300 Gb	8.5 d
	Illumina GAIIx	50–75	$6.4 \times 10^8$	≤95 Gb	7.5–14.5 d
	Illumina MiSeq	100–150	$7 \times 10^6$	≤1–2 Gb	19–27 h
	AB SOLiD 5500 system	35–75	$2.4 \times 10^9$	~100 Gb	4 d
	AB SOLiD 5500 xl system	35–75	$6 \times 10^9$	~250 Gb	7–8 d
	Ion Torrent -314 chip	100–200	$1 \times 10^6$	≥10 Mb	3.5 h
	Ion Torrent -316 chip	100–200	$6 \times 10^6$	≥100 Mb	4.7 h
	Ion Torrent -318 chip	100–200	$11 \times 10^6$	≥1 Gb	5.5 h
	SMS technologies	Helicos HeliScope	30–35	$1 \times 10^9$	~20–28 Gb
Pacific Biosciences system		≥1500	$50 \times 10^3$	~60–75 Mb	0.5 h

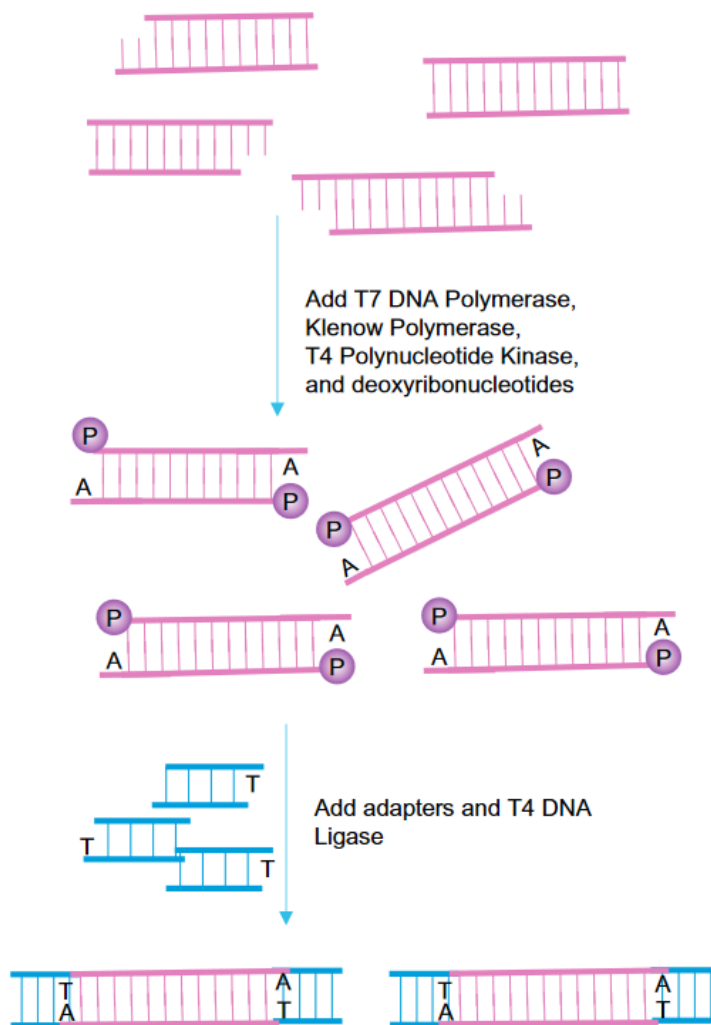
Εικόνα 5 Σύγκριση των τεχνολογιών NGS (Shokralla et al., 2012)

Συγκεκριμένα, οι διαδικασίες που εκτελούνται στις πλατφόρμες Roche 454 Pyrosequencing, AB SOLiD™ System και Ion Personal Genome Machine (PGM) μπορούν να διαχωριστούν στα εξής βήματα: 1) προετοιμασία βιβλιοθηκών, 2) Emulsion PCR (PCR γαλακτώματος) και 3) αλληλούχηση. Στη PCR γαλακτώματος δημιουργείται αρχικά μια βιβλιοθήκη DNA μέσω τυχαίου τεμαχισμού του γονιδιωματικού DNA. Στα άκρα των θραυσμάτων συνδέονται προσαρμογείς (**adapters**). Οι προσαρμογείς παρέχουν αλληλουχίες εκκίνησης τόσο για ενίσχυση όσο και για αλληλουχία. Ένας προσαρμογέας περιέχει μια ετικέτα 5'-βιοτίνης η οποία επιτρέπει την ακινητοποίηση της βιβλιοθήκης DNA σε σφαιρίδια επικαλυμμένα με στρεπταβιδίνη. Ο μη βιοτινυλιωμένος κλώνος απομακρύνεται και χρησιμοποιείται ως βιβλιοθήκη μονόκλωνου προτύπου DNA (sstDNA). Μονόκλωνα πρότυπα DNA βρίσκονται ενωμένα στην επιφάνεια των σφαιριδίων από τους προσαρμογείς οι οποίοι είναι συμπληρωματικοί με τους ολιγονουκλεοτιδικούς ανιχνευτές που βρίσκονται στην επιφάνεια. Κάθε σφαιρίδιο είναι προσαρτημένο σε ένα μόνο θραύσμα DNA από τη βιβλιοθήκη. Τα σφαιρίδια στη συνέχεια χωρίζονται σε σταγονίδια γαλακτώματος είτε νερού είτε ελαίου τα οποία περιέχουν επίσης το μίγμα αντίδρασης PCR. Στο υδατικό γαλάκτωμα νερού ή ελαίου καθένα από τα σταγονίδια που συλλαμβάνουν ένα σφαιρίδιο λειτουργεί ως μικροαντιδραστήρας PCR το οποίο παράγει χιλιάδες ενισχυμένα αντίγραφα μόνο του μοναδικού προτύπου DNA. Στη συνέχεια πραγματοποιείται η απαραίτητη διαδικασία προκειμένου να διατηρηθούν μόνο τα ενισχυμένα και να απορριφθούν τα αποτυχημένα. Τα σφαιρίδια αυτά τοποθετούνται στη συνέχεια σε πλάκα Picotiter (PTP) η οποία έχει πολλαπλά φρεάτια. Καθένα από τα φρεάτια αυτά μπορεί να συγκρατήσει μόνο ένα ενισχυμένο σφαιρίδιο. Μετά την φόρτωση των σφαιριδίων προστίθενται επιπλέον στρώματα κατασκευασμένων σφαιριδίων στο PPT. Το PPT βρίσκεται απέναντι από μια κάμερα CCD η οποία καταγράφει το εκπεμπόμενο φως από κάθε σφαιρίδιο (Mardis, 2008).

### 1.3.2.1 Κατασκευή βιβλιοθηκών

Σε κάθε ανάλυση NGS το αρχικό βήμα είναι η κατασκευή βιβλιοθηκών. Πρωτίστως γίνεται η απομόνωση του γονιδιωματικού DNA το οποίο μπορεί να ληφθεί είτε από δείγμα αίματος είτε από το ξύσιμο του εσωτερικού από το μάγουλο ενός ατόμου. Σε επόμενο βήμα της διαδικασίας θα πρέπει να γίνει ο κατακερματισμός των δειγμάτων σε μικρότερα κομμάτια δίκλωνου DNA, διαδικασία που μπορεί να γίνει με τη χρήση

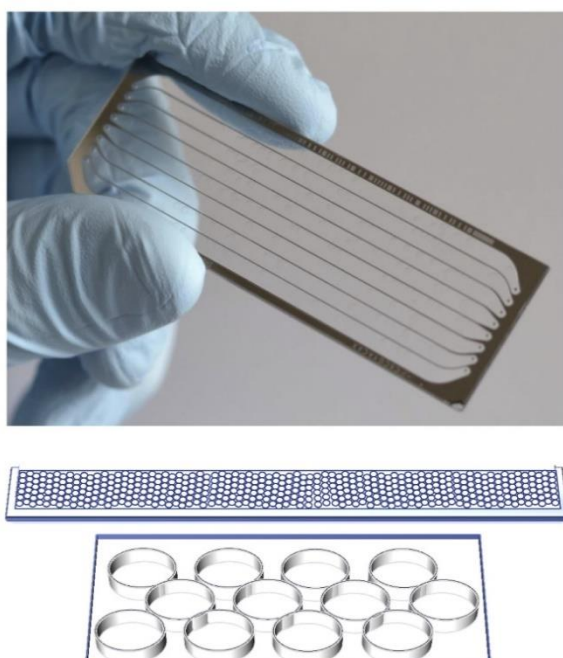
των υπερηχητικών κυμάτων υψηλότερα από αυτά που μπορεί να ακούσει το άτομο. Καθώς τα κύματα περνούν το υγρό που περιβάλλει το γονιδιωματικό DNA σχηματίζονται μικρές φυσαλίδες παγιδευμένων αερίων. Η δύναμη αυτών μπορεί να σπάσει τους μοριακούς δεσμούς και αφού γίνει η κατεργασία με τους υπερήχους γίνεται και η διάσπαση σε μικρότερα μόρια. Στη συνέχεια τα θραύσματα που προκύπτουν είναι τυχαία και έχουν διαφορετική αλληλουχία DNA καθώς επίσης και τα άκρα τους είναι ανομοιόμορφα. Αρχικά, η DNA πολυμεράση T4 γεμίζει τις μονόκλωνες περιοχές με συμπληρωματικά νουκλεοτίδια. Η Klenow πολυμεράση προσθέτει ένα μόνο A στο 3' άκρο. Η προεξοχή A βοηθάει την προσθήκη του προσαρμογέα, ο οποίος συντίθενται για να έχει μια μόνο προεξοχή T. Έτσι λοιπόν, χρησιμοποιώντας τους προσαρμογείς τα οποία είναι μικρά δίκλιωνα μόρια DNA των οποίων η αλληλουχία είναι γνωστή και αναγνωρίσιμη από το μηχάνημα αυτοί προσδένονται σε κάθε άκρο των μικρών αυτών θραυσμάτων με τη βοήθεια της DNA λιγάσης όπως φαίνεται και στην **(Εικόνα 6)** (Clark, 2019).



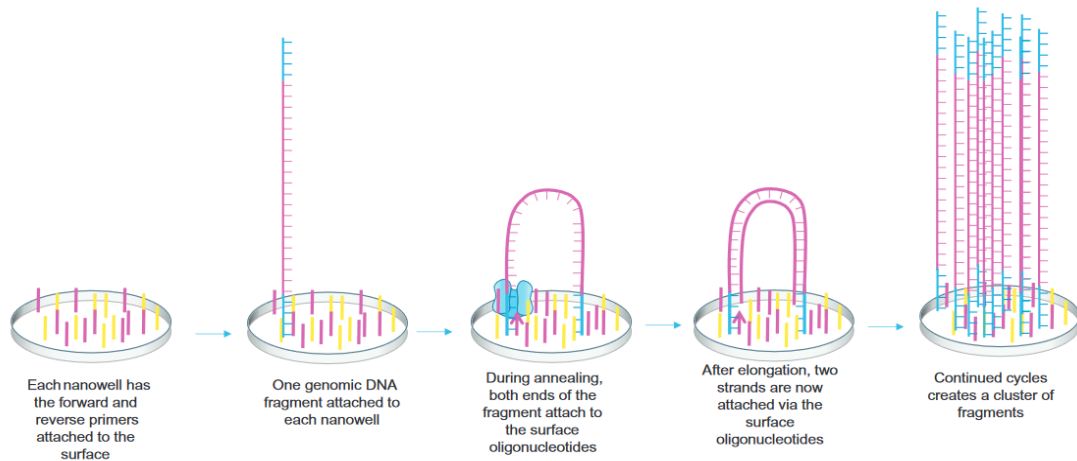
Εικόνα 6 Κατακερματισμός του γονιδιωματικού DNA σε μικρότερα θραύσματα (Clark, 2019)

### 1.3.2.2 Τεχνολογία της Illumina

Η πλατφόρμα της Illumina, παλαιότερα γνωστή και ως Solexa, παρουσίασε τον αναλυτή γονιδιώματος το 2007 η οποία λόγω της υψηλής χωρητικότητας σύντομα χρησιμοποιήθηκε για εφαρμογές αναλύσεων ολόκληρου του γονιδιώματος. Ορισμένα από τα μηχανήματα της Illumina είναι: Illumina MiniSeq, Illumina MiSeq, Illumina NextSeq, Illumina HiSeq και Illumina NovaSeq. Σε αρχικό στάδιο οι αλληλουχίες που πρόκειται να αναλυθούν προσδένονται σε μια επιφάνεια πυριτίου κυψέλη ροής (**flow cell**) η οποία μοιάζει με μια γυάλινη πλάκα. Κάθε κυψέλη ροής είναι χωρισμένη σε 8 ξεχωριστές λωρίδες (**Εικόνα 7**). Οι εσωτερικές επιφάνειες των κυψελών έχουν συνδεδεμένα ολιγονουκλεοτιδικές αλληλουχίες συμπληρωματικές με τα ολιγονουκλεοτίδια στους προσαρμογείς που συνδέονται με τα θραύσματα της βιβλιοθήκης. Τα ολιγονουκλεοτίδια που βρίσκονται στο κύτταρο ροής υβριδοποιούνται με τα ολιγονουκλεοτίδια στην επιφάνεια προκειμένου να σχηματίσουν ένα αρχικό αντίγραφο του μεμονωμένου μορίου προτύπου. Τα αρχικά μόρια της βιβλιοθήκης αφαιρούνται και τα αντιγραφόμενα θραύσματα που συνδέονται στην κυψέλη ροής χρησιμοποιούνται για τη δημιουργία συστάδων (**clusters**) πανομοιότυπων μορίων προτύπου χρησιμοποιώντας ισοθερμική ενίσχυση. Η δομή αυτή μοιάζει με γέφυρα, οπότε ο τύπος PCR καλείται ενίσχυση γέφυρας (**Bridge amplification**). Κάθε συστάδα αντιπροσωπεύει μόνο ένα μικρό μήκος πληροφοριών αλληλουχίας DNA. Όμως υπάρχουν εκατομμύρια συστάδες σε μια κυψέλη ροής παρέχοντας πληροφορίες συνολικά για ολόκληρο το γονιδίωμα (**Εικόνα 8**) (Buermans and den Dunnen, 2014; Clark, 2019).



Εικόνα 7 Κυψέλη ροής (flow cell) (Clark, 2019)

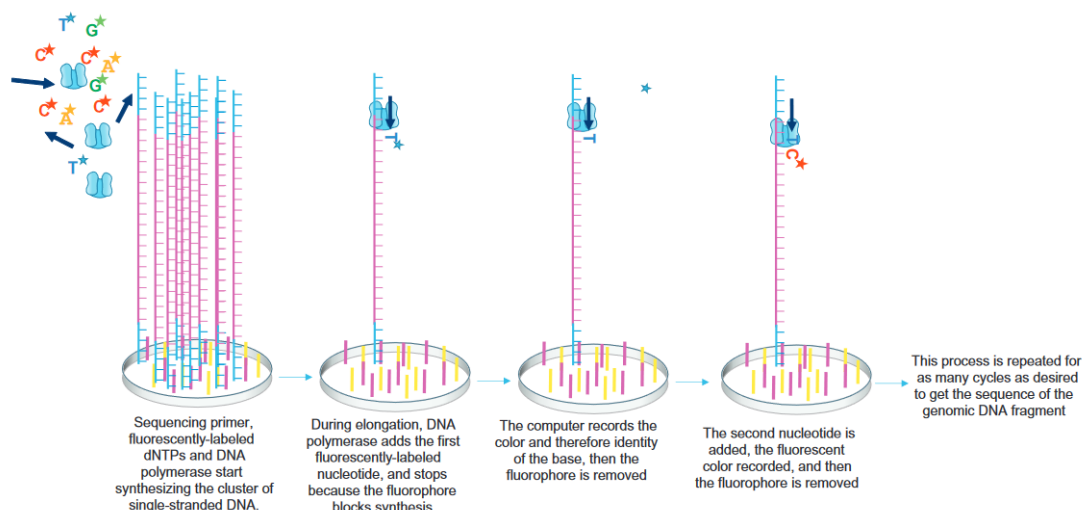


Εικόνα 8 Ενίσχυση γέφυρας για τη δημιουργία συστάδων πανομοιότυπων θραυσμάτων (Clark, 2019)

### 1.3.2.3 Αλληλούχηση RNA-seq

Η αλληλούχηση των δειγμάτων στην Illumina γίνεται με τη μέθοδο Sequencing by Synthesis (SBS) σε συνδυασμό με την ενίσχυση γέφυρας στην επιφάνεια μιας κυψέλης ροής. Η αλληλούχηση όλων των συστάδων γίνεται ταυτόχρονα βάση προς βάση παράλληλα και προσδιορίζεται με τη βοήθεια ενός εκκινητή αλληλουχίας, την DNA πολυμεράση και 4 φθορίζοντα δεοξυνουκλεοτίδια ως δομικά στοιχεία με διαφορετικό φθορισμό το καθένα. Τα νουκλεοτίδια με ετικέτα φθορισμού ένα κάθε φορά (κυκλικά) πλησιάζουν τη βάση της συστάδας και καταγράφονται εικόνες φθορισμού για να προσδιοριστεί πιο νουκλεοτίδιο ενσωματώνεται σε κάθε κύκλο. Αφού ολοκληρωθεί η διαδικασία για την πρώτη βάση και έχει ελεγχθεί ότι είναι η σωστή βάση αυτή που έχει προσδεθεί με τη συγκεκριμένη βάση της συστάδας, τότε καταγράφεται το χρώμα της φθορίζουσας και συνεχίζει στην επόμενη βάση.

Η διαδικασία αυτή συνεχίζεται για κάθε συστάδα και ο υπολογιστής καταγράφει κάθε φορά την ταυτότητα του προστιθέμενου νουκλεοτιδίου. Αποτέλεσμα αυτού είναι η δημιουργία συμπληρωματικών αλυσίδων των συστάδων που ονομάζονται αναγνώσματα (**reads**). Δεδομένου ότι υπάρχουν δεκάδες εκατομμύρια συστάδες σε μια κυψέλη ροής, πρακτικά δημιουργούνται δεκάδες εκατομμύρια αναγνώσματα. Ο αριθμός των συστάδων ισούται με τον αριθμό των αναγνωσμάτων και ο αριθμός των κύκλων ισούται με το μήκος των αναγνωσμάτων που όμως εξαρτάται από την πλατφόρμα αλληλούχησης που θα χρησιμοποιηθεί (**Εικόνα 9**). Συμπερασματικά, τα δεδομένα που προκύπτουν αποτελούνται από αρχεία με πληροφορίες για κάθε ανάγνωση και περιλαμβάνουν τα πρωτογενή δεδομένα (**raw data**) σε μορφή .fastQ με τα οποία πραγματοποιείται η RNA-seq.



Εικόνα 9 Sequencing By Synthesis (Clark, 2019)

### 1.3.2.4 Ανάλυση Διαφορικής Έκφρασης Γονιδίων (Differential Gene Expression Analysis)

Η ανάλυση διαφορικής έκφρασης είναι η διαδικασία κατά την οποία πραγματοποιείται η λήψη κανονικοποιημένων δεδομένων μέτρησης αναγνώσεων και εκτέλεση στατιστικής ανάλυσης με σκοπό την ανακάλυψη ποσοτικών αλλαγών στα επίπεδα έκφρασης μεταξύ των υπό μελέτη ομάδων. Κύριος σκοπός μιας ανάλυσης διαφορικής έκφρασης είναι η κατανόηση της διαφορικής έκφρασης γονιδίων σε μια συνθήκη μελέτης (test) ως προς τη συνθήκη ελέγχου (control), π.χ. παθολογικοί ιστοί σε σχέση με φυσιολογικούς ιστούς. Σε ένα πείραμα αλληλούχησης επόμενη γενιάς είναι σημαντική η ποσοτικοποίηση των αποτελεσμάτων. Η ποσοτικοποίηση πραγματοποιείται στη βάση του αριθμού των αναγνώσεων που αντιστοιχεί σε ένα μεταγράφο ή γονίδιο. Εξαιτίας του μήκους του μεταγράφου, ενδέχεται τα μεγάλα γονίδια να παρουσιάσουν μεγαλύτερο αριθμό αναγνώσεων και αυτό λόγω του μεγέθους και έτσι να δίνουν και μεγαλύτερες τιμές των αναγνώσεων ανά γονίδιο. Επομένως, για την εκτίμηση του βαθμού έκφρασης των γονιδίων υπολογίζεται μια διορθωμένη τιμή RPKM (Reads per Kilobase per Million) ή FPKM (Fragments per Kilobase per Million). Με αυτόν τον τρόπο εξασφαλίζεται ότι ο βαθμός έκφρασης ενός γονιδίου είναι ανεξάρτητος και από το μήκος αλλά και από το βάθος της αλληλούχησης (Νικολάου and Χουβαρδάς, 2015). Υπάρχουν διαφορετικές μέθοδοι για ανάλυση διαφορικής έκφρασης και είναι σημαντικό να λαμβάνεται υπόψη ο πειραματικός σχεδιασμός κατά την επιλογή μιας μεθόδου ανάλυσης διότι υπάρχουν ορισμένα εργαλεία διαφορικής έκφρασης τα οποία πραγματοποιούν σύγκριση μόνο κατά ζεύγη και άλλα τα οποία μπορούν να εκτελέσουν πολλαπλές συγκρίσεις.

### 1.3.2.5 Χρήση των τεχνολογιών NGS και εφαρμογές αυτών στην κλινική πράξη και έρευνα

Οι τεχνολογίες NGS έχουν ολοένα και μεγαλύτερο αντίκτυπο σε πολλαπλούς επιστημονικούς κλάδους και αποτελούν ένα από τα πιο μεγάλα και σημαντικά άλματα της Βιολογίας των τελευταίων χρόνων. Οι διάφορες τεχνικές των τεχνολογιών αυτών έχουν δώσει καλύτερη εικόνα στη Βιολογία του όγκου, επιτρέποντας έτσι τον

εντοπισμό γενετικών μεταλλάξεων που οδηγούν στην εξέλιξη του καρκίνου. Έχουν κλινικές εφαρμογές για την αποτελεσματική ανίχνευση είτε κληρονομικών είτε σωματικών μεταλλάξεων στα καρκινικά γονίδια. Αυτό δίνει την ελπίδα ότι αυτή η γνώση θα χρησιμοποιηθεί κλινικά και θα βελτιώσει τη διάγνωση, την πρόγνωση, την παρακολούθηση της νόσου καθώς και τον προσδιορισμό ορισμένων στόχων για θεραπεία. Οι τεχνολογίες NGS χρησιμοποιούν διαφορετικές μεθοδολογίες για ανάλυση του γενώματος, μεταγραφώματος και επιγενώματος. Οι διαφορετικές ωμικές τεχνολογίες είναι απαραίτητες για να αποδοθεί μια πλήρης εικόνα των γενετικών αλλαγών. Η αλληλούχηση επόμενης γενιάς μπορεί να προσφέρει πρωτοφανείς γνώσεις στην καρδιολογία, τη νευρολογία και την ψυχιατρική, σε αιματολογικές διαταραχές και τη φαρμακογονιδιωματική, αποκαλύπτοντας με αυτόν τον τρόπο την περιπλοκότητα του γονιδιώματος και του μεταγραφώματος στην έρευνα βιοδεικτών, τις μελέτες γονιδιακής έκφρασης, την ιική επιδημιολογία, την επιτήρηση ασθενών και τη δημιουργία ενός εξατομικευμένου προφίλ με απώτερο στόχο την εξατομικευμένη θεραπεία του (Demkow and Ploski, 2015).

## 2. Σκοπός

Η νόσος του Πάρκινσον είναι νευροεκφυλιστική νόσος, όπου αποτέλεσμα του εκφυλισμού είναι κάποια γονίδια να αλλάζουν έκφραση μεταξύ των παθολογικών και των φυσιολογικών ιστών στον εγκέφαλο. Σκοπός της παρούσας διπλωματικής εργασίας είναι η εύρεση γονιδίων που εκφράζονται διαφορετικά σε εγκεφάλους ασθενών με τη νόσο του Πάρκινσον και υγιών, χρησιμοποιώντας δεδομένα αλληλούχησης επόμενης γενιάς (RNA-seq). Στη συνέχεια με τη μέθοδο της μετα-ανάλυσης θα βελτιωθούν οι εκτιμήσεις κάθε ξεχωριστής μελέτης και θα ταυτοποιηθούν οι βασικές βιολογικές διεργασίες που σχετίζονται με την εμφάνιση του φαινοτύπου. Συνακόλουθα, γνωρίζοντας ποια είναι τα γονίδια αυτά που υπερεκφράζονται ή υποεκφράζονται αντίστοιχα, αυτό θα βοηθήσει την εύρεση βιοδεικτών τα οποία θα συμβάλλουν ολοένα και περισσότερο στη θεραπεία της νόσου και θα μπορούμε εκ των προτέρων να γνωρίζουμε σε ποια φάση της ασθένειας βρίσκεται ο ασθενής με σκοπό την παροχή της εξατομικευμένης θεραπείας του.

## 3. Μέθοδοι

### 3.1 Λήψη πρωτογενών δεδομένων

Ξεκινώντας την ανάλυση των δεδομένων θα πρέπει αρχικά να γίνει η σωστή διαχείριση των αρχείων αλληλούχησης και ο έλεγχος ποιότητας για κάθε ανάγνωση. Τα δεδομένα τα οποία εξάγονται από τα μηχανήματα NGS της Illumina είναι σε μορφή .fastq. Η μορφή που έχει ένα .fastq αρχείο είναι πρακτικά ένα αρχείο κειμένου το οποίο παρουσιάζει την αλληλουχία της κάθε ανάγνωσης καθώς και βαθμό ποιότητας της κάθε ανάγνωσης. Ένα ενδεικτικό παράδειγμα του τι εμπεριέχει ένα αρχείο .fastq φαίνεται παρακάτω όπου εδώ βλέπουμε μια ανάγνωση:

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*(((((***+))%%#++)(%%%) .1***-+*''))**55CCF>>>>>>CCCCCCC65
```

Κάθε αρχείο αποτελείται από 4 ξεχωριστές γραμμές. Η πρώτη γραμμή ξεκινάει με το σύμβολο «@» και ακολουθεί το όνομα της κάθε ανάγνωσης. Στη δεύτερη γραμμή περιλαμβάνεται η αλληλουχία της κάθε ανάγνωσης σε μορφή FASTA και αποτυπώνεται με τα γράμματα A, T, G, C. Στη τρίτη γραμμή το σύμβολο «+», διαχωρίζει την αλληλουχία από τις βαθμολογίες ποιότητας (quality scores), και τέλος η τέταρτη γραμμή δημιουργήθηκε στο πρόγραμμα Phred και αποτυπώνει την ποιότητα της κάθε βάσης με κωδικοποιημένη μορφή η οποία γραμμή έχει τον ίδιο αριθμό συμβόλων με τα γράμματα της ακολουθίας στη γραμμή 2. Οι βαθμολογίες ποιότητας σχετίζονται λογαριθμικά με τις πιθανότητες η συσκευή να έχει κάνει λάθος στην τοποθέτηση της κάθε βάσης (base calling). Ο υπολογισμός των βαθμολογιών της ποιότητας Phred ορίζεται από τον μαθηματικό τύπο (Ewing and Green, 1998):

$$Q_{PHRED} = -10 \cdot \log_{10}(p)$$

όπου  $p$  είναι η πιθανότητα σφάλματος.

Phred Quality Score	Πιθανότητα λανθασμένης τοποθέτησης	Ακρίβεια τοποθέτησης
10	1/10	90%
20	1/100	99%
30	1/1000	99.9%
40	1/10000	99.99%
50	1/100000	99.999%

Για παράδειγμα αν ο Phred τοποθετήσει τη βαθμολογία Q30 σε μια βάση, αυτό ισοδυναμεί με την πιθανότητα λανθασμένης τοποθέτησης 1/1000. Αυτό πρακτικά σημαίνει ότι η τοποθέτηση της βάσης αυτή είναι 99.9% ακριβής, άρα και μικρότερη πιθανότητα λάθους όπως φαίνεται και στον παρακάτω πίνακα. Όσο μεγαλύτερο είναι το Q τόσο μικρότερη είναι η πιθανότητα λάθους. Χαμηλότερο Q έχει μεγαλύτερη εσφαλμένη τοποθέτηση. Όταν η πιθανότητα της αλληλούχησης φτάσει στο Q30, τότε σχεδόν όλα τα δείγματα θα είναι πλήρως αποδεκτά. Γι' αυτό το λόγο το Q30 θεωρείται σημείο αναφοράς για την ποιότητα στην αλληλούχηση επόμενης γενιάς.

### 3.2 Βιολογικές Βάσεις Δεδομένων

Οι βιολογικές βάσεις δεδομένων αποτελούν σημαντικό κομμάτι στη διεξαγωγή μιας έρευνας, αφού μέσω αυτών ένας ερευνητής μπορεί να αντλήσει δεδομένα στα οποία θα βασίζεται και στη συνέχεια η ανάλυση του. Η ύπαρξη των δεδομένων αυτών είναι ιδιαίτερα σημαντική, καθώς στις περισσότερες περιπτώσεις ο ερευνητής ο οποίος παράγει σημαντικής ποιότητας και αξιολόγησης πρωτογενή δεδομένα, καλείται να τα αναρτήσει σε δημόσιες βάσεις με σκοπό να μπορούν να χρησιμοποιηθούν και από άλλους ερευνητές από κάθε επιστημονική κοινότητα, ακόμη και να επανελεγχθούν. Παραδείγματα των τριών μεγάλων βάσεων που είναι ελεύθερα διαθέσιμες στο διαδίκτυο και προσβάσιμες από τον κάθε χρήστη είναι οι: DNA Data Bank of Japan



(DDBJ) (Okido et al., 2022) που είναι διεθνώς αναγνωρίσιμη και περιλαμβάνει νουκλεοτιδικές αλληλουχίες, η European Nucleotide Archive (ENA) (Cummins et al., 2022) η οποία είναι η πιο μεγάλη βάση που περιλαμβάνει επίσης νουκλεοτιδικές αλληλουχίες και η GenBank (Sayers et al., 2022) η οποία και αυτή περιλαμβάνει νουκλεοτιδικές αλληλουχίες και τα δεδομένα της μπορούν να χρησιμοποιηθούν για πειραματικούς σκοπούς. Η International Nucleotide Sequence Database Collaboration (INSDC) (Arita et al., 2021) αποτελεί το συνδυασμό των τριών αυτών μεγάλων βάσεων και αντιπροσωπεύει μια από τις μεγαλύτερες επιτεύξεις στο κομμάτι της κοινής χρήσης δεδομένων δημόσιου τομέα. Πρακτικά, μέσω αυτής της βάσης δεδομένων διασφαλίζεται ότι όλες οι νουκλεοτιδικές αλληλουχίες που κατατίθενται στα αρχεία διατηρούνται ως μέρος του επιστημονικού αρχείου και είναι προσβάσιμες συγχρόνως σε όλες τις βάσεις δεδομένων. Σκοπός είναι τα δεδομένα από όλες τις βάσεις να μπορούν να συνδυάζονται για να χρησιμοποιηθούν αποτελεσματικότερα από τις διάφορες τεχνολογίες και μεθόδους αλληλούχησης (Karsch-Mizrachi et al., 2018).

### 3.3 Συλλογή Δεδομένων

Προκείμενου να συγκεντρωθούν όλες οι απαραίτητες μελέτες με τις πληροφορίες που χρειάζονται για να πραγματοποιηθεί η ανάλυση μας, χρησιμοποιήθηκε η αναζήτηση στις κατάλληλες βάσεις δεδομένων από όπου και έγινε η συλλογή των δεδομένων, καθώς και η χρήση διαφόρων εργαλείων Βιοπληροφορικής για την ανάλυση και την επεξεργασία αυτών. Οι Βάσεις Δεδομένων που χρησιμοποιήθηκαν και μας παρείχαν τις κατάλληλες πληροφορίες είναι η Bioproject (Barrett et al., 2012) και η ENA.

Εισάγουμε στη βάση δεδομένων Bioproject το παρακάτω query, προκείμενου να συλλέξουμε πληροφορίες πιθανώς σχετικές με την μελέτη μας, έρευνες που περιείχαν δεδομένα από RNA-seq πειράματα:

```
txid9606[Organism] AND "transcriptome gene expression"[Filter]  
AND Parkinson*[All Fields]
```

Συγκεντρώθηκαν συνολικά 174 μελέτες. Μελετήθηκε η καθεμία ξεχωριστά δίνοντας σημαντικές πληροφορίες για τον τρόπο διεξαγωγής της εκάστοτε μελέτης, τις τεχνικές που χρησιμοποιήθηκαν καθώς και τον κωδικό πρόσβασης (accession code) το οποίο θα χρησιμοποιηθεί στη συνέχεια στη Βάση ENA η οποία περιλαμβάνει τα .fastq αρχεία που χρειαζόμαστε. Οι μελέτες αυτές αρχικά καταγράφηκαν σε ένα συγκεντρωτικό φύλλο excel διαχωρίζοντας το σε 5 στήλες (**Πίνακας 4**). Η πρώτη στήλη περιλαμβάνει το accession code της κάθε μελέτης, η δεύτερη τα sample (yes/no) την ύπαρξη ή μη δειγμάτων στη μελέτη αυτή, η τρίτη στήλη το tissue εάν πρόκειται δηλαδή για ιστό από εγκέφαλο ή εάν πρόκειται για δείγμα αίματος ή ακόμη και εργαστηριακά δείγματα, η τέταρτη το experiment type, δηλαδή το είδος του πειράματος και τέλος τον αριθμό των δειγμάτων των υγιών και των παθολογικών.

Accession_code	Samples(Yes/No)	Tissue	Experiment type	Accept
PRJNA721677	YES	iPSC-derived dopaminergic neurons	high throughput sequencing	No
PRJNA721657	YES	neuroblastoma cells	high throughput sequencing	No
PRJNA662780	YES	postmortem midbrain tissue	high throughput sequencing	Yes
PRJNA655240	YES	Extracellular vesicles derived from plasma	high throughput sequencing	No
PRJNA637314	YES	dopaminergic neurons	Expression profiling by array	No
PRJNA283498	YES	prefrontal cortex/BA9	profiling by high throughput sequencing	YES
PRJNA266233	No records	No records	No records	No
PRJNA264625	YES	iPS cell-derived neuronal samples	Expression profiling by high throughput sequencing	No
PRJNA176709	YES	Blood	Non-coding RNA profiling by high throughput sequencing	No

Πίνακας 4 Ενδεικτικό παράδειγμα ορισμένων μελετών και διαχωρισμός αυτών με βάση τον κωδικό, την ύπαρξη ή μη δειγμάτων, τον ιστό και το είδος του πειράματος

Ακολουθήθηκαν συγκεκριμένα βήματα προκειμένου να ληφθούν οι κατάλληλες για την εργασία μας μελέτες οι οποίες θα μας έδιναν τις απαραίτητες πληροφορίες. Τα κριτήρια επιλογής/απόρριψης μιας μελέτης είναι τα εξής:

- Απορρίφθηκαν όσες μελέτες δεν είχαν δείγματα ή είχαν δείγματα μόνο από φυσιολογικούς και όχι από παθολογικούς ιστούς εγκεφάλου και το αντίστροφο
- Απορρίφθηκαν όσες μελέτες αφορούσαν stem cells, in vitro ανάλυση, miRNA, non-coding RNA και δείγματα αίματος
- Απορρίφθηκαν όσες μελέτες δεν χρησιμοποιούσαν την τεχνολογία της Illumina αλλά άλλες εταιρίες οι οποίες χρησιμοποιούν διαφορετικές τεχνικές όπως για παράδειγμα η ChIP-seq analysis. Χρειαζόμαστε δεδομένα που χρησιμοποιούν high throughput sequencing
- Επιλέχθηκαν οι μελέτες όπου πληρούσαν τα παραπάνω κριτήρια και αφορούσαν συγκεκριμένα σημεία του εγκεφάλου
- Έγινε έλεγχος των δειγμάτων προκειμένου να προέρχονται από τον ίδιο ιστό διαφορετικά γίνεται ο διαχωρισμός αυτών σε δυο ξεχωριστές μελέτες

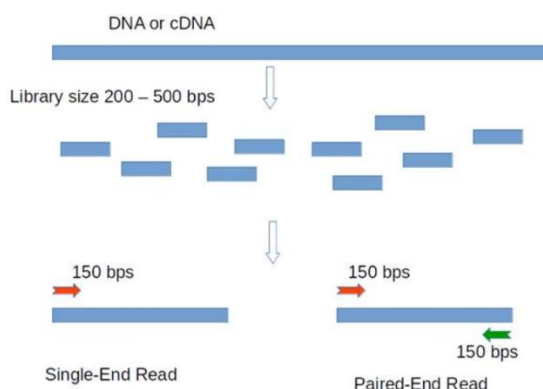
Συμπερασματικά οι μελέτες αυτές με τα παραπάνω κριτήρια επιλογής ήταν 4. Τα BioProject accession code αυτών είναι: PRJNA283498 (Dumitriu et al., 2016), PRJNA557205 (Li et al., 2020), PRJNA563007 (Xicoy et al., 2020) και PRJNA662780 (Smajić et al., 2020). Τα χαρακτηριστικά των μελετών που χρησιμοποιήθηκαν για την μετα- ανάλυση διαφαίνονται παρακάτω (Πίνακας 5).

Bioproject	Sample(yes/no)	Tissue	Experiment type	Sample count
PRJNA662780	Yes	Postmortem midbrain tissue	High throughput sequencing	6 controls 5 Parkinson
PRJNA563007	Yes	Substantia nigra/Putamen samples	High throughput sequencing	Nigra: 5 controls 5 Parkinson Putamen: 3 controls 3 Parkinson
PRJNA557205	Yes	Brain prefrontal cortex	High throughput sequencing	12 controls 24 Parkinson
PRJNA283498	Yes	Prefrontal cortex/BA9	High throughput sequencing	44 controls 29 Parkinson

Πίνακας 5 Περιγράφονται αναλυτικά οι 4 μελέτες που συμπεριλήφθηκαν στην εργασία, το είδος του ιστού της κάθε μελέτης, το είδος του πειράματος και το σύνολο των δειγμάτων(υγιών/παθολογικών)

### 3.4 Paired-end και Single-end Δεδομένα

Σε ένα πείραμα αλληλούχησης επόμενης γενιάς υπάρχει η επιλογή αλληλούχησης είτε ενός μόνο άκρου των θραυσμάτων και έτσι αναφερόμαστε στα **Single -end reads**, είτε και των δύο άκρων των θραυσμάτων που τότε αναφερόμαστε στα **Paired -end reads** δημιουργώντας έτσι δυο αρχεία των αναγνώσεων το ένα προς τα εμπρός και το άλλο προς την αντίστροφη κατεύθυνση (**Εικόνα 10**). Η paired- end αλληλούχηση διευκολύνει την ανίχνευση γονιδιωματικών αναδιατάξεων, συγχωνεύσεων και νέων μεταγράφων (Illumina Sequencing). Όλα στα συστήματα αλληλούχησης επόμενης γενιάς της Illumina παρέχουν τη δυνατότητα για paired-end αλληλούχηση. Από τις μελέτες μόνο η μελέτη με PRJNA563007 ήταν paired-end.



Εικόνα 10 <https://www.1010genome.com/illumina-sequencing/>

### 3.5 NGS Pipelines: Bcbio-nextgen

Ως ροή εργασίας (Pipeline) καλείται ένα σύνολο δεδομένων τα οποία χρησιμοποιούνται για να φιλτράρουν και να επεξεργασθούν δεδομένα. Καθώς εκτελείται μια ροή εργασίας ο κάθε χρήστης μπορεί να μετασχηματίσει τα δεδομένα και να τροποποιήσει ο ίδιος το χρόνο που εκτελείται ο αλγόριθμος, τη θερμοκρασία του υπολογιστή αλλά ακόμη και τους πυρήνες που χρησιμοποιούνται για την επεξεργασία των δεδομένων αυτών. Κάθε ροή εργασίας έχει ένα σημείο έναρξης και λήξης και ο κάθε χρήστης επιλέγει την κατάλληλη ροή εργασίας ιδανική για την ανάλυση την οποία καλείται να εκτελέσει. Η ροή εργασίας που χρησιμοποιείται στην παρούσα διπλωματική εργασία είναι το **Bcbio-nextgen** (Guimera, 2012). Η τεχνολογία RNA-Seq η οποία μελετά το συνολικό mRNA κυτταρικών πληθυσμών ή τμημάτων ιστών καλείται bulk RNA-Seq. Όπως αναφερθήκαμε και παραπάνω τα πρωτογενή δεδομένα τα οποία παράγονται από πειράματα RNA-Seq είναι αρχεία FASTQ τα οποία εμπεριέχουν τις αλληλουχίες, καθώς και μια τιμή για την ποιότητα της κάθε βάσης. Η προ επεξεργασία των δεδομένων RNA-Seq καθορίζεται από ορισμένα βήματα:

- 1) Ποιοτικό έλεγχο και περικοπή των αναγνώσεων αλληλουχίας
- 2) Αντιστοίχιση των αναγνώσεων στο γονιδίωμα αναφοράς
- 3) Παραγωγή μετρήσεων ανάγνωσης γονιδίων
- 4) Κανονικοποίηση

Το αρχικό βήμα της ροής εργασίας περιλαμβάνει τη σωστή διαχείριση των αρχείων αλληλούχησης, τον ποιοτικό έλεγχο των αλληλουχιών και αποκλεισμό όσων δεν πληρούν τις προϋποθέσεις ποιότητας με τη χρήση διαφόρων εργαλείων συμπεριλαμβανόμενων των FastQC (Andrews, 2019), MultiQC (Ewels et al., 2016) και Qualimap (Garcia-Alcalde et al., 2012). Το πιο συχνά χρησιμοποιούμενο λογισμικό είναι το FastQC το οποίο παρέχει πληροφορίες για κάθε ανάγνωση, καθώς και τις προβληματικές περιοχές αυτών. Οι βασικές λειτουργίες του FastQC είναι πως δέχεται ως είσοδο τα FASTQ αρχεία, δεδομένα από αρχεία BAM (Binary Alignment Map) (Li et al., 2009) και μας παρέχει σημαντικές πληροφορίες για: την ποιότητα της κάθε ανάγνωσης, αν και ποιες βάσεις του προσαρμογέα εντοπίζονται και τέλος, δίνει συνοπτικά γραφήματα και πίνακες για γρήγορη αξιολόγηση των δεδομένων και εξαγωγή αποτελεσμάτων σε μια σελίδα HTML με δυνατότητα αποθήκευσης. Το φιλτράρισμα των δεδομένων πραγματοποιείται με τη χρήση διαφόρων εργαλείων συμπεριλαμβανόμενων των Cutadapt (Martin, 2011), fastp (Chen et al., 2018), Trim Galore! (Krueger, 2012), Trimmomatic (Bolger et al., 2014), AdapterRemoval (Lindgreen, 2012) και Skewer (Jiang et al., 2014) με σκοπό τη βελτιστοποίηση των αλληλουχιών σε κάθε δείγμα και περικοπή όσων έχουν ποιότητα μικρότερη του 20, καθώς και την πλήρη αποκοπή των αλληλουχιών του προσαρμογέα ώστε να μπορούν να αγνοηθούν σε επόμενα στάδια. Στις περισσότερες περιπτώσεις πραγματοποίησης αναλύσεων NGS απαιτείται ένα ολοκληρωμένο και καλά σχολιασμένο γονιδίωμα αναφοράς το οποίο αποτελεί πρότυπο για την εύρεση των διαφόρων μεταλλάξεων. Ένας χαρτογράφος λαμβάνει ως είσοδο ένα γονιδίωμα αναφοράς και το σύνολο των

αναγνώσεων. Ένα γονιδίωμα αναφοράς (Genome Reference), είναι μια ψηφιακή βάση δεδομένων αλληλουχίας νουκλεϊκών οξέων που συγκεντρώθηκε από διάφορους επιστήμονες σαν ένα αντιπροσωπευτικό παράδειγμα του συνόλου των γονιδίων σε έναν εξιδανικευμένο μεμονωμένο οργανισμό ενός είδους. Η αλληλουχία αναφοράς μπορεί να ληφθεί από τις βάσεις δεδομένων Ensembl (Howe et al., 2021) και UCSC Genome Browser (Lee et al., 2022). Στην παρούσα διπλωματική εργασία το γονιδίωμα αναφοράς (GRCh38.p13) που χρησιμοποιήθηκε είναι της βάσης δεδομένων του Ensembl.

Πριν πραγματοποιηθεί η διαδικασία χρειάζονται:

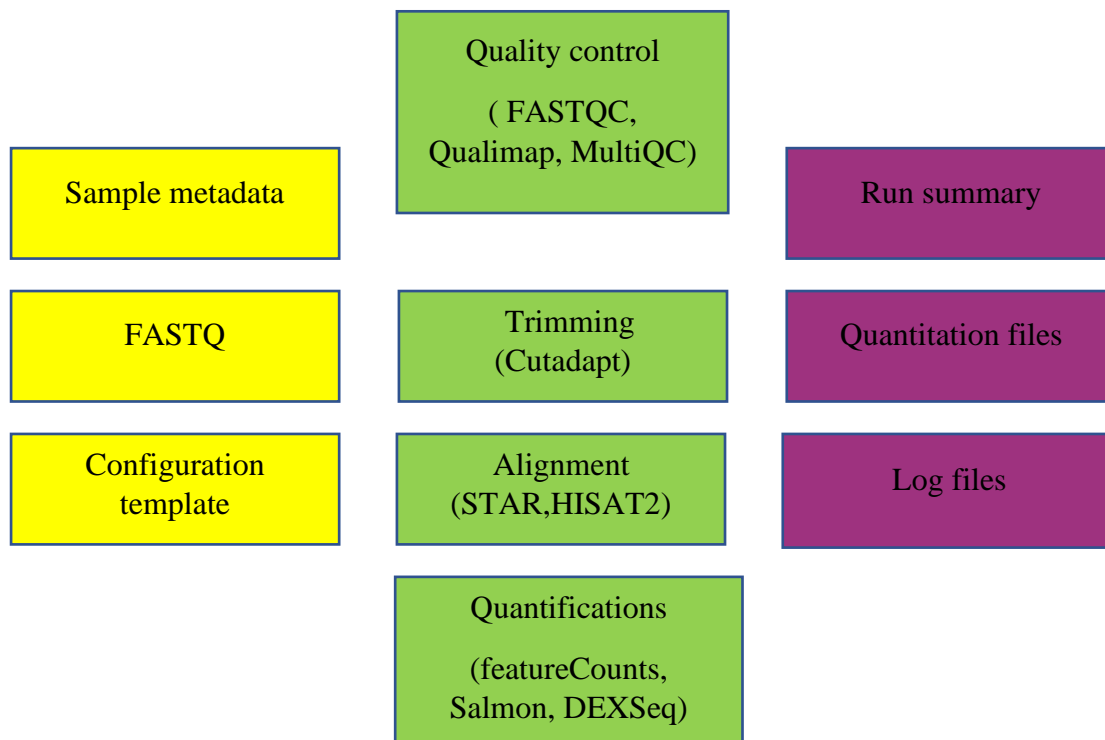
1) Ένα αρχείο σε μορφή FASTA που να περιέχει την πραγματική αλληλουχία γονιδιώματος/μεταγραφώματος και

2) Ένα αρχείο **GTF/GFF** (Reese et al., 2010) το οποίο θα περιέχει τον αντίστοιχο σχολιασμό. Δηλαδή, το αρχείο αυτό περιλαμβάνει πληροφορίες για κάθε αλληλουχία γονιδίου, μεταγράφου και εξωνίου που έχει ταυτοποιηθεί έως και σήμερα.

Η μορφοποίηση FASTA είναι μια μορφή βασισμένη σε κείμενο για την αναπαράσταση είτε αλληλουχιών νουκλεοτιδίων είτε αμινοξέων (πρωτεΐνες), στις οποίες τα νουκλεοτίδια ή τα αμινοξέα αναπαρίστανται χρησιμοποιώντας κωδικούς ενός γράμματος. Η μορφή επιτρέπει επίσης τα ονόματα και τα σχόλια να προηγούνται των αλληλουχιών. Υπάρχουν αρκετά εργαλεία που χρησιμοποιούνται για την στοίχιση των αναγνώσεων στην αλληλουχία αναφοράς ορισμένα από αυτά είναι τα: STAR (Dobin et al., 2013), HISAT2 (Kim et al., 2019), Tophat2 (Kim et al., 2013), Bowtie2 (Langmead and Salzberg, 2012) και GEM (Marco-Sola et al., 2012). Πρόκειται για μια σημαντική διαδικασία με την οποία τα αποτελέσματα ποσοτικοποιούνται. Επομένως μετά την χαρτογράφηση των αναγνώσεων μπορούμε να υπολογίσουμε και να γνωρίζουμε επακριβώς πόσες φορές «εντοπίστηκε» το κάθε νουκλεοτίδιο του υπό μελέτη γονιδιώματος. Η ποσοτικοποίηση των μετρήσεων που έχουν αντιστοιχηθεί σε κάθε γονίδιο υλοποιείται με τη χρήση του Salmon (Patro et al., 2017) και Kallisto (Bray et al., 2016). Τέλος, προκειμένου να υπολογίσουμε τις τιμές γονιδιακής έκφρασης θα πρέπει να προηγηθεί μια διαδικασία κανονικοποίησης με σκοπό να μετατρέψουμε τα δεδομένα ώστε να είναι συγκρίσιμα και να αφαιρέσουμε τυχόν συστηματικά σφάλματα που έχουν προέλθει με ποικίλους τρόπους από πολλαπλούς παράγοντες.

### **3.5.1 Workflow του Bcbio-nextgen**

Έχοντας εγκαταστήσει το bcbio-nextgen στο περιβάλλον του Linux με τα απαραίτητα εργαλεία και έχοντας φτιάξει του καταλόγους και αρχεία που απαιτούνται, τότε για να ξεκινήσει να τρέχει το πρόγραμμα θα πρέπει να ακολουθηθούν ορισμένα βήματα με τα κατάλληλα δεδομένα εισόδου προκειμένου να προχωρήσουμε και στην επεξεργασία των δεδομένων αυτών. Παρουσιάζεται σχηματικά το workflow του bcbio :



Ας περιγράψουμε αναλυτικά τα παραπάνω βήματα:

Το hcbio χρειάζεται ως είσοδο αρχικά τα FASTQ αρχεία (χωρίς να έχουν υποστεί κάποιου είδους φιλτραρίσματος) τα οποία είναι τα πρωτογενή δεδομένα που κατεβάζουμε από τη Βάση Δεδομένων ENA όπως αναλύθηκαν λεπτομερώς εκ των προτέρων. Εκτός από τα αρχεία αυτά χρειάζονται επίσης και τα .csv (comma separated value file) αρχεία τα οποία έχουμε δημιουργήσει για κάθε μια από τις μελέτες ξεχωριστά και περιλαμβάνουν τα μετα-δεδομένα (**sample metadata**). Τέλος, απαραίτητο είναι και ένα πρότυπο διαμόρφωσης (**configuration template**) το οποίο εμπεριέχει πληροφορίες σχετικά με το είδος της ανάλυσης που θα εκτελέσουμε. Το αρχείο αυτό χρησιμοποιείται σε συνδυασμό τόσο των csv αρχείων, όσο και των αρχείων FASTQ με σκοπό τελικά τη δημιουργία ενός τελικού προτύπου διαμόρφωσης. Ένα ενδεικτικό παράδειγμα (**Πίνακας 6**) ενός αρχείου .csv για την μελέτη PRJNA662780 φαίνεται στον πίνακα παρακάτω, το οποίο αποτελείται από τρεις στήλες.

samplename	description	panel
SRR12621861	IPD1	parkinson
SRR12621863	IPD2	parkinson
SRR12621864	IPD3	parkinson
SRR12621865	IPD4	parkinson
SRR12621866	IPD5	parkinson
SRR12621867	C1	control
SRR12621868	C2	control
SRR12621869	C3	control
SRR12621870	C4	control
SRR12621871	C5	control
SRR12621872	C6	control

Πίνακας 6 Το csv αρχείο για τη μελέτη PRJNA662780 με τα δείγματα που αυτό εμπεριείχε. Πέντε δείγματα παθολογικών ιστών και έξι δείγματα υγιών.

Η πρώτη στήλη (samplename) περιέχει την επικεφαλίδα (samplename) που αντιστοιχεί στα ονόματα αρχείων FASTAQ στα οποία πραγματοποιείται η ανάλυση.

Η δεύτερη στήλη (description) περιέχει την περιγραφή προκειμένου να μετατρέψουμε το όνομα του δείγματος που αρχικά παρέχεται από το όνομα του αρχείου σε αυτή την τιμή.

Η τρίτη στήλη (panel) παρέχει επιπρόσθετες πληροφορίες για κάθε δείγμα εάν δηλαδή προέρχεται από ασθενείς με τη νόσο ή υγιείς.

Εκτός από τα πρωτογενή FASTQ αρχεία και τα .csv αρχεία, το bcbio απαιτεί και ένα πρότυπο διαμόρφωσης το οποίο δίνει λεπτομέρειες σχετικά με τις επιλογές ανάλυσης. Το αρχείο αυτό χρησιμοποιείται σε συνδυασμό τόσο με τα .csv όσο και με τα FASTQ αρχεία με σκοπό τελικά τη δημιουργία ενός τελικού προτύπου διαμόρφωσης (**Πίνακας 7**).

```
# Template for human RNA-seq using Illumina prepared samples
---
details:
- analysis: RNA-seq
  genome_build: hg38
  algorithm:
    aligner: star
    strandedness: unstranded
    quantify_genome_alignments: true
upload:
  dir: ../final
```

*Πίνακας 7 Το πρότυπο διαμόρφωσης που χρησιμοποιήθηκε στην εργασία και ορίζει λεπτομέρειες για κάθε δείγμα προς επεξεργασία*

Ειδικότερα:

- **analysis:** Το είδος της ανάλυσης που εκτελούμε (λόγου χάρη, RNA-seq, ChIPseq, variant).
- **genome\_build:** Το γονιδίωμα αναφοράς που χρησιμοποιείται. Υπάρχει η δυνατότητα αναζήτησης των γονιδιωμάτων που είναι διαθέσιμα στο bcbio και πως αυτά μπορούν να καθοριστούν πληκτρολογώντας στο τερματικό την εντολή: `bcbio_setup_genome.py` το οποίο επιστρέφει μια λίστα όλων των τρεχόντων γονιδιωμάτων.
- **aligner:** Το εργαλείο που χρησιμοποιείται για την στοίχιση των αναγνώσεων στο γονιδίωμα αναφοράς.
- **upload:** Ορίζουμε το upload το οποίο είναι για το τελικό αποτέλεσμα από το bcbio.

Η βασική εντολή που τρέχουμε στο τερματικό για να τρέξει το bcbio είναι:

- **`bcbio_nextgen.py ../config/seqc.yaml -n 14 &`**

Όμως πριν εκτελέσουμε την ανάλυση θα πρέπει να αναφερθούμε και στα εργαλεία που θα χρησιμοποιηθούν προκειμένου να γίνει ο ποιοτικός έλεγχος, το φιλτράρισμα των προσαρμογών, η στοίχιση και η ποσοτικοποίηση σε επίπεδο γονιδίου.

Στο δικό μας πρότυπο διαμόρφωσης καθορίσαμε τα εξής:

algorithm:

aligner: star

strandedness: unstranded

Στη συγκεκριμένη ανάλυση δεν πραγματοποιείται φιλτράρισμα, καθώς το Star και το Salmon έχουν τη δυνατότητα να χειρίζονται τις αλληλουχίες των προσαρμογών με δικό τους τρόπο. Για το λόγο αυτόν δεν γίνεται καθόλου φιλτράρισμα. Υπάρχουν φορές που το φιλτράρισμα δημιουργεί πρόβλημα διότι σε ορισμένες περιπτώσεις μπορεί να κοπούν και χρήσιμες βάσεις, και άλλες φορές που είναι αρκετά βοηθητική στην ανάλυση. Επιπλέον, πραγματοποιείται ο ποιοτικός έλεγχος με σκοπό να συγκεντρωθούν τα αποτελέσματα σε μια έκθεση η οποία περιέχει χαρακτηριστικά των χαρτογραφημένων αναγνώσεων. Η καταμέτρηση των αναγνώσεων γίνεται με τη χρήση του featureCounts. Χρησιμοποιείται επίσης και το Salmon, το οποίο είναι μια εξαιρετικά γρήγορη μέθοδος ποσοτικοποίησης. Το βήμα αυτό είναι βασικό καθώς η καταμέτρηση απαιτείται για τη διεξαγωγή στατιστικών συμπερασμάτων στη διαφορική έκφραση. Τελικώς το output που παράγει είναι το αρχείο counts όπως αυτό δημιουργείται αυτόματα από το πρόγραμμα στο οποίο εμπεριέχετε η καταμέτρηση των αναγνώσεων που έχουν αντιστοιχηθεί σε κάθε γονίδιο.

### 3.5.2 Διαφορική έκφραση γονιδίων

Αφού έχει ολοκληρωθεί η αρχική ανάλυση των δεδομένων, επόμενη διαδικασία είναι η εκτίμηση της διαφορικής έκφρασης των γονιδίων που διαθέτουμε από τις 4 μελέτες: PRJNA283498, PRJNA557205, PRJNA563007 και PRJNA662780. Στην ανάλυση διαφορικής έκφρασης η σύγκριση πραγματοποιείται μεταξύ δυο καταστάσεων. Συγκεκριμένα για ένα σύνολο γονιδίων προσδιορίζονται τα γονίδια εκείνα τα οποία ενεργοποιούνται, δηλαδή υπερεκφράζονται, τα γονίδια εκείνα τα οποία υποεκφράζονται καθώς και τα γονίδια εκείνα των οποίων η έκφραση δεν αλλάζει ή με άλλα λόγια δεν εκφράζονται. Πριν προχωρήσουμε στην εκτίμηση της διαφορικής έκφρασης των γονιδίων θα πρέπει να προηγηθεί μια διαδικασία κανονικοποίησης με σκοπό να μετατρέψουμε τα δεδομένα σε μια κλίμακα ώστε να είναι συγκρίσιμα και να αφαιρέσουμε τυχόν σφάλματα που έχουν προκύψει. Ο πιο συνηθισμένος τρόπος κανονικοποίησης είναι η τυποποίηση z-score (σχέση 3.1)(Νικολάου and Χουβαρδάς, 2015). Υπάρχουν και άλλοι τρόποι κανονικοποίησης οπότε ανάλογα με τις απαιτήσεις του πειράματος επιλέγουμε αντίστοιχα.

$$Z(g) = \frac{E(g) - \mu}{\sigma} \quad (3.1)$$

Όπου,

g: το υπο μελέτη γονίδιο

Z(g): κανονικοποιημένη τιμή έκφρασης



E(g): πραγματική τιμή έκφρασης

μ: μέση τιμή του δείγματος

σ: τυπική απόκλιση

Το εργαλείο για τη διαφορική έκφραση που χρησιμοποιήθηκε στην παρούσα εργασία είναι το DESeq2 (Love et al., 2014). Για να υπολογίσουμε το βαθμό που μεταβάλλεται η έκφραση ενός γονιδίου μεταξύ των δύο καταστάσεων, υπολογίζεται ο λογάριθμος της συνθήκης μελέτης που στην προκειμένη περίπτωση είναι η παθολογική κατάσταση (test) δια της φυσιολογικής κατάστασης (control). Επομένως, απώτερος στόχος μιας ανάλυσης διαφορικής έκφρασης είναι να κατανοήσουμε κατά πόσο διαφορετικά εκφράζονται συγκεκριμένα γονίδια σε μια παθολογική κατάσταση, έναντι μιας φυσιολογικής. Η εκτίμηση του μεγέθους της επίδρασης υπολογίζεται από τον παρακάτω τύπο:

$$\log_2FC(g) = \frac{Eg(test)}{Eg(control)}, \text{ όπου } g \text{ το γονίδιο}$$

Η τιμή αυτή αποτυπώνει το πόσο φαίνεται να έχει αλλάξει η έκφραση του γονιδίου ή του μεταγράφου μεταξύ των δυο καταστάσεων. Η τιμή αυτή αναφέρεται σε λογαριθμική κλίμακα στη βάση 2.

Οι θετικές τιμές  $\log_2FC$  υποδεικνύουν την αύξηση της έκφρασης του γονιδίου  $g$  στη παθολογική κατάσταση, ενώ αρνητικές τιμές  $\log_2FC$  τη μείωση της έκφρασης του.

Στόχος μιας ανάλυσης διαφορικής έκφρασης είναι η παραγωγή μιας λίστας πολλών γονιδίων που έχουν υποστεί πολλαπλές δοκιμές και τα οποία έχουν ταξινομηθεί βάση του επιπέδου σημαντικότητας τους, p-value.

Για παράδειγμα ένα αρχείο αποτελεσμάτων του DESeq2 είναι ένα αρχείο csv το οποίο περιέχει τις εξής πληροφορίες (**Πίνακας 8**):

- rowname: Η πρώτη στήλη περιλαμβάνει το όνομα του κάθε γονιδίου στη μορφή της βάσης που χρησιμοποιήθηκε σε προηγούμενα βήματα (Ensembl, πχ. ENSG00000000005)
- baseMean: Η δεύτερη στήλη περιλαμβάνει το μέσο όρο των κανονικοποιημένων μετρήσεων που λαμβάνεται σε όλα τα δείγματα
- log2FoldChange: Η τρίτη στήλη αναφέρεται στην εκτίμηση του μεγέθους επίδρασης
- lfcSE: Η τέταρτη στήλη παρουσιάζει την εκτίμηση τυπικού σφάλματος για την εκτίμηση της λογαριθμικής αλλαγής  $\log_2$  φορές
- p-value: Η πέμπτη στήλη περιλαμβάνει την τιμή της δοκιμής για το συγκεκριμένο γονίδιο
- padj: Η έκτη στήλη περιλαμβάνει τις προσαρμοσμένες τιμές p-value για πολλαπλές δοκιμές για το γονίδιο

rowname	baseMean	log <sub>2</sub> FoldChange	lfcSE	pvalue	padj
ENSG00000000003	328.4101	0.166321	0.138427	0.150763	0.277565
ENSG00000000005	3.439692	-0.32405	0.310738	0.053323	0.128546
ENSG00000000419	554.4034	-0.24015	0.117952	0.020723	0.062366
ENSG00000000457	307.6646	-0.36033	0.094504	4.22E-05	0.000584
ENSG00000000460	115.5179	0.176526	0.13087	0.11348	0.225759
ENSG00000000938	235.1435	0.929073	0.198913	2.59E-07	1.69E-05
ENSG00000000971	749.4382	0.945554	0.288603	7.83E-05	0.00091
ENSG00000001036	401.7555	-0.00998	0.080249	0.894739	0.939079
ENSG00000001084	1223.339	-0.171	0.115874	0.09315	0.194981
ENSG00000001167	556.8368	-0.07083	0.089721	0.394555	0.552214
ENSG00000001460	606.1504	-0.18052	0.06586	0.004181	0.018317
ENSG00000001461	3266.809	-0.20949	0.128073	0.056166	0.133528
ENSG00000001497	1783.104	0.088187	0.083943	0.262155	0.415584
ENSG00000001561	1551.148	-0.23387	0.118975	0.025474	0.072909
ENSG00000001617	267.1973	0.931435	0.20748	5.90E-07	2.99E-05
ENSG00000001626	47.76537	-0.66084	0.312456	0.002327	0.011612
ENSG00000001629	1754.161	-0.08939	0.106666	0.357224	0.515585
ENSG00000001630	2482.541	-0.15229	0.09221	0.073092	0.162931
ENSG00000001631	937.485	-0.22994	0.060813	9.08E-05	0.001012
ENSG00000002016	444.8344	0.309115	0.156721	0.016676	0.052797
ENSG00000002079	1.087154	0.015782	0.231138	0.81533	0.886782
ENSG00000002330	1646.777	0.123812	0.084617	0.118001	0.232494
ENSG00000002549	1062.042	0.017902	0.104566	0.851187	0.910835
ENSG00000002586	1706.605	0.6709	0.185791	3.13E-05	0.000472

Πίνακας 8 Παράδειγμα εξόδου του DESeq2 συμπεριλαμβανομένων 24 σειρών από την ανάλυση διαφορικής έκφρασης γονιδίων στα δεδομένα της μελέτης PRJNA283498

Ένα από τα βασικά προβλήματα που προκύπτουν σε μια RNA-Seq ανάλυση στα οποία ελέγχεται μεγάλος αριθμός γονιδίων, είναι η εμφάνιση των εσφαλμένων ‘ανακαλύψεων’ (false discoveries), εξαιτίας των πολλαπλών δοκιμών που πραγματοποιούνται. Αυτό αντιμετωπίζεται με την εκτίμηση του FDR (False Discovery Rate), χρησιμοποιώντας τις διορθώσεις του Benjamini-Hochberg (Benjamini and Hochberg, 1995). Τα βήματα της μεθόδου είναι τα εξής:

- Κατάταξη των p-values από το μικρότερο προς το μεγαλύτερο
- Αρίθμηση των p-values
- Υπολογισμός των adjusted p-values

Για τον υπολογισμό του adj p-values χρησιμοποιούμε τον παρακάτω τύπο (σχέση 3.2):

$$\text{Current } p - \text{ values} * \left( \frac{\text{total \# of } p - \text{ values}}{p - \text{ value rank}} \right) \quad (3.2)$$

Για να γίνει κατανοητή η μέθοδος Benjamini-Hochberg θα παρουσιάσουμε ένα παράδειγμα το οποίο περιλαμβάνει 6 δείγματα και τα αντίστοιχα p-values αυτών.

Samples	P-values
A	0.91
B	0.11
C	0.71
D	0.31
E	0.01
F	0.45

Το αρχικό βήμα είναι να κατατάξουμε τα p-values από το μικρότερο προς το μεγαλύτερο, επομένως θα είναι:

Samples	P-values
E	0.01
B	0.11
D	0.31
F	0.45
C	0.71
A	0.91

Παρατηρούμε λοιπόν ότι το  $p\text{-value} = 0.01 < 0.05$  είναι ψευδώς θετικό. Στη συνέχεια κάνουμε την αρίθμηση (rank) αυτών από το 1-6.

Samples	Rank	P-values
E	1	0.01
B	2	0.11
D	3	0.31
F	4	0.45
C	5	0.71
A	6	0.91

Για τον υπολογισμό των adjusted p-values χρησιμοποιούμε την σχέση (3.2) και έτσι έπειτα από υπολογισμούς έχουμε:

Samples	Rank	P-values	Adj p-values
E	1	0.01	0.06
B	2	0.11	0.33
D	3	0.31	0.62
F	4	0.45	0.675
C	5	0.71	0.852
A	6	0.91	0.91

Παρατηρούμε πλέον ότι το ψευδώς στατιστικά σημαντικό p-value δεν είναι στατιστικά σημαντικό αφού η τιμή του  $0.06 > 0.05$ . Για το λόγο αυτόν, είναι σημαντική η διαδικασία καθώς απορρίπτονται ορισμένα p-values τα οποία στην αρχή θεωρούνταν

ως στατιστικά σημαντικά και απομένουν οι τιμές αυτών που είναι στατιστικά σημαντικά και θα μας οδηγήσουν σε πιο ακριβή αποτελέσματα.

### 3.5.3 Μετα-ανάλυση δεδομένων γονιδιακής έκφρασης

Έχοντας υλοποιήσει όλα τα παραπάνω βήματα και έχοντας διαθέσιμη τη λίστα με τα διαφορικά εκφρασμένα γονίδια, χρησιμοποιούμε τη μέθοδο της μετα-ανάλυσης για να βρούμε τα γονίδια εκείνα που εκφράζονται διαφορετικά μεταξύ των υπο μελέτη μελετών. Η μέθοδος της μετα-ανάλυσης συνδυάζει μη προσαρμοσμένες τιμές p-values της κάθε μελέτης για κάθε γονίδιο χρησιμοποιώντας μια έκδοση της μετα-ανάλυσης Stouffer (Stouffer et al., 1949), που προτάθηκε από τους Mosteller & Bush (Mosteller and Bush, 1954). Για κάθε γονίδιο και μελέτη γίνεται η μετατροπή του two tailed p-value σε one tailed p-value βάση του πρόσημου του αντίστοιχου  $\log_2FC$ . Για κάθε one tailed p-value το αντίστοιχο z-score υπολογίστηκε χρησιμοποιώντας τη συνάρτηση κανονικής κατανομής ( $\Phi^{-1}$ ). Το p-value για κάθε γονίδιο υπολογίζεται με βάση τον τύπο (3.3).

$$p = \Phi\left(\frac{\sum_{i=1}^k ni\Phi^{-1}(pi)}{\sqrt{\sum_{i=1}^k ni^2}}\right) \quad (3.3)$$

Όπου,  $p_i$  είναι η τιμή που προέρχεται από το DESeq2,  $n_i$  είναι ο αριθμός των δειγμάτων της μελέτης  $i$  και  $k$  ο αριθμός των μελετών. Τέλος οι τιμές  $p$  υποβλήθηκαν σε FDR και το 0.05 ορίστηκε ως κατώφλι στατιστικής σημαντικότητας.

### 3.5.4 Δίκτυο αλληλεπίδρασης σύμφωνα με τη βάση STRING και WebGestalt

Προκειμένου να γίνουν αντιληπτές οι διεργασίες και ο τρόπος λειτουργίας των κυττάρων είναι απαραίτητο να γνωρίζει κανείς το ρόλο των πρωτεϊνών, τις μεταξύ τους αλληλεπιδράσεις-συσχετίσεις και γενικότερα την πιθανή αλληλεπίδραση αυτών με πολλαπλή διαφορετικά μόρια του κυττάρου. Όπως αναφέρθηκε εκτενώς στα προηγούμενα κεφάλαια η πρωτεΐνη α-συνουκλειΐνη παίζει κεντρικό ρόλο στην εξέλιξη της νόσου του Πάρκινσον. Επομένως είναι σημαντικό να γνωρίζουμε τη συμπεριφορά των πρωτεϊνών τόσο σε κύτταρα υγιά όσο και σε ασθενή τα οποία σε συνδυασμό με διάφορες πειραματικές και κλινικές μελέτες και παρατηρήσεις θα συμβάλλουν τόσο στην αντίληψη των βλαβών που σχετίζονται είτε από τη γήρανση είτε από την εμφάνιση διαφόρων ασθενειών. Για να έχουμε μια πλήρη εικόνα των διαφορικά εκφρασμένων γονιδίων που προέκυψαν και την μεταξύ τους συσχέτιση, χρησιμοποιήθηκε η βάση δεδομένων STRING (<https://string-db.org/>). Η STRING (Search Tool for the Retrieval of Interacting Genes/proteins) είναι μια ολοκληρωμένη βάση δεδομένων η οποία αποσκοπεί την παροχή κριτικής αξιολόγησης και ολοκλήρωσης των αλληλεπιδράσεων μεταξύ γονιδίων/πρωτεϊνών οι οποίες προέρχονται από διαφορετικές πηγές, όπως λόγου χάρη γενετικές αναλύσεις, βιοχημικές, αναλύσεις συν έκφρασης κ.α.(Szkarczyk et al., 2015). Επιπλέον χρησιμοποιήθηκε και το διαδικτυακό εργαλείο WebGestalt (Zhang et al., 2005) (<http://www.webgestalt.org/>) προκειμένου να αποκτήσουμε μια συνολική εικόνα των λιστών των γονιδίων που υπερ εκφράζονται και υπο εκφράζονται διαφορικά με στατιστικά σημαντικό τρόπο. Το εργαλείο WebGestalt έχει σχεδιαστεί κυρίως για γονιδιωματικές, γονιδιακές, πρωτεομικές καθώς και γενετικές μελέτες υψηλής κλίμακας από τις οποίες δημιουργούνται σύνολα δεδομένων υψηλής απόδοσης, όπως

είναι για παράδειγμα τα διαφορετικά εκφρασμένα γονίδια, γονίδια που συν εκφράζονται κ.ο.κ.

#### 4. Αποτελέσματα

Όπως αναλύθηκε εκτενώς στα προηγούμενα κεφάλαια οι μελέτες που τελικά συμπεριλήφθηκαν στην παρούσα διπλωματική εργασία, γιατί πληρούσαν τα κριτήρια επιλεξιμότητας και περιείχαν δείγματα και δεδομένα για την ανάλυσή μας, είναι 4. Μια από τις μελέτες αυτές, η PRJNA563007, χωρίστηκε σε δυο υπομελέτες, την PRJNA563007\_Nigra και την PRJNA563007\_Putamen, γιατί υπάρχουν δείγματα τα οποία προέρχονται από διαφορετικό τμήμα του εγκεφάλου. Μετά τη συλλογή των δεδομένων μας, πραγματοποιήθηκε η καταγραφή των δειγμάτων της κάθε μελέτης ξεχωριστά. Συγκεκριμένα, η μελέτη PRJNA283498 έχει συνολικά 73 δείγματα, τα 29 προέρχονται από παθολογικούς ιστούς και τα 44 από φυσιολογικούς. Αντίστοιχα, η PRJNA557205 έχει συνολικά 36 δείγματα, 24 από παθολογικούς ιστούς και 12 από φυσιολογικούς, η υπομελέτη PRJNA563007\_Nigra έχει συνολικά 10 δείγματα, 5 από παθολογικούς και 5 φυσιολογικούς ιστούς, ενώ η υπομελέτη PRJNA563007\_Putamen έχει συνολικά 6 δείγματα, 3 παθολογικούς και 3 φυσιολογικούς και η μελέτη PRJNA662780 έχει στο σύνολο 11 δείγματα, 5 που προέρχονται από παθολογικούς και 6 από φυσιολογικούς ιστούς (Πίνακας 9).

Μελέτη	Τύπος κυττάρου	Παθολογικά δείγματα	Φυσιολογικά δείγματα	Σύνολο Δειγμάτων
PRJNA283498	BA9/ Προμετωπιαίος φλοιός	29	44	73
PRJNA557205	Προμετωπιαίος φλοιός	24	12	36
PRJNA563007 NIGRA	Μέλαινα ουσία	5	5	10
PRJNA563007 PUTAMEN	Κέλυφος	3	3	6
PRJNA662780	Μεσεγκέφαλος	5	6	11

Πίνακας 9 Χαρακτηριστικά των υπό μελέτη μελετών που χρησιμοποιήθηκαν στη μετα-ανάλυση

Πιο συγκεκριμένα, το εργαλείο DESeq2 χρησιμοποιήθηκε για τον υπολογισμό της διαφορικής έκφρασης σε κάθε μελέτη (Πίνακας 10).

Μελέτη	Σύνολο γονιδίων	Στατιστικά σημαντικά	Υπερ εκφρασμένα	Υπο εκφρασμένα
PRJNA283498	35143	8635	4138	4497
PRJNA557205	42041	1503	924	579
PRJNA563007 Nigra	28884	177	89	88
PRJNA563007 Putamen	26777	8	5	3
PRJNA662780	21706	26	12	14

Πίνακας 10 Συγκεντρωτικός πίνακας αποτελεσμάτων βάση των υπό μελέτη μελετών

Η μελέτη PRJNA283498 κατέγραψε συνολικά **35143** γονίδια. Ελέγχοντας την τιμή του p-value του κάθε γονιδίου, αφού εφαρμόσαμε την τροποποίηση FDR, βρέθηκαν συνολικά **8635** γονίδια τα οποία είναι στατιστικώς σημαντικά διαφορικά εκφρασμένα. Στα αποτελέσματα που βρέθηκαν, ελέγχοντας επίσης και τη στήλη που απεικονίζεται

το  $\log_2FC$  για κάθε γονίδιο, εντοπίζουμε τα γονίδια εκείνα που υπερεκφράζονται και υποεκφράζονται βάσει του προσήμου: Θετικές τιμές  $\log_2FC$  υποδηλώνουν αύξηση της έκφρασης του εκάστοτε γονιδίου στην παθολογική κατάσταση, ενώ αρνητικές τιμές  $\log_2FC$  υποδηλώνουν μείωση της έκφρασής του. Από τα 8635 γονίδια τα οποία είναι διαφορετικά εκφρασμένα, βρέθηκαν **4497** γονίδια που υποεκφράζονται και **4138** που υπερεκφράζονται (**Πίνακας 11**).

rowname	baseMean	log2FoldChange	lfcSE	pvalue	padj
ENSG00000086570	34.99029	-6.38899	0.347963	5.99E-05	0.000754
ENSG00000145863	24.57235	-3.70856	0.267582	0.000293	0.002432
ENSG00000174788	1.332471	-3.41872	0.497536	0.005952	0.023838
ENSG00000110887	11.37489	-2.67437	0.448123	0.012493	0.042185
ENSG00000181195	104.1987	-2.51258	0.317252	0.000683	0.004574
ENSG00000102924	178.0918	-2.33711	0.165103	0.010274	0.036288
ENSG00000261499	4.304646	-2.2942	0.860193	0.000216	0.001936
ENSG00000223669	1.525649	-1.90907	0.809544	0.000897	0.005625
ENSG00000174948	6.056105	-1.87793	0.476051	3.73E-06	0.000106
ENSG00000217455	1.173008	-1.81339	0.824952	0.001492	0.008247
ENSG00000224566	1.497898	-1.8034	0.671236	0.000336	0.002686
ENSG00000259318	1.373368	-1.78407	0.786626	0.00091	0.005688
ENSG00000219881	3.800937	-1.76205	0.586444	9.90E-05	0.001077
ENSG00000211896	26.82495	2.554933	0.43669	0.001607	0.008749
ENSG00000204388	20596.72	2.575473	0.32018	6.06E-17	5.58E-13
ENSG00000149257	1035.745	2.611772	0.307954	1.62E-18	4.47E-14
ENSG00000240032	2.260954	2.61537	0.606381	8.50E-07	3.79E-05
ENSG00000124107	48.71624	2.646905	0.481644	1.20E-05	0.000239
ENSG00000226777	1.465166	2.679963	0.752721	0.001002	0.006125
ENSG00000275405	3.73748	2.933677	1.647234	0.004855	0.02056
ENSG00000204389	29251.57	2.948632	0.355646	8.67E-18	1.20E-13
ENSG00000225217	91.12602	3.119707	0.342922	6.13E-10	2.61E-07
ENSG00000204936	27.55515	3.383007	0.477489	0.000951	0.005887
ENSG00000224259	1.502816	3.5643	0.986145	0.002958	0.013958
ENSG00000226673	4.062052	4.264211	0.69898	7.52E-11	4.83E-08
ENSG00000173110	153.1006	5.715856	0.348195	2.06E-13	6.34E-10

*Πίνακας 11 Παρουσίαση ενός τμήματος των διαφορικά εκφρασμένων και στατιστικά σημαντικών γονιδίων της μελέτης PRJNA283498. Ελέγχοντας το πρόσημο του  $\log_2FC$  συμπεραίνουμε εάν το γονίδιο αυτό υποεκφράζεται ή υπερεκφράζεται*

Αντίστοιχα, ελέγχοντας τα αποτελέσματα της κάθε μελέτης ξεχωριστά καταλήγουμε στα εξής:

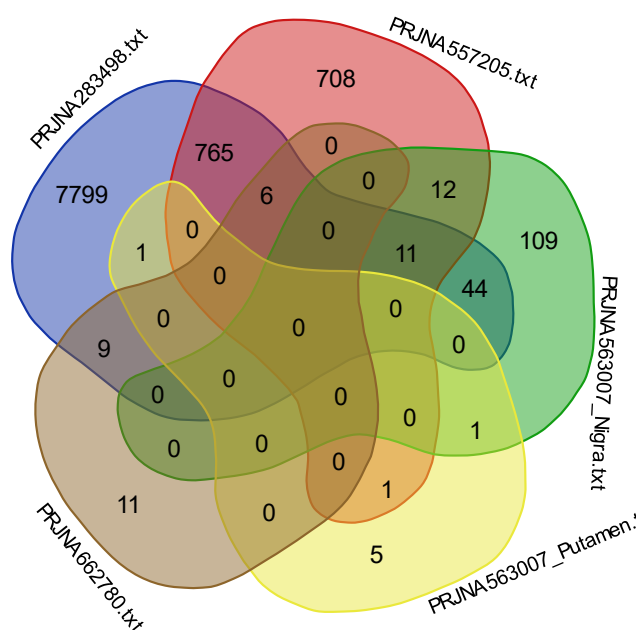
Η δεύτερη μελέτη, η PRJNA557205, κατέγραψε συνολικά **42041** γονίδια εκ των οποίων διαφορικά εκφρασμένα είναι τα **1503**, από τα οποία τα **579** υποεκφράζονται και τα **924** γονίδια υπερεκφράζονται.

Η υπομελέτη PRJNA563007\_Nigra ανέλυσε συνολικά **28884** γονίδια εκ των οποίων τα **177** ήταν διαφορικά εκφρασμένα γονίδια: Τα **88** υποεκφράζονται, ενώ τα **89** υπερεκφράζονται.

Η υπομελέτη PRJNA563007\_Putamen ταυτοποίησε συνολικά **26777** γονίδια αλλά μόνο τα **8** από αυτά ήταν διαφορετικά εκφρασμένα, με **3** από αυτά να υποεκφράζονται και **5** να υπερεκφράζονται.

Τέλος, η μελέτη PRJNA662780 κατέγραψε **21706** συνολικά γονίδια με μόνο 26 από αυτά να είναι διαφορετικά εκφρασμένα. Από τα 26, τα **14** υποεκφράζονται και τα υπόλοιπα **12** υπερεκφράζονται.

Τα διαγράμματα Venn είναι σημαντικά εργαλεία τα οποία χρησιμοποιούνται ευρέως σε διάφορους ερευνητικούς τομείς για τη γραφική απεικόνιση των ενώσεων, των τομών και των διακρίσεων μεταξύ πολλαπλών συνόλων δεδομένων με απώτερο στόχο την οπτικοποίηση των συνόλων αυτών και την ανάλυση των σχέσεών τους. Μια σύγκριση μεταξύ των στατιστικά σημαντικώς διαφορικώς εκφρασμένων γονιδίων τα οποία βρέθηκαν από κάθε μελέτη παρουσιάζονται στο διάγραμμα Venn. Κάθε μια από τις μελέτες χρωματίζεται με διαφορετικό χρώμα. 11 γονίδια βρέθηκαν κοινά μεταξύ των μελετών PRJNA283498, PRJNA557205 και PRJNA563007\_Nigra. Μεταξύ των μελετών PRJNA283498, PRJNA557205 και PRJNA662780 κοινά γονίδια ήταν μόλις 6. Ωστόσο μεταξύ των δυο μελετών οι οποίες έχουν πολλά διαφορετικά εκφρασμένα γονίδια, δηλαδή των μελετών PRJNA283498 και PRJNA557205 βρέθηκαν 765 κοινά γονίδια εξαιτίας του πλήθους των δειγμάτων στις μελέτες αυτές. Αξιοσημείωτο είναι ότι με βάση το διάγραμμα Venn δεν υπάρχουν κοινά γονίδια σε όλες ή έστω σε 4 εκ των μελετών (**Εικόνα 11**).



Εικόνα 11 Διάγραμμα Venn των διαφορετικά εκφρασμένων και στατιστικά σημαντικών γονιδίων σε κάθε μελέτη

Στη συνέχεια χρησιμοποιείται η μέθοδος της μετα-ανάλυσης προκειμένου να αυξηθεί η στατιστική ισχύ για την ανίχνευση γονιδίων που οδηγούν τελικά στην εμφάνιση του φαινοτύπου του Πάρκινσον και μπορούν να χρησιμοποιηθούν ως βιοδείκτες διαγνωστικής, προγνωστικής ή φαρμακολογικής σημασίας. Η πρώτη μετα-ανάλυση χρησιμοποιώντας το σύνολο των 5 μελετών ταυτοποίησε συνολικά 43709 γονίδια. Εξαιτίας της ύπαρξης των πολλαπλών δειγμάτων σε δύο από τις συμπεριλαμβανόμενες

μελέτες αυτό οδήγησε σε περαιτέρω προσαρμογές. Για να προχωρήσουμε σε δεύτερη μετα-ανάλυση λόγω των προηγούμενων μεροληπτικών αποτελεσμάτων, θεωρήσαμε ότι ένα γονίδιο δεν μπορεί να θεωρείται εκφρασμένο σε μια μελέτη (άρα και διαφορεικά εκφρασμένο σε αυτήν) εάν η μέση έκφρασή του μεταξύ όλων των δειγμάτων (baseMean) ανήκει στο κατώτατο 10% των μέσων εκφράσεων της μελέτης αυτής. Επίσης, τα γονίδια που έχουν  $|\log_2FC| < 1$  θεωρήθηκαν μη διαφορεικά εκφρασμένα. Ως εκ τούτου, για όλα αυτά τα γονίδια κάθε μελέτης ορίσαμε ανεξάρτητα από τις υπολογισμένες από το DESeq2 τιμές τους, το  $\log_2FC=0$  και το  $p\text{-value}=1$ . Έτσι, με τη μέθοδο της μετα-ανάλυσης ταυτοποιήθηκαν τελικά **339** διαφορεικά εκφρασμένα γονίδια από τα οποία τα **269** είναι υπερεκφρασμένα γονίδια (**Πίνακας 12**) και τα **70** είναι υποεκφρασμένα (**Πίνακας 13**). Στη συνέχεια θέλοντας να αντλήσουμε πληροφορίες για το κάθε γονίδιο ξεχωριστά χρησιμοποιούμε τη **Biomart** του Ensembl (Smedley et al., 2009). Η Biomart δίνει τη δυνατότητα στους χρήστες να πραγματοποιούν προηγμένες αναζητήσεις πηγών βιολογικών δεδομένων μέσω μιας ενιαίας διεπαφής ιστού. Η πρώτη στήλη περιγράφει τον κωδικό του κάθε γονιδίου στη βάση Ensembl, η δεύτερη στήλη περιγράφει το διεθνές αναγνωρισμένο όνομα του γονιδίου και η τρίτη στήλη τη περιγραφή του κάθε γονιδίου. Παρατηρώντας λοιπόν και τους δυο αυτούς πίνακες, βλέπουμε ότι πολλά από τα γονίδια είναι long non-coding RNAs (lncRNA), δηλαδή γονίδια τα οποία μεταγράφονται αλλά τελικά δεν μεταφράζονται σε πρωτεΐνες. Επίσης, πολλά από τα γονίδια που έχουμε καταγράψει δεν έχουν χαρακτηριστεί στο παρελθόν επομένως είναι άγνωστη ακόμη η λειτουργία τους. Επιπλέον, άλλα γονίδια είναι αντιπληροφορικά (antisense). Ένα αντιπληροφορικό γονίδιο ενός πληροφοριακού (κωδικοποιητικού ή μη κωδικοποιητικού) γονιδίου μεταγράφεται από τον αντίθετο κλώνο από τον κλώνο από τον οποίο μεταγράφεται το πληροφοριακό γονίδιο. Τα αντιπληροφορικά γονίδια εμπλέκονται κυρίως στη ρύθμιση της γονιδιακής έκφρασης, παρότι μπορούν να εκπληρώσουν διάφορους ρόλους (Pelechano and Steinmetz, 2013). Βλέπουμε επίσης ότι πολλά γονίδια χαρακτηρίζονται ως ψευδογονίδια (pseudogene), δηλαδή μη λειτουργικές αλληλουχίες DNA οι οποίες μοιάζουν με διάφορες αλληλουχίες γνωστών γονιδίων. Συμπεραίνουμε έτσι ότι πολλά από τα γονίδια παίζουν κυρίως ρυθμιστικό ρόλο.

Gene stable ID	Gene name	Gene description
ENSG00000006128	TAC1	tachykinin precursor 1
ENSG00000072315	TRPC5	transient receptor potential cation channel subfamily C member 5
ENSG00000086717	PPEF1	protein phosphatase with EF-hand domain 1
ENSG00000104327	CALB1	calbindin 1
ENSG00000116652	DLEU2L	deleted in lymphocytic leukemia 2 like
ENSG00000116981	NT5C1A	5'-nucleotidase, cytosolic 1A
ENSG00000119973	PRLHR	prolactin releasing hormone receptor
ENSG00000120875	DUSP4	dual specificity phosphatase 4
ENSG00000122025	FLT3	fms related receptor tyrosine kinase 3
ENSG00000122548	<b>KIAA0087</b>	KIAA0087 lncRNA
ENSG00000124678	TCP11	t-complex 11
ENSG00000128564	VGF	VGF nerve growth factor inducible
ENSG00000135824	RGS8	regulator of G protein signaling 8
ENSG00000140057	AK7	adenylate kinase 7
ENSG00000141433	ADCYAP1	adenylate cyclase activating polypeptide 1



ENSG00000145681	HAPLN1	hyaluronan and proteoglycan link protein 1
ENSG00000147571	CRH	corticotropin releasing hormone
ENSG00000152495	CAMK4	calcium/calmodulin dependent protein kinase IV
ENSG00000157005	SST	somatostatin
ENSG00000162913	<b>OBSCN-AS1</b>	OBSCN antisense RNA 1
ENSG00000164616	<b>FBXL21P</b>	F-box and leucine rich repeat protein 21, pseudogene
ENSG00000168843	FSTL5	follistatin like 5
ENSG00000169213	<b>RAB3B</b>	RAB3B, member RAS oncogene family
ENSG00000170160	CCDC144A	coiled-coil domain containing 144A
ENSG00000170629	<b>DPY19L2P2</b>	DPY19L2 pseudogene 2
ENSG00000174576	NPAS4	neuronal PAS domain protein 4
ENSG00000174948	GPR149	G protein-coupled receptor 149
ENSG00000175426	PCSK1	proprotein convertase subtilisin/kexin type 1
ENSG00000176697	BDNF	brain derived neurotrophic factor
ENSG00000180440	SERTM1	serine rich and transmembrane domain containing 1
ENSG00000181800	<b>CELF2-AS1</b>	CELF2 antisense RNA 1
ENSG00000184611	KCNH7	potassium voltage-gated channel subfamily H member 7
ENSG00000185008	ROBO2	roundabout guidance receptor 2
ENSG00000186479	RGS7BP	regulator of G protein signaling 7 binding protein
ENSG00000188211	NCR3LG1	natural killer cell cytotoxicity receptor 3 ligand 1
ENSG00000188263	IL17REL	interleukin 17 receptor E like
ENSG00000188517	COL25A1	collagen type XXV alpha 1 chain
ENSG00000188729	OSTN	ostecrin
ENSG00000188730	VWC2	von Willebrand factor C domain containing 2
ENSG00000189238	<b>LINC00943</b>	long intergenic non-protein coding RNA 943
ENSG00000196549	MME	membrane metalloendopeptidase
ENSG00000197705	KLHL14	kelch like family member 14
ENSG00000197882	<b>OR7E13P</b>	olfactory receptor family 7 subfamily E member 13 pseudogene
ENSG00000203668	CHML	CHM like Rab escort protein
ENSG00000205488	<b>CALML3-AS1</b>	CALML3 antisense RNA 1
ENSG00000231752	<b>EMBP1</b>	embigin pseudogene 1
ENSG00000236714	LINC01844	long intergenic non-protein coding RNA 1844
ENSG00000236806	RPL7AP15	ribosomal protein L7a pseudogene 15 Symbol;Acc:HGNC:36873]
ENSG00000237720		novel transcript
ENSG00000245384	CXXC4-AS1	CXXC4 antisense RNA 1
ENSG00000248050	SYBU-AS1	SYBU antisense RNA 1
ENSG00000250305	TRMT9B	tRNA methyltransferase 9B (putative)
ENSG00000253649	PRSS51	serine protease 51
ENSG00000253678		<b>novel transcript</b>
ENSG00000257647		novel transcript
ENSG00000257941		density-regulated protein (DENR) pseudogene
ENSG00000266405	CBX3P2	chromobox 3 pseudogene 2
ENSG00000275740		novel readthrough transcript

ENSG00000276672		novel transcript, sense intronic to RFC3
ENSG00000278962		TEC
ENSG00000279513		TEC
ENSG00000279952		novel transcript
ENSG00000280234		TEC
ENSG00000281183	NPTN-IT1	NPTN intronic transcript 1
ENSG00000281852	LINC00891	long intergenic non-protein coding RNA 891
ENSG00000285601		novel transcript
ENSG00000285735	LINC02717	long intergenic non-protein coding RNA 2717
ENSG00000285844	KCNQ5-DT	KCNQ5 divergent transcript
ENSG00000285918		novel transcript, antisense to WWOX
ENSG00000285940		novel transcript, antisense to CELF4

Πίνακας 12 Αναλυτική περιγραφή των γονιδίων που υπο εκφράζονται διαφορεικά με στατιστικά σημαντικό τρόπο με τη χρήση του Biomart

Gene stable ID	Gene name	Gene description
ENSG00000006327	TNFRSF12A	TNF receptor superfamily member 12A
ENSG00000010327	STAB1	stabilin 1
ENSG00000011422	PLAUR	plasminogen activator, urokinase receptor
ENSG00000015520	NPC1L1	NPC1 like intracellular cholesterol transporter 1
ENSG00000018280	SLC11A1	solute carrier family 11 member 1
ENSG00000021488	SLC7A9	solute carrier family 7 member 9
ENSG00000042062	RIPOR3	RIPOR family member 3
ENSG00000042304	C2orf83	chromosome 2 open reading frame 83
ENSG00000070495	JMJD6	jumonji domain containing 6, arginine demethylase and lysine hydroxylase
ENSG00000072682	P4HA2	prolyl 4-hydroxylase subunit alpha 2
ENSG00000072694	FCGR2B	Fc gamma receptor IIb
ENSG00000077150	NFKB2	nuclear factor kappa B subunit 2
ENSG00000077238	IL4R	interleukin 4 receptor
ENSG00000087510	TFAP2C	transcription factor AP-2 gamma
ENSG00000088836	SLC4A11	solute carrier family 4 member 11
ENSG00000089327	FXYD5	FXYD domain containing ion transport regulator 5
ENSG00000091879	ANGPT2	angiopoietin 2
ENSG00000092607	TBX15	T-box transcription factor 15
ENSG00000099875	MKNK2	MAPK interacting serine/threonine kinase 2
ENSG00000100292	HMOX1	heme oxygenase 1
ENSG00000100985	MMP9	matrix metalloproteinase 9
ENSG00000101187	SLCO4A1	solute carrier organic anion transporter family member 4A1
ENSG00000101276	SLC52A3	solute carrier family 52 member 3
ENSG00000101916	TLR8	toll like receptor 8
ENSG00000102287	GABRE	gamma-aminobutyric acid type A receptor subunit epsilon
ENSG00000102359	SRPX2	sushi repeat containing protein X-linked 2
ENSG00000102886	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3
ENSG00000104415	CCN4	cellular communication network factor 4

ENSG00000104894	CD37	CD37 molecule
ENSG00000104972	LILRB1	leukocyte immunoglobulin like receptor B1
ENSG00000105366	SIGLEC8	sialic acid binding Ig like lectin 8
ENSG00000105668	UPK1A	uroplakin 1A
ENSG00000106211	HSPB1	heat shock protein family B (small) member 1
ENSG00000106397	PLOD3	procollagen-lysine,2-oxoglutarate 5-dioxygenase 3
ENSG00000106404	CLDN15	claudin 15
ENSG00000110077	MS4A6A	membrane spanning 4-domains A6A
ENSG00000110719	TCIRG1	T cell immune regulator 1, ATPase H+ transporting V0 subunit a3
ENSG00000110852	CLEC2B	C-type lectin domain family 2 member B
ENSG00000111057	KRT18	keratin 18
ENSG00000111087	GLI1	GLI family zinc finger 1
ENSG00000111181	SLC6A12	solute carrier family 6 member 12
ENSG00000112299	VNN1	vanin 1
ENSG00000112559	MDFI	MyoD family inhibitor
ENSG00000112936	C7	complement C7
ENSG00000114124	GRK7	G protein-coupled receptor kinase 7
ENSG00000115596	WNT6	Wnt family member 6
ENSG00000116096	SPR	sepiapterin reductase
ENSG00000117601	SERPINC1	serpin family C member 1
ENSG00000121966	CXCR4	C-X-C motif chemokine receptor 4
ENSG00000123405	NFE2	nuclear factor, erythroid 2
ENSG00000123689	G0S2	G0/G1 switch 2
ENSG00000124107	SLPI	secretory leukocyte peptidase inhibitor
ENSG00000124253	PCK1	phosphoenolpyruvate carboxykinase 1
ENSG00000124440	HIF3A	hypoxia inducible factor 3 subunit alpha
ENSG00000125089	SH3TC1	SH3 domain and tetratricopeptide repeats 1
ENSG00000125148	MT2A	metallothionein 2A
ENSG00000125531	FNDC11	fibronectin type III domain containing 11
ENSG00000125733	TRIP10	thyroid hormone receptor interactor 10
ENSG00000125735	TNFSF14	TNF superfamily member 14
ENSG00000125878	TCF15	transcription factor 15
ENSG00000127530	OR7C1	olfactory receptor family 7 subfamily C member 1
ENSG00000129437	KLK14	kallikrein related peptidase 14
ENSG00000129667	RHBDF2	rhomoid 5 homolog 2
ENSG00000130600	H19	H19 imprinted maternally expressed transcript
ENSG00000132002	DNAJB1	DnaJ heat shock protein family (Hsp40) member B1
ENSG00000134061	CD180	CD180 molecule
ENSG00000135736	CCDC102A	coiled-coil domain containing 102A
ENSG00000138964	PARVG	parvin gamma
ENSG00000139914	FITM1	fat storage inducing transmembrane protein 1
ENSG00000142227	EMP3	epithelial membrane protein 3
ENSG00000142627	EPHA2	EPH receptor A2
ENSG00000143127	ITGA10	integrin subunit alpha 10
ENSG00000143199	ADCY10	adenylate cyclase 10
ENSG00000144837	PLA1A	phospholipase A1 member A

ENSG00000145040	UCN2	urocortin 2
ENSG00000145832	SLC25A48	solute carrier family 25 member 48
ENSG00000148426	PROSER2	proline and serine rich 2
ENSG00000148773	MKI67	marker of proliferation Ki-67
ENSG00000148926	ADM	adrenomedullin
ENSG00000149257	SERPINH1	serpin family H member 1
ENSG00000149451	ADAM33	ADAM metallopeptidase domain 33
ENSG00000149573	MPZL2	myelin protein zero like 2
ENSG00000149781	FERMT3	FERM domain containing kindlin 3
ENSG00000149922	TBX6	T-box transcription factor 6
ENSG00000150782	IL18	interleukin 18
ENSG00000151650	VENTX	VENT homeobox
ENSG00000151929	BAG3	BAG cochaperone 3
ENSG00000152049	KCNE4	potassium voltage-gated channel subfamily E regulatory subunit 4
ENSG00000152779	SLC16A12	solute carrier family 16 member 12
ENSG00000153404	PLEKHG4B	pleckstrin homology and RhoGEF domain containing G4B
ENSG00000157551	KCNJ15	potassium inwardly rectifying channel subfamily J member 15
ENSG00000157833	GAREM2	GRB2 associated regulator of MAPK1 subtype 2
ENSG00000159167	STC1	stanniocalcin 1
ENSG00000160447	PKN3	protein kinase N3
ENSG00000162383	SLC1A7	solute carrier family 1 member 7
ENSG00000163141	BNIP1	BCL2 interacting protein like
ENSG00000163221	S100A12	S100 calcium binding protein A12
ENSG00000163501	IHH	Indian hedgehog signaling molecule
ENSG00000163638	ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif 9
ENSG00000163874	ZC3H12A	zinc finger CCCH-type containing 12A
ENSG00000164761	TNFRSF11B	TNF receptor superfamily member 11b
ENSG00000164867	NOS3	nitric oxide synthase 3
ENSG00000166523	CLEC4E	C-type lectin domain family 4 member E
ENSG00000166592	RRAD	RRAD, Ras related glycolysis inhibitor and calcium channel regulator
ENSG00000166825	ANPEP	alanyl aminopeptidase, membrane
ENSG00000167676	PLIN4	perilipin 4
ENSG00000167711	SERPINF2	serpin family F member 2
ENSG00000167895	TMC8	transmembrane channel like 8
ENSG00000168209	DDIT4	DNA damage inducible transcript 4
ENSG00000168404	MLKL	mixed lineage kinase domain like pseudokinase
ENSG00000168874	ATOH8	atonal bHLH transcription factor 8
ENSG00000169248	CXCL11	C-X-C motif chemokine ligand 11
ENSG00000169397	RNASE3	ribonuclease A family member 3
ENSG00000171790	SLFNL1	schlafen like 1
ENSG00000172037	LAMB2	laminin subunit beta 2
ENSG00000172578	KLHL6	kelch like family member 6
ENSG00000173040	EVC2	EvC ciliary complex subunit 2

ENSG00000173110	HSPA6	heat shock protein family A (Hsp70) member 6
ENSG00000173372	C1QA	complement C1q A chain
ENSG00000173530	TNFRSF10D	TNF receptor superfamily member 10d
ENSG00000173976	RAX2	retina and anterior neural fold homeobox 2
ENSG00000173991	TCAP	titin-cap
ENSG00000175832	ETV4	ETS variant transcription factor 4
ENSG00000176046	NUPR1	nuclear protein 1, transcriptional regulator
ENSG00000177045	SIX5	SIX homeobox 5
ENSG00000177234	LINC01561	long intergenic non-protein coding RNA 1561
ENSG00000177464	GPR4	G protein-coupled receptor 4
ENSG00000178343	SHISA3	shisa family member 3
ENSG00000178602	OTOS	otospiralin
ENSG00000179528	LBX2	ladybird homeobox 2
ENSG00000180229	HERC2P3	HERC2 pseudogene 3
ENSG00000180447	GAS1	growth arrest specific 1
ENSG00000181016	LSMEM1	leucine rich single-pass membrane protein 1
ENSG00000181885	CLDN7	claudin 7
ENSG00000182162	P2RY8	P2Y receptor family member 8
ENSG00000183615	FAM167B	family with sequence similarity 167 member B
ENSG00000183963	SMTN	smoothelin
ENSG00000184371	CSF1	colony stimulating factor 1
ENSG00000184414	IRS3P	insulin receptor substrate 3, pseudogene
ENSG00000185022	MAFF	MAF bZIP transcription factor F
ENSG00000185291	IL3RA	interleukin 3 receptor subunit alpha
ENSG00000185507	IRF7	interferon regulatory factor 7
ENSG00000185633	NDUFA4L2	NDUFA4 mitochondrial complex associated like 2
ENSG00000185955	SPACDR	sperm acrosome developmental regulator
ENSG00000187513	GJA4	gap junction protein alpha 4
ENSG00000187554	TLR5	toll like receptor 5
ENSG00000187634	SAMD11	sterile alpha motif domain containing 11
ENSG00000187800	PEAR1	platelet endothelial aggregation receptor 1
ENSG00000187922	LCN10	lipocalin 10
ENSG00000188379	IFNA2	interferon alpha 2
ENSG00000188425	NANOS2	nanos C2HC-type zinc finger 2
ENSG00000188488	SERPINA5	serpin family A member 5
ENSG00000188511	MIR3667HG	MIR3667 host gene
ENSG00000188662	H1-9P	H1.9 linker histone, pseudogene
ENSG00000188959	C9orf152	chromosome 9 open reading frame 152
ENSG00000196126	HLA-DRB1	major histocompatibility complex, class II, DR beta 1
ENSG00000196154	S100A4	S100 calcium binding protein A4
ENSG00000198113	TOR4A	torsin family 4 member A
ENSG00000198417	MT1F	metallothionein 1F
ENSG00000198695	MT-ND6	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 6
ENSG00000204388	HSPA1B	heat shock protein family A (Hsp70) member 1B

ENSG00000204389	HSPA1A	heat shock protein family A (Hsp70) member 1A
ENSG00000204839	MROH6	maestro heat like repeat family member 6
ENSG00000205307	SAP25	Sin3A associated protein 25
ENSG00000211892	IGHG4	immunoglobulin heavy constant gamma 4 (G4m marker)
ENSG00000213060		ribosomal protein S15 (RPS15) pseudogene
ENSG00000213445	SIPA1	signal-induced proliferation-associated 1
ENSG00000213654	GPSM3	G protein signaling modulator 3
ENSG00000213700	RPL17P50	ribosomal protein L17 pseudogene 50
ENSG00000215915	ATAD3C	ATPase family AAA domain containing 3C
ENSG00000217078		pseudogene similar to part of eukaryotic translation initiation factor 2 subunit 2 EIF2S2
ENSG00000219186	FTH1P19	ferritin heavy chain 1 pseudogene 19
ENSG00000225217	HSPA7	heat shock protein family A (Hsp70) member 7 (pseudogene)
ENSG00000225693	LAGE3P1	<b>L antigen family member 3 pseudogene 1</b>
ENSG00000226673	LINC01108	long intergenic non-protein coding RNA 1108
ENSG00000227764	LINC01693	long intergenic non-protein coding RNA 1693
ENSG00000228058	LINC01736	long intergenic non-protein coding RNA 1736
ENSG00000228275	ARMCX3-AS1	ARMCX3 antisense RNA 1
ENSG00000228459	LINC01546	long intergenic non-protein coding RNA 1546
ENSG00000230528	NOS2P3	nitric oxide synthase 2 pseudogene 3
ENSG00000231304	SGO1-AS1	SGO1 antisense RNA 1
ENSG00000231355		novel transcript
ENSG00000233232	NPIPB7	nuclear pore complex interacting protein family member B7
ENSG00000234584		<b>novel transcript</b>
ENSG00000234776	C11orf94	chromosome 11 open reading frame 94
ENSG00000235111		novel transcript
ENSG00000236423	LINC01134	long intergenic non-protein coding RNA 1134
ENSG00000236498		novel transcript
ENSG00000236710		proteasome (prosome, macropain) subunit, alpha type, 1 (PSMA1) pseudogene
ENSG00000237575	<b>PYY2</b>	peptide YY 2 (pseudogene)
ENSG00000237669		HLA complex group 4 pseudogene 3
ENSG00000237927		novel transcript
ENSG00000240032	<b>LNCsRLR</b>	lncRNA sorafenib resistance in renal cell carcinoma associated
ENSG00000241170	RPL31P46	ribosomal protein L31 pseudogene 46
ENSG00000241728		novel transcript, sense overlapping C21orf2
ENSG00000241765	RPS26P45	ribosomal protein S26 pseudogene 45
ENSG00000243649	CFB	complement factor B
ENSG00000245156		novel transcript
ENSG00000248208	WDR45P1	WD repeat domain 45 pseudogene 1
ENSG00000253930	TNFRSF10A-AS1	TNFRSF10A antisense RNA 1
ENSG00000254239		novel transcript
ENSG00000255629		novel transcript
ENSG00000255986	MT1JP	metallothionein 1J, pseudogene

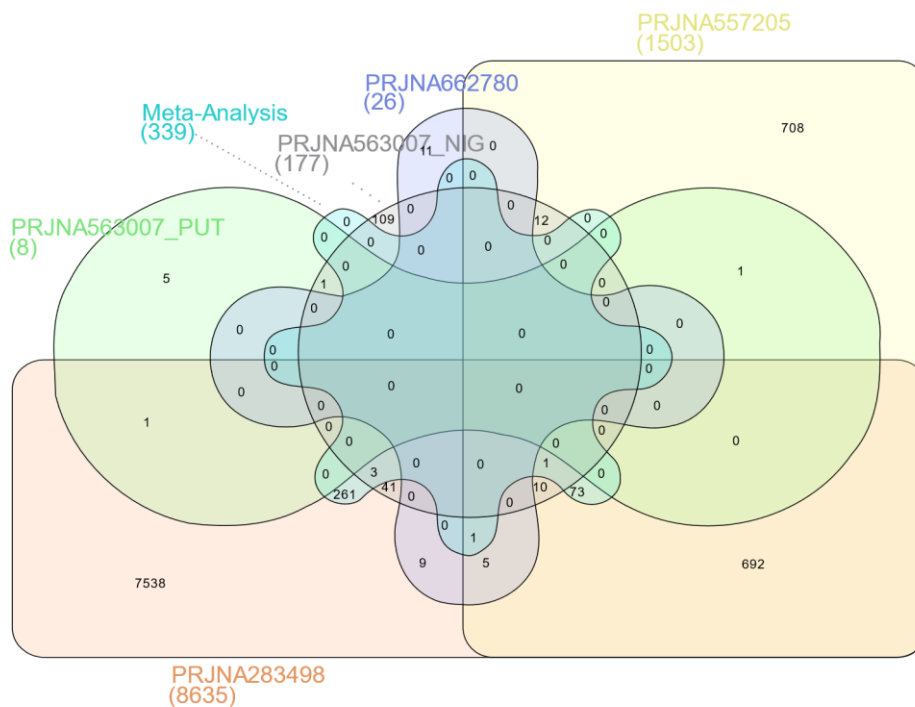
ENSG00000256007	ARAP1-AS1	ARAP1 antisense RNA 1
ENSG00000256597	LINC02393	long intergenic non-protein coding RNA 2393
ENSG00000256955		novel transcript
ENSG00000257702	LBX2-AS1	LBX2 antisense RNA 1
ENSG00000258232		novel transcript, antisense to TUBA1C
ENSG00000258384	RCCD1-AS1	RCCD1 and UNC45A antisense RNA 1
ENSG00000258471		novel transcript, antisense to SLC39A2
ENSG00000258811		novel transcript
ENSG00000258900	HNRNPCP1	heterogeneous nuclear ribonucleoprotein C pseudogene 1
ENSG00000260325	HSPB9	heat shock protein family B (small) member 9
ENSG00000260418		novel transcript
ENSG00000260549	MT1L	metallothionein 1L, pseudogene
ENSG00000260552	COSMOC	cell fate and sterol metabolism associated divergent transcript of MOCOS
ENSG00000262097	LINC02185	long intergenic non-protein coding RNA 2185
ENSG00000265206	MIR142HG	MIR142 host genes
ENSG00000266378		novel transcript
ENSG00000266903	CEACAM16-AS1	CEACAM16, CEACAM19 and PVR antisense RNA 1
ENSG00000267107	PCAT19	prostate cancer associated transcript 19
ENSG00000267206	LCN6	lipocalin 6
ENSG00000267291	MAP4K1-AS1	MAP4K1 antisense RNA 1
ENSG00000267388		novel transcript
ENSG00000267519	MIR23AHG	miR-23a/27a/24-2 cluster host gene
ENSG00000269720	CCDC194	coiled-coil domain containing 194
ENSG00000269959	SPACA6-AS1	SPACA6 antisense RNA 1
ENSG00000270164	LINC01480	long intergenic non-protein coding RNA 1480
ENSG00000271239		novel transcript
ENSG00000271653		novel transcript
ENSG00000271766		novel transcript
ENSG00000272183		novel transcript, antisense to LBX2
ENSG00000272217		novel transcript
ENSG00000272434		novel transcript, antisense to IMPDH2 and QRICH1
ENSG00000272512		novel transcript
ENSG00000272582		novel transcript, antisense to C22orf23
ENSG00000272783		novel transcript
ENSG00000273100		novel transcript
ENSG00000273113		novel transcript
ENSG00000273233		novel transcript
ENSG00000273356	LINC02019	long intergenic non-protein coding RNA 2019
ENSG00000274156		novel transcript
ENSG00000274286	ADRA2B	adrenoceptor alpha 2B
ENSG00000275329		novel transcript
ENSG00000275484		novel transcript

ENSG00000275910		novel transcript, antisense to PDXDC1
ENSG00000276043	UHRF1	ubiquitin like with PHD and ring finger domains 1
ENSG00000276445		novel transcript
ENSG00000276867		novel transcript
ENSG00000276980		novel transcript, sense intronic C3
ENSG00000277218		novel transcript
ENSG00000278000		novel transcript
ENSG00000278716		novel transcript
ENSG00000278737		novel transcript
ENSG00000278966		TEC
ENSG00000279196		novel transcript
ENSG00000279690		TEC
ENSG00000279821		TEC
ENSG00000280109	PLAC4	placenta enriched 4
ENSG00000280287		novel transcript
ENSG00000281327	LINC01338	long intergenic non-protein coding RNA 1338
ENSG00000281491	DNAJB5-DT	DNAJB5 divergent transcript Symbol;Acc:HGNC:49846]
ENSG00000284631	SCYGR4	small cysteine and glycine repeat containing 4
ENSG00000284640		novel transcript
ENSG00000284747		novel transcript, antisense to ERFFI1
ENSG00000285972	CERNA2	competing endogenous lncRNA 2 for microRNA let-7b

Πίνακας 13 Αναλυτική καταγραφή των 269 υπερ εκφρασμένων γονιδίων με τη χρήση του Biomart

Τα 339 διαφορετικώς εκφρασμένα γονίδια που προέκυψαν από την μετα-ανάλυση υποθέτουμε ότι σχετίζονται με τη νόσο του Πάρκινσον. Χρησιμοποιώντας το εργαλείο InteractiVenn (Heberle et al., 2015) οπτικοποιήσαμε τα 339 διαφορετικά εκφρασμένα γονίδια των μελετών (**Εικόνα 12**). Με το σχήμα του γκραντζιού παρατίθενται τα αποτελέσματα της μετα-ανάλυσης. Συγκεκριμένα παρατηρούμε ότι 261 διαφορετικώς εκφρασμένα γονίδια της μετα-ανάλυσης βρέθηκαν μόνο στη μελέτη PRJNA283498. 3 διαφορετικά εκφρασμένα γονίδια της μετα-ανάλυσης βρέθηκαν να είναι κοινά σε δυο μελέτες, τις PRJNA283498 και PRJNA563007\_Nigra. Ένα διαφορετικά εκφρασμένο γονίδιο της μετα-ανάλυσης βρέθηκε κοινό μεταξύ των μελετών PRJNA283498, PRJNA557205 και της PRJNA662780. Επίσης, 73 διαφορετικά εκφρασμένα γονίδια της μετα-ανάλυσης είναι κοινά μεταξύ μόνο των μελετών PRJNA283498 και PRJNA557205. Ένα διαφορετικά εκφρασμένο γονίδιο της μετα-ανάλυσης είναι κοινό στις μελέτες PRJNA563007\_Nigra, PRJNA283498 και PRJNA557205. Από τα παραπάνω προκύπτει ότι όλα τα διαφορετικά εκφρασμένα γονίδια της μετα-ανάλυσης αποτελούν διαφορετικά εκφρασμένα γονίδια της διαφορετικής ανάλυσης PRJNA283498 (**Εικόνα 12**).





Εικόνα 12 Αναπαράσταση των επικαλύψεων των γονιδίων μεταξύ των 5 μελετών χρησιμοποιώντας το πρόγραμμα InteractiVenn

Στη συνέχεια χρησιμοποιήθηκε το εργαλείο WebGestalt (Khatri et al., 2012) για την ανάλυση υπερεκπροσώπισης βιολογικών όρων που περιγράφουν τα διαφορετικά εκφρασμένα γονίδια που προκύπτουν από τη μετα-ανάλυση. Ορισμένα από τα πλεονεκτήματα του WebGestalt είναι ότι έχει την ικανότητα ανάκτησης περισσότερων πληροφοριών για όλα τα γονίδια σε ένα σύνολο γονιδίων, περισσότερους τρόπους οργάνωσης ενός συνόλου γονιδίων, καθώς επίσης έχει πολύ καλή οπτικοποίηση των αποτελεσμάτων. Χρησιμοποιούμε τη μέθοδο ανάλυσης υπερεκπρόσωπισης (Over-Representation Analysis -ORA) που παρέχει το WebGestalt. Η μέθοδος αυτή αξιολογεί στατιστικά τα γονίδια σε ένα συγκεκριμένο μονοπάτι που βρίσκεται μεταξύ του συνόλου των γονιδίων που εμφανίζουν αλλαγές στην έκφραση. Με την επιλογή διαφόρων βάσεων δεδομένων και διορθώνοντας το p-value χρησιμοποιώντας FDR, το εργαλείο αυτό εντοπίζει σε γονιδιακές λίστες, υπερεκπροσωπημένους όρους, όπως GO (Gene Ontology, 2015), KEGG (Kanehisa et al., 2002), WikiPathways (Kutmon et al., 2016) κ.ά.

Οι τρεις επιμέρους συνιστώσες της γονιδιακής οντολογίας (Gene Ontology) είναι:

- Βιολογική διεργασία (biological process)
- Μοριακή λειτουργία (molecular function)
- Κυτταρικός εντοπισμός (cellular component)

Η βιολογική διεργασία αναφέρεται στις προγραμματισμένες λειτουργίες του οργανισμού οι οποίες απαιτούν την υλοποίηση πολλαπλών μοριακών διαδικασιών συνεισφέροντας τελικά στην επίτευξη ενός βιολογικού σκοπού.

Η μοριακή λειτουργία αναφέρεται στις διεργασίες που εκτελεί ένα μακρομόριο αλληλοεπιδρώντας με διάφορες μοριακές οντότητες. Συνδέεται τόσο με τη βιοχημεία όσο και με τον γενικότερο ρόλο που έχει το μακρομόριο είτε μεμονωμένα είτε σε ένα γενικευμένο σύστημα ή μια διαδικασία.

Ο κυτταρικός εντοπισμός αποτελεί τη θέση στην οποία εντοπίζεται το μακρομόριο το οποίο εκτελεί μια βιολογική διεργασία όπως λόγω χάρη αυτή είναι τα ριβοσώματα, η πλασματική μεμβράνη ή τα μιτοχόνδρια.

Παρακάτω παρατίθεται ο πίνακας με τους πιο υπερεκπροσωπημένους όρους των 70 υποεκφρασμένων γονιδίων, με βάση το FDR αυτών των όρων. Συμπεραίνουμε ότι σχεδόν όλοι οι όροι σχετίζονται με την μνήμη και τις νοητικές διαδικασίες, τις συνάψεις και γενικότερα τη μετάδοση του νευρικού σήματος, καθώς και τη σχέση αυτών με διάφορα χαρακτηριστικά του εγκεφάλου που τυγχάνει να είναι καίρια για την εμφάνιση της νευροεκφυλιστικής νόσου του Πάρκινσον (**Πίνακας 14**). Όμως ο πιο σημαντικός όρος με ιδιαίτερη σημασία είναι η μακροπρόθεσμη μνήμη (long-term memory). Ελέγχοντας και τις υπόλοιπες πτυχές της γονιδιακής οντολογίας παρατηρούμε ότι αναφορικά με τον υποκυτταρικό εντοπισμό των γονιδίων ο πιο στατιστικά σημαντικός όρος είναι το δένδριτικό διαμέρισμα (**Πίνακας 15**). Εξετάζοντας και τη μοριακή λειτουργία παρατηρούμε ότι ο όρος με ιδιαίτερη σημασία είναι η δραστηριότητα των ορμονών (**Πίνακας 16**). Περαιτέρω ανάλυση και έλεγχος στη σύνδεση των γονιδίων με ασθένειες καταδεικνύει ότι τα γονίδια αυτά σχετίζονται με την ψυχική κατάθλιψη χρησιμοποιώντας την πλατφόρμα DisGeNet (Pinero et al., 2020) (**Πίνακας 17**). Η βάση δεδομένων KEGG περιλαμβάνει γονίδια που είναι καταχωρημένα σε μονοπάτια. Επομένως πραγματοποιήσαμε και την ανάλυση εμπλουτισμού ομάδων γονιδίων με σκοπό να ελέγξουμε τα βιοχημικά/ μεταβολικά μονοπάτια KEGG στα οποία εμπλέκονται τα γονίδια αυτά. Ένα μεταβολικό μονοπάτι είναι μια σειρά από διάφορες χημικές αντιδράσεις που πραγματοποιούνται ενός του κύτταρου σε διαφορετικές χρονικές στιγμές. Η ανάλυση εμπλουτισμού για τα μονοπάτια KEGG των διαφορετικά υπο εκφρασμένων γονιδίων δεν επέστρεψε αποτελέσματα με όριο FDR<0.05 όπως έχει οριστεί.

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GO:0007616	<b>long-term memory</b>	32	0.063363	63.13	4.40E-07	<b>0.004004</b>
GO:0007613	<b>memory</b>	111	0.21979	22.75	2.4483E-06	0.011129
GO:0071385	cellular response to glucocorticoid stimulus	60	0.1188	33.67	5.7437E-06	0.0133
GO:0007611	<b>learning or memory</b>	244	0.48314	12.42	7.3645E-06	0.0133
GO:0071384	cellular response to corticosteroid stimulus	64	0.12673	31.56	7.4423E-06	0.0133
GO:0009725	response to hormone	962	1.9048	5.25	1.0721E-05	0.0133
GO:0007610	behavior	582	1.1524	6.942	1.3512E-05	0.0133

GO:0050804	modulation of chemical synaptic transmission	416	0.82371	8.498	1.3987E-05	0.0133
GO:0099177	<b>regulation of trans-synaptic signaling</b>	417	0.82569	8.478	1.4206E-05	0.0133
GO:0050806	positive regulation of synaptic transmission	160	0.31681	15.78	0.00001463	0.0133
GO:0050890	Cognition	283	0.56036	10.71	1.7123E-05	0.014151
GO:0009719	response to endogenous stimulus	1595	3.1582	3.8	2.9668E-05	0.019668
GO:0071495	cellular response to endogenous stimulus	1347	2.6672	4.124	3.3721E-05	0.019668
GO:0008628	hormone-mediated apoptotic signaling pathway	5	0.0099	202	0.00003788	0.019668
GO:0035934	corticosterone secretion	5	0.0099	202	0.00003788	0.019668
GO:0008585	female gonad development	98	0.19405	20.61	4.0355E-05	0.019668
GO:0007268	<b>chemical synaptic transmission</b>	685	1.3564	5.898	4.3547E-05	0.019668
GO:0098916	anterograde trans-synaptic signaling	685	1.3564	5.898	4.3547E-05	0.019668
GO:0032870	cellular response to hormone stimulus	689	1.3643	5.864	4.5385E-05	0.019668
GO:0007218	neuropeptide signaling pathway	101	0.19999	20	4.5423E-05	0.019668

Πίνακας 14 Αποτελέσματα της ανάλυσης εμπλουτισμού των όρων της γονιδιακής οντολογίας και των βιολογικών μονοπατιών των διαφορικά υπο εκφρασμένων γονιδίων

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GO:0036477	<b>somatodendritic compartment</b>	826	1.2674	7.8899	2.00E-07	<b>0.00017511</b>
GO:0097458	neuron part	1690	2.5932	5.0131	3.12E-07	0.00017511
GO:0043025	neuronal cell body	486	0.74574	10.728	4.47E-07	0.00017511
GO:0044297	cell body	551	0.84548	9.4621	0.000001153	0.00033869
GO:0030424	axon	595	0.91299	8.7624	2.0494E-06	0.00048162
GO:0045202	synapse	1154	1.7707	5.6473	4.2656E-06	0.00083535
GO:0043005	neuron projection	1288	1.9764	5.0598	0.000011346	0.0019044
GO:0044456	synapse part	921	1.4132	5.6608	0.000049758	0.0073082
GO:0098793	presynapse	483	0.74113	8.0957	0.000075136	0.0098094
GO:0044463	cell projection part	1425	2.1866	4.116	0.00017883	0.019103

GO:0120038	plasma membrane bounded cell projection part	1425	2.1866	4.116	0.00017883	0.019103
GO:0033267	axon part	374	0.57388	8.7126	0.00023188	0.020958
GO:0030133	transport vesicle	374	0.57388	8.7126	0.00023188	0.020958

Πίνακας 15 Αποτελέσματα της ανάλυσης εμπλουτισμού της γονιδιακής οντολογίας και τον υποκυτταρικό εντοπισμό των γονιδίων

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GO:0005179	hormone activity	117	0.21756	22.982	2.29E-06	0.004307
GO:0005184	neuropeptide hormone activity	27	0.050207	59.753	1.65E-05	0.012881
GO:0071855	neuropeptide receptor binding	29	0.053926	55.632	2.06E-05	0.012881

Πίνακας 16 Αποτελέσματα ανάλυσης εμπλουτισμού της γονιδιακής οντολογίας και των γονιδιακών προϊόντων σε μοριακό επίπεδο

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
C0011570	<b>Mental Depression</b>	260	0.55301	10.85	1.13E-05	0.041585

Πίνακας 17 Αποτελέσματα σύνδεσης των γονιδίων με ασθένειες

Ομοίως, ακολουθείται η ίδια διαδικασία και για τα 269 διαφορεικά υπερεκφρασμένα γονίδια λαμβάνοντας τα εξής αποτελέσματα: Αναφορικά με τις βιολογικές διεργασίες στις οποίες εμπλέκονται τα υπερεκφρασμένα γονίδια, ο κυριότερος εκπροσωπημένος και στατικά σημαντικός όρος είναι η φλεγμονή (Πίνακας 18). Σχετικά με τη μοριακή λειτουργία δεν βρέθηκε κάποιος σημαντικός όρος υπερεκπροσώπισης με  $FDR < 0.05$  όπως έχει οριστεί. Σχετικά με τον υποκυτταρικό εντοπισμό των γονιδίων αυτών παρατηρούμε ότι ο πιο σημαντικός όρος είναι η εξωκυττάρια ουσία (Extracellular matrix) (Πίνακας 19). Πραγματοποιώντας και την ανάλυση εμπλουτισμού ομάδων γονιδίων με σκοπό να ελέγξουμε τα βιοχημικά/ μεταβολικά μονοπάτια KEGG που εμπλέκονται τα γονίδια αυτά, η βάση επέστρεψε το μονοπάτι των καταρρακτών συμπληρώματος και πήξης (complement and coagulation cascades) (Πίνακας 20). Τέλος, ο έλεγχος για τη σύνδεση των γονιδίων με ασθένειες δεν επέστρεψε κανένα αποτέλεσμα με  $FDR < 0.05$ .

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GO:0006954	<b>inflammatory response</b>	717	6.1951	4.0354	2.32E-09	<b>0.000021132</b>
GO:0051707	response to other organism	897	7.7504	3.4837	1.14E-08	0.00003617

GO:0043207	response to external biotic stimulus	899	7.7677	3.4759	1.19E-08	0.00003617
GO:0009607	response to biotic stimulus	926	8.001	3.3746	2.22E-08	0.000050376
GO:0006952	defense response	1518	13.116	2.6685	5.13E-08	0.000093197
GO:0009617	response to bacterium	595	5.141	3.8903	2.00E-07	0.00030289
GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	248	2.1428	6.0668	2.40E-07	0.0003113
GO:0001775	cell activation	1335	11.535	2.6875	3.07E-07	0.00034862
GO:0048646	anatomical structure formation involved in morphogenesis	1067	9.2192	2.9287	3.92E-07	0.00037692
GO:0035239	tube morphogenesis	806	6.9641	3.3026	4.15E-07	0.00037692
GO:0007155	cell adhesion	1369	11.829	2.6208	5.31E-07	0.00043877
GO:0022610	biological adhesion	1377	11.898	2.6055	6.02E-07	0.00045627
GO:0010594	regulation of endothelial cell migration	151	1.3047	7.6646	7.52E-07	0.00052594
GO:0001525	angiogenesis	487	4.2078	4.0401	1.06E-06	0.00068827
GO:0002366	leukocyte activation involved in immune response	693	5.9878	3.3401	2.16E-06	0.0009756
GO:0051051	negative regulation of transport	458	3.9573	4.0432	2.23E-06	0.0009756
GO:0002286	T cell activation involved in immune response	99	0.85539	9.3524	2.26E-06	0.0009756
GO:0002274	myeloid leukocyte activation	634	5.478	3.4684	2.30E-06	0.0009756
GO:0048514	blood vessel morphogenesis	574	4.9596	3.6294	2.31E-06	0.0009756

GO:0010632	regulation of epithelial cell migration	212	1.8318	6.0052	2.32E-06	0.0009756
GO:0002443	leukocyte mediated immunity	760	6.5667	3.198	2.34E-06	0.0009756
GO:0002263	cell activation involved in immune response	697	6.0223	3.321	2.36E-06	0.0009756
GO:0009611	response to wounding	642	5.5471	3.4252	2.76E-06	0.0010924
GO:0032101	regulation of response to external stimulus	771	6.6617	3.1523	2.93E-06	0.0011092
GO:0001568	blood vessel development	654	5.6508	3.3624	3.62E-06	0.0013153
GO:0035295	tube development	992	8.5712	2.8001	4.17E-06	0.0014574
GO:0010631	epithelial cell migration	274	2.3675	5.0687	4.63E-06	0.0014764
GO:0002252	immune effector process	1141	9.8586	2.6373	4.65E-06	0.0014764
GO:0006955	immune response	1919	16.581	2.1712	4.71E-06	0.0014764
GO:0090132	epithelium migration	277	2.3934	5.0138	5.18E-06	0.0015688
GO:0090130	tissue migration	283	2.4452	4.9075	6.44E-06	0.0018203
GO:0045765	regulation of angiogenesis	283	2.4452	4.9075	6.44E-06	0.0018203
GO:0001944	vasculature development	682	5.8927	3.2243	6.61E-06	0.0018203
GO:0043542	endothelial cell migration	195	1.6849	5.9352	7.54E-06	0.0020153
GO:0072358	cardiovascular system development	691	5.9705	3.1823	7.96E-06	0.0020682
GO:0045087	innate immune response	827	7.1456	2.9389	8.65E-06	0.0021838
GO:0001817	regulation of cytokine production	642	5.5471	3.2449	1.08E-05	0.0026602
GO:0001816	cytokine production	712	6.1519	3.0885	1.21E-05	0.0029063
GO:0042060	wound healing	532	4.5967	3.4808	1.47E-05	0.0033938
GO:0098542	defense response to other organism	473	4.0869	3.6703	1.49E-05	0.0033938
GO:0007596	blood coagulation	322	2.7822	4.3131	2.34E-05	0.0047331

GO:0002237	response to molecule of bacterial origin	330	2.8513	4.2086	2.98E-05	0.0053132
GO:0002684	positive regulation of immune system process	979	8.4589	2.6008	3.44E-05	0.0060079
GO:0097190	apoptotic signaling pathway	582	5.0287	3.1817	4.38E-05	0.0073698
GO:0046903	secretion	1605	13.868	2.0912	9.69E-05	0.013932
GO:0043535	regulation of blood vessel endothelial cell migration	88	0.76035	7.8911	1.14E-04	0.015634
GO:0046685	response to arsenic-containing substance	30	0.25921	15.431	1.23E-04	0.016121
GO:0022407	regulation of cell-cell adhesion	383	3.3093	3.6262	1.24E-04	0.016121
GO:0019221	cytokine-mediated signaling pathway	705	6.0914	2.7908	1.26E-04	0.016121
GO:0009967	positive regulation of signal transduction	1571	13.574	2.0628	1.65E-04	0.019491
GO:1902106	negative regulation of leukocyte differentiation	95	0.82083	7.3096	1.73E-04	0.019688
GO:0043299	leukocyte degranulation	529	4.5707	3.063	2.00E-04	0.021077
GO:0002701	negative regulation of production of molecular mediator of immune response	34	0.29377	13.616	2.03E-04	0.021077
GO:0009408	response to heat	139	1.201	5.8284	2.05E-04	0.021077
GO:0030198	extracellular matrix organization	347	2.9982	3.6689	2.16E-04	0.021831
GO:0009628	response to abiotic stimulus	1123	9.7031	2.2673	2.53E-04	0.024563
GO:1903531	negative regulation of secretion by cell	192	1.6589	4.8223	2.66E-04	0.025412

GO:0042089	cytokine biosynthetic process	105	0.90724	6.6135	2.99E-04	0.027179
GO:0016064	immunoglobulin mediated immune response	110	0.95044	6.3129	3.84E-04	0.032314
GO:0043618	regulation of transcription from RNA polymerase II promoter in response to stress	72	0.6221	8.0372	3.95E-04	0.032314
GO:0030097	hemopoiesis	791	6.8345	2.4874	4.82E-04	0.035646
GO:0032943	mononuclear cell proliferation	266	2.2983	3.9159	5.12E-04	0.03697
GO:0097746	regulation of blood vessel diameter	123	1.0628	5.6457	6.95E-04	0.046831
GO:0070370	cellular heat acclimation	5	0.043202	46.294	7.29E-04	0.047329

Πίνακας 18 Αποτελέσματα της ανάλυσης εμπλουτισμού των όρων της γονιδιακής οντολογίας και των βιολογικών μονοπατιών των διαφορικά υπερεκφρασμένων γονιδίων

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GO:0031012	<b>extracellular matrix</b>	496	3.3826	4.7301	2.61E-07	<b>0.000307</b>
GO:0009986	cell surface	782	5.333	3.3752	5.75E-06	0.002192
GO:0031226	intrinsic component of plasma membrane	1673	11.409	2.4541	5.94E-06	0.002192
GO:0005887	integral component of plasma membrane	1596	10.884	2.4806	7.46E-06	0.002192
GO:0062023	collagen-containing extracellular matrix	366	2.496	4.407	4.16E-05	0.009439
GO:0030141	secretory granule	831	5.6672	2.9997	4.82E-05	0.009439
GO:0044433	cytoplasmic vesicle part	1462	9.9704	2.3068	0.000123	0.020635
GO:0099503	secretory vesicle	976	6.6561	2.5541	0.000333	0.048877

Πίνακας 19 Υποκυτταρικός εντοπισμός των διαφορικά υπερεκφρασμένων γονιδίων



Gene Set	Description	Size	Expect	Ratio	P Value	FDR
hsa04610	<b>Complement and coagulation cascades</b>	79	0.90963	7.6955	3.11E-05	<b>0.005878</b>
hsa05134	Legionellosis	55	0.63328	9.4744	3.61E-05	0.005878
hsa04060	Cytokine-cytokine receptor interaction	294	3.3852	3.2494	0.000512	0.04544
hsa04151	PI3K-Akt signaling pathway	354	4.076	2.944	0.000683	0.04544
hsa05164	Influenza A	171	1.9689	4.0631	0.000744	0.04544
hsa04640	Hematopoietic cell lineage	97	1.1169	5.3721	0.000836	0.04544

Πίνακας 20 Αποτελέσματα της βάσης KEGG σχετικά με τα μοριακά και βιοχημικά μονοπάτια που εμπλέκονται τα γονίδια

Εκτός από τη χρήση του εργαλείου WebGestalt, χρησιμοποιήσαμε και τη βάση δεδομένων STRING, προκείμενου να δούμε τις συσχετίσεις των πρωτεϊνών μεταξύ των γονιδίων και να συγκρίνουμε τα αποτελέσματα των δύο αυτών εργαλείων. Έτσι εισάγουμε αρχικά τα 70 διαφορεικά υποεκφρασμένα γονίδια που προέκυψαν από τη μετα-ανάλυση (**Εικόνα 13**). Σε αυτό το σημείο θα πρέπει να τονίσουμε ότι δεν υπάρχει κάποια εισαγωγή εγγραφών στην βάση STRING η οποία να είναι αναγνωριστικό ασθένειας ή χαρακτηριστικό του εγκεφάλου. Από τα 70 διαφορεικά υποεκφρασμένα γονίδια που εισήχθησαν στη βάση, μόνο τα 40 αναγνωρίστηκαν και συσχετίστηκαν με διαφορετικά χαρακτηριστικά του εγκεφάλου (ορμόνες, νευροπλαστικότητα, συνάψεις, ικανότητα μάθησης κ.α.). Αυτό συμβαίνει διότι 30 διαφορεικά υποεκφρασμένα γονίδια ανήκουν σε non-coding RNAs και εργαλεία δικτύων αλληλεπιδράσεων μεταξύ πρωτεϊνών (PPI), όπως η βάση δεδομένων STRING, δεν μελετούν non-coding RNAs (Zogoroulos et al., 2023). Στα 40 διαφορεικά εκφρασμένα γονίδια που αναγνωρίστηκαν (κόμβοι), ο αριθμός των σχέσεων μεταξύ τους (ακμές) είναι 30 (**Εικόνα 13**). Από τα 40 διαφορεικά υποεκφρασμένα γονίδια, όπως παρουσιάζεται στην Εικόνα 13, κεντρικό ρόλο φαίνεται να διαδραματίζει ο εγκεφαλικός νευροτροφικός παράγοντας BDNF (Brain Derived Neurotrophic Factor), ο οποίος ανήκει στις νευροτροφίνες (NTs), μια οικογένεια πρωτεϊνών που υποστηρίζουν τη λειτουργία του ΚΝΣ. Οι νευροτροφίνες συντίθενται κυρίως στο ΚΝΣ, αλλά και στα μη νευρωνικά περιφερικά κύτταρα, όπως τα T και B λεμφοκύτταρα, τα μονοκύτταρα, τα αγγειακά ενδοθηλιακά λεία και σκελετικά μυϊκά κύτταρα. Η έκφρασή του έχει επιβεβαιωθεί στον ιππόκαμπο, τον μετωπιαίο φλοιό, τον μεσεγκέφαλο, στην αμυγδαλή καθώς και στον υποθάλαμο και υποστηρίζει την ανάπτυξη του νευρικού συστήματος με ενεργή κυτταρική διαφοροποίηση, νευρωνική ανάπτυξη, νευρογένεση καθώς και συναπτική πλαστικότητα. Επιπλέον, προάγει τη νευροπροστασία και τη νευροαναγέννηση και βελτιώνει τη ντοπαμινεργική νευροδιαβίβαση (Palasz et al., 2020).

Οι νευροχημικές συνάψεις του συγκεκριμένου γονιδίου είναι περισσότερες από το αναμενόμενο και πιο συγκεκριμένα εμπλέκεται με το σύμπλεγμα των υποεκφρασμένων γονιδίων *TAC1*, *CRH*, *SST*, *CALB1*, τα οποία με τη σειρά τους έχουν σημαντικές

νευροχημικές συνάψεις μεταξύ τους. Αναλυτικότερα, το γονίδιο *TAC1* κωδικοποιεί τέσσερα προϊόντα της οικογένειας ορμονών πεπτιδίων ταχυκινίνης, την ουσία P και τη νευροκινίνη A, καθώς και το νευροπεπτίδιο K και γάμμα. Πρόκειται για ορμόνες που λειτουργούν ως νευροδιαβιβαστές και αλληλοεπιδρούν με τους νευρικούς υποδοχείς και τα λεία μυϊκά κύτταρα, με την ουσία P να αποτελεί ένα αντιμικροβιακό πεπτίδιο με αντιβακτηριακές και αντιμυκητιασικές ιδιότητες. (Zheng et al., 2021).

Το γονίδιο *CRH* κωδικοποιεί ένα μέλος της οικογένειας του παράγοντα απελευθέρωσης κορτικοτροπίνης, μια κωδικοποιημένη προπροπρωτεΐνη, η οποία υποβάλλεται σε πρωτεολυτική επεξεργασία για την παραγωγή νευροπεπτιδικής ορμόνης. Ως απόκριση του στρες, αυτή η ορμόνη εκκρίνεται στον παρακοιλιακό πυρήνα του υποθαλάμου, συνδέεται με τους υποδοχείς της ορμόνης απελευθέρωσης κορτικοτροπίνης και διεγείρει την απελευθέρωση της φλοιοεπινεφριδικής ορμόνης στην υπόφυση (Koutmani et al., 2013).

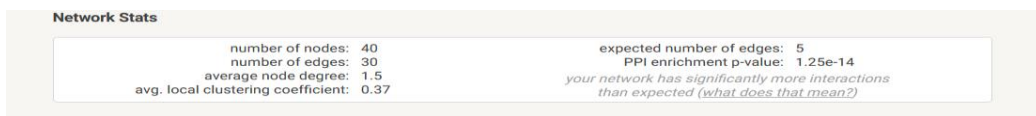
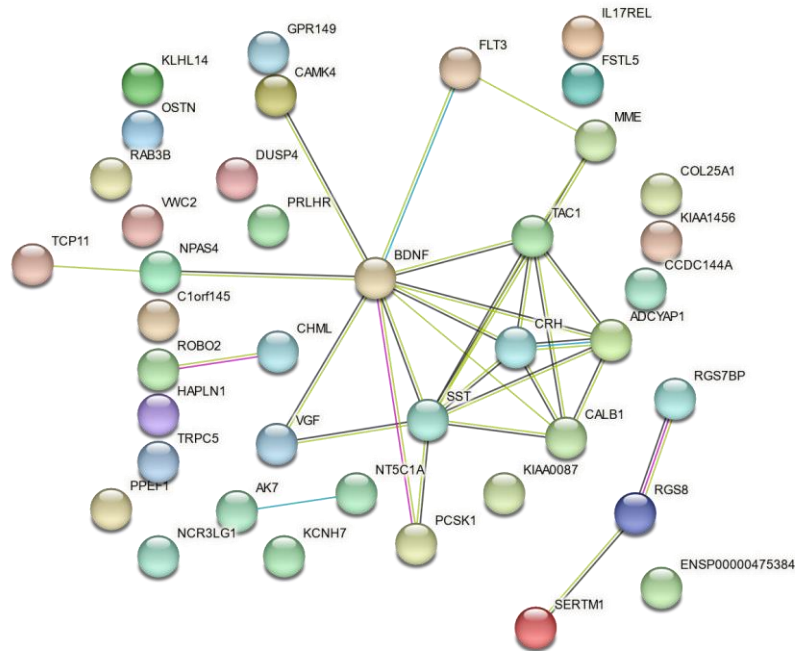
Το γονίδιο *SST*, δηλαδή το νευροπεπτίδιο σωματοστατίνη, εκφράζεται σε όλο τον εγκέφαλο και πιο συγκεκριμένα στον φλοιό σε ένα υποσύνολο νευρώνων που είναι γνωστός ως πρωτεϊνικός δείκτης ανασταλτικών ενδονευρώνων. Η σωματοστατίνη εμπλέκεται στη ρύθμιση κυκλωμάτων του εγκεφαλικού φλοιού. Επιπλέον, εκφράζεται και σε όλο τον οργανισμό, αναστέλλοντας την απελευθέρωση πολυάριθμων δευτερογενών ορμονών και αποτελεί έναν σημαντικό ρυθμιστή του ενδοκρινικού συστήματος, μέσω των αλληλεπιδράσεων της με την αυξητική ορμόνη της υπόφυσης, την ορμόνη διέγερσης του θυρεοειδούς και τις περισσότερες ορμόνες του γαστρεντερικού σωλήνα. Εκτός αυτών, επηρεάζει τους ρυθμούς νευροδιαβίβασης στο ΚΝΣ (McGregor and Nelson, 2019).

Όσον αφορά στο γονίδιο *CALBI*, μια πρωτεΐνη που δεσμεύει το ασβέστιο, η έκφρασή του παρατηρείται σε υψηλότερα επίπεδα στους νευρώνες ντοπαμίνης προερχόμενους από την κοιλιακή καλυπτική περιοχή (VTA) και όχι από τη γειτονική μέλαινα ουσία. Δρα μέσω μηχανισμών που υποστηρίζουν τόσο την απελευθέρωση όσο και την πρόσληψη νευρώνων ντοπαμίνης στην κοιλιακή περιοχή αλλά όχι στο ραβδωτό (Brimblecombe et al., 2019).

Τα ανωτέρω γονίδια σε φυσιολογικούς εγκεφάλους υπερεκφράζονται ενώ, όπως βρέθηκε επί νόσου Πάρκινσον, η υποέκφρασή τους και οι πολλαπλές νευροχημικές συνάψεις μεταξύ τους εξηγούν τα ελλείμματα στη διαμόρφωση χημικής συναπτικής μετάδοσης, στη μνήμη, στους νευρώνες αλλά και στην απορρύθμιση των ορμονών.

Τα αποτελέσματά μας φαίνονται να είναι αξιόπιστα, καθότι οι συνδέσεις-συσχετίσεις μεταξύ των γονιδίων είναι πολύ περισσότερες από αυτές που θα δημιουργούνταν αν στη θέση των 40 γονιδίων τοποθετούνταν προς ανάλυση 40 τυχαία γονίδια, καθώς η βάση επιστρέφει ότι ο εκτιμώμενος αριθμός σχέσεων-συσχετίσεων μεταξύ των τυχαίων γονιδίων θα ήταν 5. Ένας τέτοιος εμπλουτισμός που η πιθανότητα p-value να προκύψει τυχαία είναι  $1.25 \times 10^{-14}$ , δείχνει ότι οι πρωτεΐνες είναι τουλάχιστον εν μέρει βιολογικά συνδεδεμένες, ως ομάδα. Παρακάτω βλέπουμε τους υπερεκπροσωπημένους όρους των τριών επιμέρους οντολογιών της γονιδιακής οντολογίας (**Πίνακας 21**, **Πίνακας 22**, **Πίνακας 23**). Συγκεκριμένα, ο πιο στατιστικά σημαντικός υπερεκπροσωπημένος βιολογικός όρος είναι η διαμόρφωση χημικής συναπτικής μετάδοσης (modulation of chemical synaptic transmission). Έτσι βλέπουμε ότι από τα

40 διαφορεικά υποεκφρασμένα γονίδια που αναγνωρίζει η βάση δεδομένων, 9 από τα γονίδια αυτά σχετίζονται με τον όρο αυτόν, όταν συνολικά υπάρχουν 446 γονίδια αυτού του όρου κατατεθειμένα στη βάση. Παρόμοια, από τα 40 διαφορεικά υποεκφρασμένα γονίδια, 4 από αυτά σχετίζονται με την μακροπρόθεσμη μνήμη, όταν η βάση συνολικά περιλαμβάνει 34 τέτοια γονίδια. Τα αποτελέσματα επιβεβαιώθηκαν και με το εργαλείο WebGestalt. Ελέγχοντας επίσης, τόσο τον υποκυτταρικό όσο και τον κυτταρικό εντοπισμό των γονιδίων παρατηρούμε ότι σχεδόν όλοι οι όροι σχετίζονται με διάφορα χαρακτηριστικά του εγκεφάλου που σχετίζονται με την εμφάνιση του φαινοτύπου της νευροεκφυλιστικής νόσου του Πάρκινσον (Πίνακας 22, Πίνακας 23). Σχετικά με τη μοριακή λειτουργία, ο όρος που περιλαμβάνεται είναι η δραστηριότητα των ορμονών. Αντίστοιχα αποτελέσματα εξάγουμε και από το εργαλείο WebGestalt. Καταληκτικά, μέσω της βάσης αυτής εντοπίζονται τόσο οι σχέσεις-συνδέσεις των 40 διαφορεικά υποεκφρασμένα γονιδίων μεταξύ τους, όσο και οι συσχετίσεις αυτών με διάφορα χαρακτηριστικά του Πάρκινσον.



Εικόνα 13 Δίκτυο αλληλεπιδράσεων πρωτεϊνών- πρωτεϊνών των 70 διαφορεικά υπο εκφρασμένων γονιδίων με χρήση του προγράμματος STRING

Term description	Observed Gene count	Background gene count	FDR	matching proteins in your network (labels)
<b>Modulation of chemical synaptic transmission</b>	<b>9</b>	<b>446</b>	0.0033	VGF,CALB1,CRH,NPAS4,TAC1,RAB3B,BDNF,MME,ADCYAP1
Learning or memory	7	259	0.0043	CALB1,CRH,CAMK4,NPAS4,TAC1,BDNF,MME

<b>Long-term memory</b>	<b>4</b>	<b>34</b>	0.00 43	CALB1,CAMK4,NPAS4,TAC1
Behavior	9	572	0.00 51	PRLHR,CALB1,CRH,CAMK4,NPAS4,TAC1,BDNF,MME,ADCYAP1
Regulation of synaptic plasticity	6	191	0.00 58	VGF,CALB1,CRH,NPAS4,BDNF,MME
Memory	5	122	0.00 93	CALB1,CAMK4,NPAS4,TAC1,BDNF
Cellular response to glucocorticoid stimulus	4	59	0.01 16	FLT3,CRH,NPAS4,ADCYAP1
Positive regulation of synaptic transmission	5	147	0.01 62	CRH,TAC1,RAB3B,MME,ADCYAP1
Multicellular organismal response to stress	4	71	0.01 74	CRH,TAC1,BDNF,ADCYAP1
Regulation of membrane potential	7	440	0.02 94	TRPC5,NPAS4,TAC1,KCNH7,RGS7BP,BDNF,ADCYAP1
Response to hormone	9	849	0.03 71	FLT3,VGF,CRH,SST,NPAS4,TAC1,OSTN,ROBO2,ADCYAP1
Positive regulation of corticosterone secretion	2	3	0.03 71	CRH,TAC1
Regulation of synaptic transmission, gabaergic	3	36	0.04 88	NPAS4,TAC1,BDNF
Regulation of postsynaptic membrane potential	4	104	0.04 88	NPAS4,RGS7BP,BDNF,ADCYAP1

Πίνακας 21 Βιολογικοί όροι γονιδιακής οντολογίας οι οποίοι είναι περισσότερο εμπλουτισμένοι στο δίκτυο

Term description	Observed gene count	Background gene count	FDR	matching proteins in your network (labels)
<b>Neuron projection</b>	9	565	0.00 42	RGS8,CALB1,CRH,SST,NPAS4,TAC1,BDNF,ROBO2,MME
<b>Dendrite</b>	6	209	0.00 53	RGS8,CALB1,SST,NPAS4,BDNF,MME
Plasma membrane bounded cell projection	11	1243	0.01 34	PRLHR,RGS8,CALB1,CRH,SST,TC P11,NPAS4,TAC1,BDNF,ROBO2,MME
Distal dendrite	2	5	0.02 79	SST,NPAS4

Πίνακας 22 Υποκυτταρικός εντοπισμός ορισμένων πρωτεϊνών από τα γονίδια που εισήχθησαν

Term description	Observed gene count	Background gene count	FDR	matching proteins in your network (labels)
<b>Synapse</b>	14	1351	0.00042	RGS8,CALB1,HAPLN1,CRH,CA MK4,SST,NPAS4,TAC1,RGS7B P,VWC2,RAB3B,BDNF,MME, ADCYAP1
Neuronal cell body	9	518	0.00076	RGS8,TRPC5,CALB1,CRH,SST, TAC1,KLHL14,MME,ADCYAP1
<b>Somatodendritic compartment</b>	10	852	0.0028	RGS8,TRPC5,CALB1,CRH,SST, TAC1,KLHL14,BDNF,MME,AD CYAP1
Neuron projection	12	1366	0.0041	RGS8,TRPC5,CALB1,CRH,PCS K1,TAC1,KLHL14,GPR149,BD NF,ROBO2,MME,ADCYAP1
Plasma membrane bounded cell projection	15	2193	0.0041	PRLHR,RGS8,TRPC5,CALB1,A K7,CRH,PCSK1,TCP11,TAC1,K LHL14,GPR149,BDNF,ROBO2, MME,ADCYAP1
Axon	8	646	0.0083	TRPC5,CALB1,CRH,TAC1,BDN F,ROBO2,MME,ADCYAP1

Πίνακας 23 Κυτταρικός εντοπισμός ορισμένων πρωτεϊνών από τα γονίδια που εισήχθησαν στη βάση

Term description	Observed gene count	Background gene count	FDR	matching proteins in your network (labels)
<b>Hormone activity</b>	5	124	0.0202	VGF,CRH,SST,OSTN,ADCYAP 1

Πίνακας 24 Αποτελέσματα ανάλυσης σε μοριακό επίπεδο

Αντίστοιχο δίκτυο εξάγουμε και για τα 269 διαφορεικά και στατιστικά σημαντικά υπερεκφρασμένα γονίδια (**Εικόνα 14**). Τα περισσότερα υπερεκφρασμένα γονίδια που βρέθηκαν, σχετίζονται με την φλεγμονή και το οξειδωτικό στρες με πολλαπλές νευροχημικές συνάψεις μεταξύ τους με τα γονίδια *IL-18*, *CSF1*, *HSPA1B*, *CXCR4*, *TLR5* και *HSPB1* να παρουσιάζουν ιδιαίτερο ενδιαφέρον. Πιο συγκεκριμένα, η *IL-18* αποτελεί μια προφλεγμονώδη κυτοκίνη που εμπλέκεται στον φλεγμονώδη καταρράκτη. Αποτελεί πρόδρομο στο κυτταρόπλασμα μιας ποικιλίας κυττάρων συμπεριλαμβανομένων και των μακροφάγων (Liu et al., 2016). Το γονίδιο *CSF1* κωδικοποιεί την πρωτεΐνη κυτοκίνη που ελέγχει την παραγωγή, διαφοροποίηση και λειτουργία των μακροφάγων, με τη δραστική της μορφή να εντοπίζεται εξωκυτταρικά, ως ένα ομοδιμερές συνδεδεμένο με δισουλφίδιο (Chang et al., 2019).

Το γονίδιο *HSPA1B* αποκωδικοποιεί την πρωτεΐνη θερμικού σοκ, 70kDa, μέλος της οικογένειας πρωτεϊνών 70. Ο ρόλος της είναι η σταθεροποίηση των πρωτεϊνών έναντι της συσσωμάτωσης και μεσολαβεί στην αναδίπλωση νεομεταφρασμένων πρωτεϊνών στο κυτταρόπλασμα. Επιπλέον, σχετίζεται με την μοριακή σύνοδο πρωτεΐνη (Joshi et

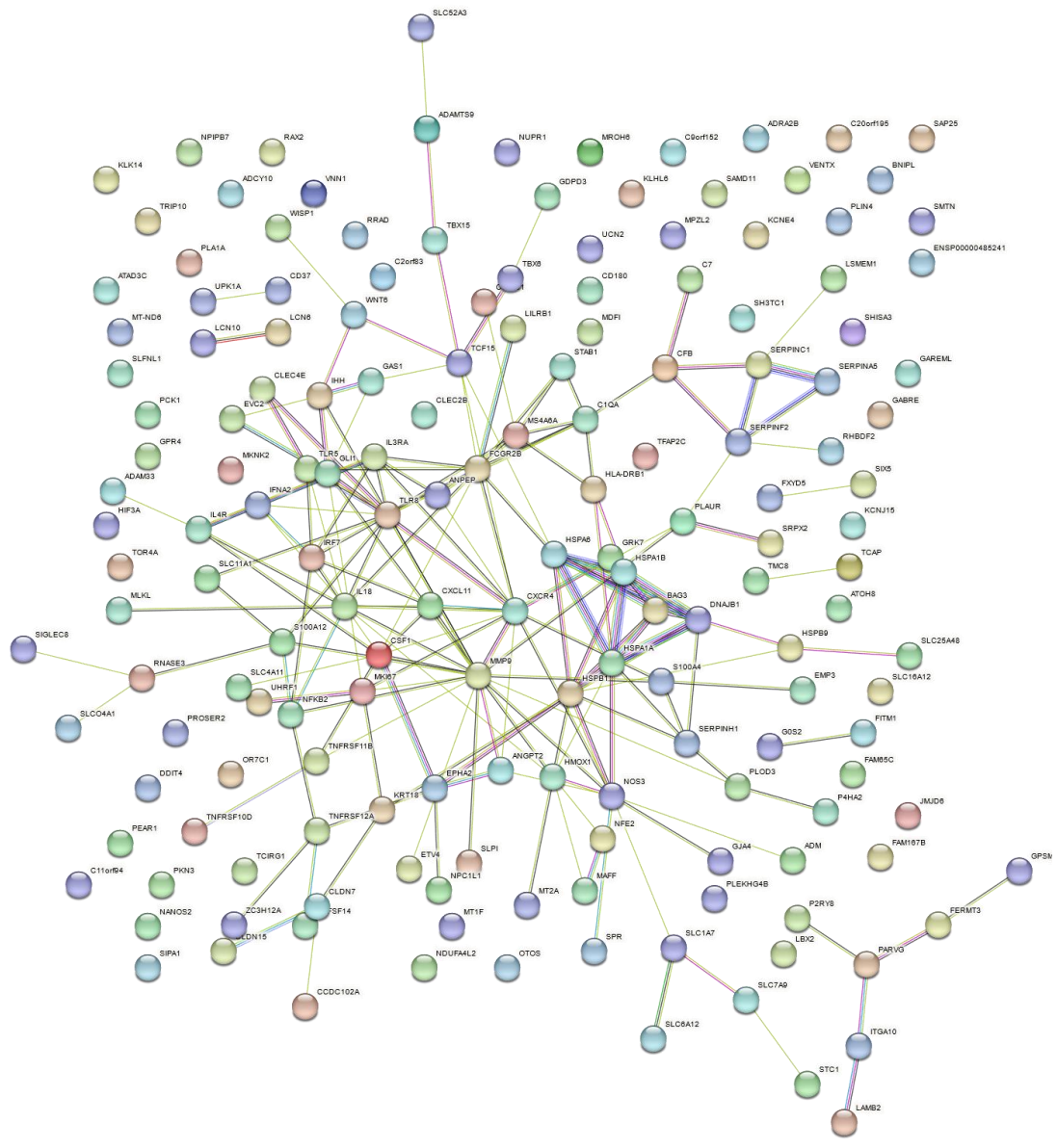
al., 2020), η οποία βρέθηκε και αυτή υπερεκφρασμένη. Στις πρωτεΐνες θερμικού σοκ συμπεριλαμβάνεται και το γονίδιο *HSPB1*, που εκφράζονται στις μοριακές σύνοδες πρωτεΐνες και είναι απαραίτητες για την κυτταρική ομοιόσταση και προστατεύουν τα κύτταρα από οξειδωτικό στρες (Navarro-Zaragoza et al., 2021).

Ο επίσης υπερεκφρασμένος υποδοχέας CXCR 4 συσχετίζεται με την παρουσία φλεγμονής και απόκρισης στο οξειδωτικό στρες (Bagheri et al., 2018). Το γονίδιο TLR5 επίσης συσχετίζεται με τη φλεγμονή καθώς ενεργοποιεί τις έμφυτες ανοσοαποκρίσεις του οργανισμού και οι υποδοχείς του αναγνωρίζουν διακριτά μοριακά μοτίβα που σχετίζονται με παθογόνα, τα οποία εκφράζονται σε μολυσματικούς παράγοντες. Η ενεργοποίηση των υποδοχέων αυτών ενεργοποιεί ένα πλήθος γονιδίων-στόχων που σχετίζονται με τη φλεγμονή (Zhang et al., 2023). Το γονίδιο αυτό εμπλέκεται και με την κάθαρση των ολιγομερών της α-συνουκλεΐνης δρώντας σε διαφορετικά σημεία σηματοδότησης (Scheiblich et al., 2021).

Τα ανωτέρω υπερεκφρασμένα γονίδια και οι νευροχημικές μεταξύ τους συνάψεις επιβεβαιώνουν την ύπαρξη φλεγμονής και οξειδωτικού στρες στη νόσο Πάρκινσον.

Ομοίως, η βάση αναγνώρισε μόνο τα 169 και οι μεταξύ τους συνδέσεις είναι 186. Από τα αποτελέσματα που εξάγουμε από τη βάση παρατηρούμε ότι η τιμή  $p$  εμπλουτισμού PPI είναι τόσο μικρή ( $<1.0 \times 10^{-16}$ ) που αυτό πρακτικά υποδηλώνει ότι ο παρατηρούμενος αριθμός των μεταξύ τους σχέσεων είναι πολύ σημαντικός και μεγάλος σε σχέση με το εάν διαλέγαμε τυχαία 169 γονίδια από όλο το γονιδίωμα. Ενώ στα τυχαία 169 γονίδια θα περιμέναμε 74 συνδέσεις, η βάση μας επέστρεψε 186. Επομένως, τα γονίδια αυτά δεν είναι τυχαία και ο αριθμός των σχέσεών τους είναι εξίσου σημαντικός, καθώς φαίνεται ότι υπάρχει βιολογική σύνδεση μεταξύ τους. Οι βιολογικές διεργασίες των γονιδίων αυτών σχετίζονται με τη φλεγμονή, όπως αντίστοιχα αποτελέσματα εξάγουμε και με το εργαλείο WebGestalt. Από τα 515 γονίδια συνολικά που εμπλέκονται στη φλεγμονή, 25 από τα 169 που εισήχθησαν βρέθηκαν να σχετίζονται με αυτήν. Με κόκκινο χρώμα παρουσιάζονται και άλλοι σημαντικοί βιολογικοί όροι, λόγω χάρη η απόκριση στο οξειδωτικό στρες, κυτταρικοί υποδοχείς που συμμετέχουν στη μετάδοση σήματος καθώς και άλλοι όροι οι οποίοι δεν έχουν άμεση σχέση με το Πάρκινσον. Στην πλειονότητα τους οι όροι αυτοί αποτελούν νευρολογικούς και νευροχημικούς όρους (**Πίνακας 25**). Ελέγχοντας και τις μοριακές λειτουργίες με τις οποίες εμπλέκονται, εμφανίζονται 5 από τις συνολικά 26 πρωτεΐνες-συννοδούς (chaperone) (**Πίνακας 26**). Συμπερασματικά, τα αποτελέσματά μας

επιβεβαιώνονται και από τα δυο εργαλεία τα οποία χρησιμοποιήσαμε, βάσει του αριθμού των γονιδίων όπου αυτά αναγνώρισαν.



number of nodes: 169	expected number of edges: 74
number of edges: 186	PPI enrichment p-value: < 1.0e-16
average node degree: 2.2	<i>your network has significantly more interactions than expected (what does that mean?)</i>
avg. local clustering coefficient: 0.403	

Εικόνα 14 Δίκτυο αλληλεπιδράσεων πρωτεϊνών- πρωτεϊνών των 269 διαφορεικά υπερ εκφρασμένων γονιδίων με χρήση του προγράμματος STRING

Term description	Observed gene count	Background gene count	FDR
<b>Inflammatory response</b>	25	515	7.99E-08
<b>Response to stress</b>	66	3485	5.65E-07
Response to other organism	35	1256	3.87E-06
Defense response	35	1296	4.51E-06
Response to external stimulus	48	2310	1.02E-05
Tube morphogenesis	23	656	2.72E-05
Defense response to other organism	27	900	3.02E-05
<b>Signaling</b>	69	5239	0.0121
<b>Chaperone-mediated protein folding</b>	5	55	0.0254

Πίνακας 25 Ορισμένες από τις βιολογικές διεργασίες στις οποίες εμπλέκονται τα διαφορετικά υπερ εκφρασμένα και στατιστικά σημαντικά γονίδια με την φλεγμονή να αποτελεί τον κυριότερο εκπροσωπημένο όρο

Term description	Observed gene count	Background gene count	FDR	matching proteins in your network (labels)
<b>Protein folding chaperone</b>	5	26	0.0199	HSPB1,DNAJB1,HSPA6,HSPA1B,HSPA1A

Πίνακας 26 Μοριακός εντοπισμός των γονιδίων και η σημαντική συσχέτιση αυτών με τη νόσο του Πάρκινσον

## 5. Συζήτηση

Ο σκοπός της παρούσας μελέτης είναι η εύρεση γονιδίων που εκφράζονται διαφορετικά σε εγκεφάλους ασθενών με τη νόσο του Πάρκινσον και υγιών ατόμων, με τη χρήση δεδομένων αλληλούχησης επόμενης γενιάς. Μέσω της μεθόδου της μετα-ανάλυσης βελτιώθηκαν οι εκτιμήσεις των 4 μελετών που συμπεριελήφθησαν στην παρούσα και ταυτοποιήθηκαν οι βιολογικές διεργασίες που σχετίζονται με την εμφάνιση του φαινοτύπου. Απώτερος στόχος είναι να ευρεθούν συγκεκριμένοι βιοδείκτες, που θα μπορέσουν να συμβάλλουν περαιτέρω στη διάγνωση, πρόγνωση ή θεραπεία της νόσου και να διευκολυνθεί η παροχή εξατομικευμένης θεραπείας βάσει της γνώσης της φάσης της ασθένειας που βρίσκεται ο ασθενής.



Στην παρούσα μελέτη ταυτοποιήθηκαν 339 διαφορεικά εκφρασμένα γονίδια, 269 υπερεκφρασμένα και 70 υποεκφρασμένα, εκ των οποίων τα περισσότερα, παρά το γεγονός ότι μεταγράφονται, δεν μεταφράζονται σε πρωτεΐνες και πολλά από τα καταγραφόμενα γονίδια δεν έχουν χαρακτηριστεί στο παρελθόν, ως εκ τούτου η λειτουργία τους παραμένει άγνωστη. Επιπλέον, τα αντιπληροφοριακά γονίδια που ανιχνεύθηκαν εμπλέκονται κυρίως στη ρύθμιση της γονιδιακής έκφρασης και πολλά ακόμη γονίδια χαρακτηρίζονται ως ψευδογονίδια, με το συμπέρασμα που προκύπτει να είναι ότι πολλά από τα γονίδια έχουν κυρίως ρυθμιστικό ρόλο. Ωστόσο, βρέθηκαν υπερεκπροσωπημένοι όροι υποεκφρασμένων γονιδίων που στην πλειοψηφία τους σχετίζονται με τη μνήμη, τις νοητικές διαδικασίες, τις συνάψεις και γενικότερα με τη μετάδοση του νευρικού σήματος, καθώς και τη σχέση αυτών με διάφορα χαρακτηριστικά του εγκεφάλου, τα οποία συνάδουν με την νευροεκφυλιστική νόσο του Πάρκινσον. Τα κυριότερα υποεκφρασμένα γονίδια που είχαν σημαντικές σχέσεις μεταξύ τους βρέθηκαν να είναι τα *BDNF*, *TAC1*, *CRH*, *SST* και *CALB1*.

Ο BDNF υποστηρίζει τη λειτουργία του Κεντρικού Νευρικού Συστήματος (ΚΝΣ) και η έκφρασή του διαδραματίζει βασικό ρόλο στην ανάπτυξη του νευρικού συστήματος με ενεργή κυτταρική διαφοροποίηση, νευρωνική ανάπτυξη και συναπτική πλαστικότητα και έχει συνδεθεί τόσο με νευροεκφυλιστικές όσο και νευροψυχιατρικές νόσους, οι οποίες ενδέχεται να προκαλούνται εν μέρει από την ύπαρξη ελαττωμάτων στη συναπτική πλαστικότητα, που σχετίζονται με την ανεπαρκή νευρωνική παροχή BDNF και άλλων νευροτροφικών παραγόντων. Επιπλέον, ο BDNF προάγει τη νευροπροστασία και ενισχύει την επιβίωση των ντοπαμινεργικών νευρώνων, βελτιώνει τη ντοπαμινεργική νευροδιαβίβαση αλλά και την κινητική απόδοση επί νόσου Πάρκινσον. Έχει προταθεί ότι η αύξηση αυτού του παράγοντα, που υποεκφράζεται στη νόσο, σε συμφωνία με την παρούσα μελέτη, μπορεί να αποτελέσει μια στρατηγική όχι μόνο αντιμετώπισης αλλά και πρόληψης της νόσου (Palasz *et al.*, 2020). Επιπλέον, σε συμφωνία με την παρούσα, ο παράγοντας αυτός σχετίζεται και με τη βιολογία των διαταραχών της διάθεσης, ιδιαίτερα της ψυχικής κατάθλιψης (Castren and Kojima, 2017) καταστάσεις που συνοδεύουν τόσο τη νόσο του Πάρκινσον, όσο και το Αλτσχάιμερ και τη νόσο Huntington (Jiang *et al.*, 2019), όπου ο παράγοντας αυτός βρέθηκε σε χαμηλότερα επίπεδα στον ορό του αίματος ασθενών με Πάρκινσον συγκριτικά με υγιά πληθυσμό. Ενδιαφέρον έχει ότι η διαρινική χορήγηση BDNF σχετίζεται με την αύξηση των επιπέδων του μετά από μόλις 30 λεπτά, κυρίως στην περιοχή του ιππόκαμπου και του μεσεγκεφάλου, με επιπλέον συσχέτιση με τη μείωση της απώλειας ντοπαμινεργικών νευρώνων (Jiang *et al.*, 2018).

Ένα ακόμη ενδιαφέρον εύρημα της παρούσας είναι η υποέκφραση της πρωτεΐνης CALB1 που εκτός του ότι σχετίζεται με τη νευροφλεγμονή, μπορεί να αποτελέσει ένα χρήσιμο βιοδείκτη για την εκτίμηση των δυσλειτουργικών νευρώνων επί νόσου Πάρκινσον, καθώς έχει βρεθεί ότι σχετίζεται με ντοπαμινεργική νευρωνική βλάβη σε συνδυασμό με τον μειωμένο αριθμό ολιγοδενδροκυττάρων, στοιχεία που υποστηρίζουν την ύπαρξη τραυματισμού (Liu *et al.*, 2015).

Η υποέκφραση της CRH που αποτελεί κύριο μεσολαβητή της απόκρισης στο στρες σχετίζεται με τη νόσο Πάρκινσον. Ωστόσο, επί αύξησής της μπορούν να αναστραφούν οι καταστροφικές επιδράσεις των γλυκοκορτικοειδών στα νευρικά βλαστοκύτταρα, καθώς η γενετική της ανεπάρκεια παρεμποδίζει τις νευροπροστατευτικές επιδράσεις

της. Συνεπώς, μπορεί να υποστηριχθεί ότι η αύξηση της έκφρασης του γονιδίου αυτού έχει φαρμακολογική εφαρμογή σε ασθένειες που χαρακτηρίζονται από αλλοιωμένη νευρωνική ομοιότητα, όπως η κατάθλιψη, η άνοια αλλά και οι νευροεκφυλιστικές ασθένειες (Koutmani *et al.*, 2013).

Η υποέκφραση της SST, όπως βρέθηκε στην παρούσα, έχει συσχετισθεί όχι μόνο με τη νόσο Αλτσχάιμερ, αλλά και με τις νόσους Πάρκινσον και Huntington καθώς και με την κατάθλιψη, τη διπολική διαταραχή και τη σχιζοφρένεια. Έχει βρεθεί ότι η μείωση των επιπέδων mRNA της SST σε ασθενείς με τη νόσο του Πάρκινσον που φέρουν μεταλλάξεις στο γονίδιο *PARK2*, προκαλεί μια διεγερτική/ανασταλτική ανισορροπία, που σχετίζεται με την εμφάνιση κινητικών και μη κινητικών συμπτωμάτων που παρατηρούνται επί της νόσου (Iwasawa *et al.*, 2019). Επιπλέον, υποστηρίζεται ότι οι βασικές λειτουργίες της SST εμπλέκονται στη ρύθμιση των κυκλωμάτων του εγκεφαλικού φλοιού και στη γνωστική λειτουργία και γι' αυτό το λόγο παρατηρείται μειωμένη έκφρασή της επί νόσων με τέτοια χαρακτηριστικά όπως η νόσος Αλτσχάιμερ, η νόσος Πάρκινσον και η κατάθλιψη, υποστηρίζοντας ότι η αντιμετώπιση αυτής της υποέκφρασης ενδεχομένως να μπορεί να οδηγήσει στην ανάπτυξη νέων φαρμάκων για την αντιμετώπιση νευρολογικών διαταραχών (Song *et al.*, 2021).

Μια επίσης ενδιαφέρουσα θεραπευτική στρατηγική ενδέχεται να είναι η διόρθωση της υποέκφρασης του γονιδίου *TAC1*, που στην παρούσα συσχετίστηκε σημαντικά με τη νόσο Πάρκινσον, καθώς πρόκειται για νευροπεπτίδια των οποίων οι υποδοχείς στο σύστημα της μέλαινας ουσίας και του ραβδωτού διαδραματίζουν σημαντικό ρόλο στην εξέλιξη της νόσου. Αυτό παρατηρείται διότι τα περισσότερα νευροπεπτίδια δρουν νευροπροστατευτικά, μειώνουν το οξειδωτικό στρες που σχετίζεται με τα μιτοχόνδρια και αυξάνουν τη βιογένεση των μιτοχονδρίων αναστέλλοντας ταυτόχρονα τη μικρογλοιακή ενεργοποίηση (Zheng *et al.*, 2021).

Όσον αφορά στα διαφορικά υπερεκφρασμένα γονίδια βρέθηκε ότι αυτά ανέρχονται στα 269 με κυριότερη εκπροσώπηση αυτών που σχετίζονται με τη φλεγμονή, με τα 169 να συνάπτουν 186 σχέσεις μεταξύ τους και στην πλειονότητά τους οι όροι που αναδείχθηκαν σημαντικοί είναι νευρολογικοί και νευροχημικοί. Επιπλέον, 5 από τις πρωτεΐνες-συνοδούς βρέθηκαν να αφορούν τις μοριακές λειτουργίες που εμπλέκονται στη νόσο Πάρκινσον.

Τα κυριότερα υπερεκφρασμένα γονίδια που είχαν σημαντικές σχέσεις μεταξύ τους βρέθηκαν να είναι τα *CSF1*, *IL-18*, *TLR5*, *CXCR4*, *HSPA1B* και *HSPB1*.

Ο ρόλος της φλεγμονής στη νόσο Πάρκινσον αρχίζει και αναδεικνύεται ως ένας σημαντικός παράγοντας εμφάνισης αλλά και εξέλιξης της νόσου και στην παρούσα αρκετά γονίδια που σχετίζονται με τη φλεγμονή αναδείχθηκαν σημαντικά λόγω της υπερέκφρασής τους. Δεν έχουν βρεθεί στατιστικά σημαντικές διαφορές στις αλληλικές και γονοτυπικές συχνότητες του γονιδίου *CSF1*, αλλά υποστηρίζεται ότι υπάρχει μια γενετική συσχέτιση μεταξύ αυτού του γονιδίου και τη νόσο Πάρκινσον. Πιο συγκεκριμένα, ο γονότυπος TT του πολυμορφισμού rs1058885 εμφανίζονταν σε μικρότερη συχνότητα σε ασθενείς με Πάρκινσον σε σύγκριση με την ομάδα ελέγχου και το αλληλόμορφο *CSF1* rs1058885 T βρέθηκε σε μικρότερη συχνότητα σε ασθενείς με Πάρκινσον σε σύγκριση με τα άτομα ελέγχου (Chang *et al.*, 2019).

Η φλεγμονή επί νόσου Πάρκινσον ενισχύεται από την υπερέκφραση της IL-18. Κατά τη διερεύνηση ορισμένων πολυμορφισμών του γονιδίου αυτού δεν βρέθηκε κάποιος συσχετισμός όσον αφορά τη συμβολή της κυρίως στη γνωστική έκπτωση, αλλά λόγω του ότι η IL-18 αποτελεί μια κυτοκίνη που επάγει φλεγμονώδη αντίδραση, ενδέχεται να σχετίζεται με μια σειρά τραυματισμών που μπορεί να οδηγήσουν στην γνωστική έκπτωση επί νόσου Πάρκινσον και στη μακροπρόθεσμη μνήμη (Liu *et al.*, 2016), όπως αναδείχθηκε στην παρούσα.

Εκτός των ανωτέρω, η φλεγμονή ενδεχομένως να συνδέεται και με την υπερέκφραση του γονιδίου *TLR5* που συνδέεται με την ενεργοποίηση του φλεγμονώδους μικρογλοιακού *NLRP3*, το οποίο με τη σειρά του σχετίζεται με την αύξηση των ολιγομερών της  $\alpha$ -συνουκλεΐνης (Scheiblich *et al.*, 2021).

Ένας ακόμη φλεγμονώδης βιοδείκτης που βρέθηκε να σχετίζεται με την ύπαρξη φλεγμονής επί Πάρκινσον είναι η υπερέκφραση του υποδοχέα CXCR4 και σε συμφωνία με προηγούμενες μελέτες (Bagheri *et al.*, 2018) μπορεί να χρησιμοποιηθεί με πολλούς τρόπους στη νόσο Πάρκινσον, όπως για παράδειγμα στην πρόβλεψη για τα συμπτώματα της νόσου όσο και στις θεραπευτικές στρατηγικές για την αντιμετώπισή τους.

Συνεχίζοντας με τη μοριακή συνοδό πρωτεΐνη, όπως βρέθηκε και στην παρούσα, η διαταραγμένη έκφραση και λειτουργία της σχετίζεται με πολλές νευροεκφυλιστικές νόσους, συμπεριλαμβανομένης και της νόσου Πάρκινσον (Joshi *et al.*, 2020) και το ίδιο υποστηρίζεται για το γονίδιο *HSP20* του οποίου οι μεταλλάξεις σχετίζονται με την περιφερική κληρονομική κινητική νευροπάθεια. Επιπλέον, η αυτόνομη δυσλειτουργία λόγω της παρουσίας συσσωματωμάτων της μοριακής συνοδού πρωτεΐνης στο κεντρικό και περιφερικό νευρικό σύστημα, προκαλεί τον εκφυλισμό των ντοπαμινεργικών νευρώνων της μέλαινας ουσίας, οδηγώντας στους σημαντικούς μη κινητικούς φαινοτύπους της νόσου. Λόγω του ότι αυτές οι πρωτεΐνες αποτρέπουν τη συσσώρευση  $\alpha$ -συνουκλεΐνης και την κυτταροτοξικότητα μπορούν να αποτελέσουν κατάλληλο θεραπευτικό στόχο τόσο για την πρόληψη όσο και για την καθυστέρηση εμφάνισης κινητικών και μη συμπτωμάτων της νόσου, λαμβάνοντας υπόψη ότι οι διαθέσιμες θεραπευτικές προσεγγίσεις που βασίζονται στην ντοπαμίνη έχουν αποτύχει ως προς αυτή την κατεύθυνση (Navarro-Zaragoza *et al.*, 2021).

Βάσει των ανωτέρω, είναι σημαντική η διερεύνηση των σπανίων γενετικών μεταλλάξεων επί νόσου Πάρκινσον που σχετίζονται με πάνω από 20 γονίδια (Blauwendraat *et al.*, 2020), όπως αναδείχθηκε και στην παρούσα, δίνοντας όμως έμφαση στις μοναδικές και μάλλον εξατομικευμένες συνάψεις μεταξύ των γονιδίων αυτών. Σημαντική είναι επίσης και η έρευνα αναφορικά με το διαχωρισμό της οικογενούς και της σποραδικής νόσου οι οποίες σχετίζονται με σημειακές μεταλλάξεις στο γονίδιο *SNCA* και τους πολυμορφισμούς του προαγωγού που σχετίζονται με την πρόκληση ανώμαλου σχηματισμού σωμάτων Lewy (Whiffin *et al.*, 2020; Yoshino *et al.*, 2017). Η γενετική έρευνα θα μπορέσει να συμβάλει στην ανάπτυξη μιας θεραπείας που θα λαμβάνει υπόψη της τους διαφορετικούς μηχανισμούς της νόσου βάσει των γενετικών χαρακτηριστικών αλλά και σημαντικών βιοδεικτών από τους ασθενείς. Ωστόσο, επί του παρόντος οι GWAS δεν είναι ακόμη αρκετές προκειμένου να ανακαλυφθούν όλοι οι γενετικοί παράγοντες που συμβάλλουν στην εμφάνιση της νόσου (Blauwendraat *et al.*, 2020; Caruso *et al.*, 2014), αλλά στοχεύουν στη σωστή

κατεύθυνση διότι οι τρέχουσες θεραπείες δεν αντιμετωπίζουν την υποκείμενη αιτία της αλλά στοχεύουν στην επιβράδυνση της εξέλιξής της. Το όφελος μιας γονιδιακής θεραπείας στο εγγύς μέλλον θα είναι ότι θα παρέχει τη δυνατότητα αντιμετώπισης των ελλείψεων των τρεχουσών φαρμακευτικών θεραπειών, διευρύνοντας τη νευροπροστασία και την ιατρική ακρίβεια (Polissidis et al., 2020).

Μια σωστή κατεύθυνση θα μπορούσε να είναι και η αντιμετώπιση της νευροφλεγμονής επί της νόσου, λαμβάνοντας υπόψη ότι η φλεγμονή στην παρούσα βρέθηκε να σχετίζεται με την υπερέκφραση αρκετών γονιδίων. Επιπλέον, βρέθηκε ότι αρκετά γονίδια σχετίζονται με το ρόλο του οξειδωτικού στρες και θα μπορούσαν να αποτελέσουν βιοδείκτες για τη νόσο του Πάρκινσον. Αυτοί οι βιοδείκτες ενδεχομένως να συμβάλλουν στην ανάδειξη νέων αντιοξειδωτικών θεραπειών έναντι των ήδη υπαρχόντων, όπως της βιταμίνης E και του συνενζύμου Q10, η χορήγηση των οποίων επί Πάρκινσον δεν έχει συσχετισθεί με τη μείωση του οξειδωτικού στρες (Chang and Chen, 2020). Θα πρέπει να ληφθεί υπόψη ότι η φλεγμονή αποτελεί έναν εξαιρετικά ρυθμιζόμενο μηχανισμό έναντι παθογόνων ερεθισμάτων ή τραυματισμού των ιστών, προστατεύοντας από ενδεχόμενη βλάβη. Ενώ το ΚΝΣ θεωρείται ως ένας ιστός με προνομιακή ανοσία λόγω του διαχωρισμού του από το περιφερικό ανοσοποιητικό σύστημα μέσω του αιματοεγκεφαλικού φραγμού, στην πραγματικότητα ενδέχεται να προκληθεί μια έμφυτη ανοσολογική απόκριση που σχετίζεται με βλάβη, ιδιαίτερα με χρόνια νευροφλεγμονή, η οποία έχει αρχίσει να αναγνωρίζεται ως παράγοντας εξέλιξης της νόσου (Stolzenberg et al., 2017), γεγονός που υποστηρίζεται από την παρούσα μελέτη, στην οποία βρέθηκε υπερέκφραση σημαντικών γονιδίων που σχετίζονται με αυτή.

Αυτό που προκύπτει από τα ανωτέρω είναι ότι ο στόχος της εξατομικευμένης γονιδιακής θεραπείας έγκειται στη γονιδιακή διόρθωση που επί νόσου Πάρκινσον, θα πρέπει να επικεντρωθεί στα γονίδια που υποεκφράζονται ή υπερεκφράζονται και συσχετίζονται με τη νόσο, όπως αυτά που αναδείχθηκαν στην παρούσα μελέτη. Είναι όμως σημαντικό να σημειωθεί ότι θα πρέπει να λαμβάνονται υπόψη και οι πιθανοί περιβαλλοντικοί παράγοντες ή οι πρόσθετες γενετικές ευαισθησίες προκειμένου η γονιδιακή διόρθωση να αποτελέσει μια ικανοποιητική θεραπευτική στρατηγική για τη νόσο Πάρκινσον (Caruso et al., 2014; Nguyen et al., 2018; Singleton and Hardy, 2011; Toffoli et al., 2020; Whiffin et al., 2020).

## 6. Παράρτημα

Οι 174 μελέτες που πάρθηκαν μετά την αναζήτηση στη βάση BioProject. Με μπλε χρώμα αναπαρίστανται οι μελέτες που συμπεριελήφθησαν στην ανάλυση και πληρούσαν τα κριτήρια επιλεξιμότητας.

---

PRJNA721677	PRJNA588234	PRJNA548929	PRJNA471671	PRJNA354947
PRJNA721657	PRJNA588191	PRJNA544416	PRJNA434419	PRJNA354375
PRJNA718884	PRJNA587794	PRJNA541279	PRJNA434416	PRJNA353642
PRJNA701960	PRJNA579533	PRJNA530121	PRJNA434415	PRJNA350311
PRJNA698999	PRJNA563836	PRJNA526676	PRJNA430219	PRJNA349023
PRJNA693202	<b>PRJNA563007</b>	PRJNA526151	PRJNA417394	PRJNA342487
PRJNA675864	PRJNA560904	PRJNA515626	PRJNA415550	PRJNA339538
<b>PRJNA662780</b>	PRJNA559913	PRJNA515615	PRJNA407941	PRJNA330836
PRJNA655240	PRJNA557595	PRJNA513044	PRJNA394739	PRJNA319440
PRJNA637314	PRJNA557593	PRJNA512523	PRJNA390560	PRJNA312783
PRJNA635098	<b>PRJNA557205</b>	PRJNA506225	PRJNA388551	PRJNA311175
PRJNA630786	PRJNA556869	PRJNA498528	PRJNA387727	PRJNA311174
PRJNA629496	PRJNA556201	PRJNA494382	PRJNA387411	PRJNA311168
PRJNA615805	PRJNA556200	PRJNA493614	PRJNA387280	PRJNA302148
PRJNA608280	PRJNA555099	PRJNA492470	PRJNA382827	PRJNA295431
PRJNA602673	PRJNA552906	PRJNA485495	PRJNA381204	PRJNA293554
PRJNA599014	PRJNA551224	PRJNA484721	PRJNA377152	PRJNA290626
PRJNA597220	PRJNA551119	PRJNA482747	PRJNA369675	PRJNA290524
PRJNA589018	PRJNA549947	PRJNA477337	PRJNA361401	PRJNA289673
PRJNA147055	PRJNA202312	<b>PRJNA283498</b>	PRJNA129807	PRJNA98879
PRJNA144343	PRJNA201039	PRJNA270576	PRJNA129805	PRJNA95655
PRJNA144289	PRJNA191867	PRJNA266233	PRJNA128683	PRJNA93155
PRJNA143623	PRJNA186561	PRJNA264625	PRJNA125581	
PRJNA143167	PRJNA185755	PRJNA262848	PRJNA125567	
PRJNA141657	PRJNA185132	PRJNA246570	PRJNA125557	
PRJNA140647	PRJNA183952	PRJNA236993	PRJNA125401	
PRJNA139443	PRJNA176709	PRJNA236096	PRJNA122363	
PRJNA136075	PRJNA175940	PRJNA231956	PRJNA121173	
PRJNA135681	PRJNA167759	PRJNA231784	PRJNA119215	
PRJNA132861	PRJNA167464	PRJNA230454	PRJNA117397	
PRJNA132387	PRJNA161893	PRJNA228920	PRJNA112099	
PRJNA131367	PRJNA153881	PRJNA225814	PRJNA111757	
PRJNA130803	PRJNA153367	PRJNA221640	PRJNA104051	
PRJNA129845	PRJNA152497	PRJNA218467	PRJNA104049	
PRJNA129839	PRJNA151407	PRJNA213115	PRJNA104047	
PRJNA129837	PRJNA149951	PRJNA212736	PRJNA103753	
PRJNA129811	PRJNA149275	PRJNA209490	PRJNA101445	
PRJNA129809	PRJNA148685	PRJNA208359	PRJNA100241	

**Ο κώδικας του προγράμματος DESeq2 που χρησιμοποιήθηκε στην παρούσα ανάλυση για τον υπολογισμό των διαφορικά εκφρασμένων γονιδίων:**

```
params_title <- "Parkinson vs Control Differential expression with DESeq2"
params_final_dir <- "/data/software/rnaseq/seqc/final"
params_design <- formula("~ panel")
params_contrasts <- list("parkinson_vs_control" = c(factor = "panel", numerator =
"parkinson", denominator = "control"))
params_alpha_threshold <- 0.05
params_lfc_threshold <- 0
params_bcb_file <- "/data/software/rnaseq/seqc/objects/object.rds"
params_data_dir <- "/data/software/rnaseq/seqc/objects/"
params_results_dir <- "/data/software/rnaseq/seqc/results/DifferentialExpression"
#library(tidyverse)
suppressPackageStartupMessages({
  library(goalie)
  library(bcbioRNASeq)
  library(DESeq2)
  library(DESeqAnalysis)
})
object <- bcbioRNASeq(
  uploadDir = params_final_dir,
  interestingGroups = c("panel"),
  organism = "Homo sapiens"
)
saveData(object, dir = params_data_dir)
invisible(lapply(
  X = c(params_data_dir, params_results_dir),
  FUN = initDir
))
file <- params_bcb_file
object <- import(file)
name <- basenameSansExt(file)
```

```

rm(file)
assert(
  is(object, "bcbioRNASeq"),
  validObject(object),
  is.character(name)
)
print(object)
dds <- as(object, "DESeqDataSet")
assert(
  is(dds, "DESeqDataSet"),
  validObject(dds)
)
print(dds)
design(dds) <- params_design
keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep, ]
print(dds)
suppressMessages(dds <- DESeq(dds))
suppressMessages(saveRDS(
  object = dds,
  file = file.path(
    params_data_dir,
    paste0(name, "_dds.rds")
  )
))
dt <- varianceStabilizingTransformation(dds)
assert(is(dt, "DESeqTransform"))
interestingGroups(dt) <- interestingGroups(object)
suppressMessages(saveRDS(
  object = dt,
  file = file.path(

```

```

    params_data_dir,
    paste0(name, "_dt.rds")
  )
))
res_list_unshrunk <- mapply(
  FUN = DESeq2::results,
  contrast = params_contrasts,
  MoreArgs = list(
    "object" = dds,
    "alpha" = params_alpha_threshold,
    "lfcThreshold" = params_lfc_threshold
  ),
  SIMPLIFY = FALSE,
  USE.NAMES = FALSE
)
names(res_list_unshrunk) <- names(params_contrasts)
suppressMessages(saveRDS(
  object = res_list_unshrunk,
  file = file.path(
    params_data_dir,
    paste0(name, "_res_list_unshrunk.rds")
  )
))
suppressMessages(res_list_shrunk <- mapply(
  FUN = DESeqAnalysis::apeglmResults,
  res = res_list_unshrunk,
  contrast = params_contrasts,
  MoreArgs = list(
    "object" = dds
  ),
  SIMPLIFY = FALSE,

```



```

    USE.NAMES = TRUE
  ))
  suppressMessages(saveRDS(
    object = res_list_shrunken,
    file = file.path(
      params_data_dir,
      paste0(name, "_res_list_shrunken.rds")
    )
  ))
  deseq <- DESeqAnalysis(
    data = dds,
    transform = dt,
    results = res_list_unshrunken,
    lfcShrink = res_list_shrunken
  )
  saveRDS(
    object = deseq,
    file = file.path(
      params_data_dir,
      paste0(name, "_deseqanalysis.rds")
    )
  )
  pdf("/data/software/rnaseq/seqc/results/DifferentialExpression/all.pdf")
  invisible(mapply(
    FUN = function(object, i) {
      markdownHeader(text = i, level = 3L, asis = TRUE)
      print(plotMA(object = object, i = i, ntop = 10L))
    },
    i = resultsNames(deseq),
    MoreArgs = list("object" = deseq),
    SIMPLIFY = FALSE,

```

```

    USE.NAMES = FALSE
  ))
dev.off()
pdf("/data/software/rnaseq/seqc/results/DifferentialExpression/volcano.pdf")
invisible(mapply(
  FUN = function(object, i) {
    markdownHeader(text = i, level = 3L, asis = TRUE)
    print(plotVolcano(object = object, i = i, ntop = 10L))
  },
  i = resultsNames(deseq),
  MoreArgs = list("object" = deseq),
  SIMPLIFY = FALSE,
  USE.NAMES = FALSE
))
dev.off()
pdf("/data/software/rnaseq/seqc/results/DifferentialExpression/DEGPCA.pdf")
invisible(mapply(
  FUN = function(object, i) {
    markdownHeader(text = i, level = 3L, asis = TRUE)
    print(plotDEGPCA(object = object, i = i, label = TRUE))
  },
  i = resultsNames(deseq),
  MoreArgs = list("object" = deseq),
  SIMPLIFY = FALSE,
  USE.NAMES = FALSE
))
dev.off()
pdf("/data/software/rnaseq/seqc/results/DifferentialExpression/DEGHeatmap.pdf")
invisible(mapply(
  FUN = function(object, i, ...) {
    markdownHeader(text = i, level = 3L)

```

```

    plotDEGHeatmap(object = object, i = i, ...)
  },
  i = resultsNames(deseq),
  MoreArgs = list(
    "object" = deseq,
    "clusteringMethod" = "ward.D2",
    "scale" = "row"
  ),
  SIMPLIFY = FALSE,
  USE.NAMES = FALSE
))
dev.off()
invisible(mapply(
  FUN = topTables,
  i = resultsNames(deseq),
  MoreArgs = list(
    "object" = deseq,
    "n" = 20L
  ),
  SIMPLIFY = FALSE,
  USE.NAMES = FALSE
))
suppressMessages(res_tbl_list <- resultsTables(deseq, i = 1L))
write.table(res_tbl_list[["all"]], file = file.path(params_results_dir, "ALL.tsv"), row.names =
FALSE, sep = "\t")
write.table(res_tbl_list[["up"]], file = file.path(params_results_dir, "UP.tsv"), row.names =
FALSE, sep = "\t")
write.table(res_tbl_list[["down"]], file = file.path(params_results_dir, "DOWN.tsv"),
row.names = FALSE, sep = "\t")
write.table(res_tbl_list[["both"]], file = file.path(params_results_dir, "BOTH.tsv"), row.names
= FALSE, sep = "\t")

```

## 7. Βιβλιογραφία

Aarsland, D., Pahlhagen, S., Ballard, C.G., Ehrt, U., and Svenningsson, P. (2011). Depression in Parkinson disease--epidemiology, mechanisms and management. *Nat Rev Neurol* 8, 35-47. 10.1038/nrneurol.2011.189.

Aarsland, D., Zaccai, J., and Brayne, C. (2005). A systematic review of prevalence studies of dementia in Parkinson's disease. *Mov Disord* 20, 1255-1263. 10.1002/mds.20527.

Alonso, A., Zaidi, T., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2001). Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc Natl Acad Sci U S A* 98, 6923-6928. 10.1073/pnas.121119298.

Anderson, J.P., Walker, D.E., Goldstein, J.M., de Laat, R., Banducci, K., Caccavello, R.J., Barbour, R., Huang, J., Kling, K., Lee, M., et al. (2006). Phosphorylation of Ser-129 is the dominant pathological modification of alpha-synuclein in familial and sporadic Lewy body disease. *J Biol Chem* 281, 29739-29752. 10.1074/jbc.M600933200.

Andrews, S. (2019). FastQC - A quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Antonini, A., Abbruzzese, G., Barone, P., Bonuccelli, U., Lopiano, L., Onofri, M., Zappia, M., and Quattrone, A. (2008). COMT inhibition with tolcapone in the treatment algorithm of patients with Parkinson's disease (PD): relevance for motor and non-motor features. *Neuropsychiatr Dis Treat* 4, 1-9. 10.2147/ndt.s2404.

Ariga, H., Takahashi-Niki, K., Kato, I., Maita, H., Niki, T., and Iguchi-Ariga, S.M. (2013). Neuroprotective function of DJ-1 in Parkinson's disease. *Oxid Med Cell Longev* 2013, 683920. 10.1155/2013/683920.

Arita, M., Karsch-Mizrachi, I., and Cochrane, G. (2021). The international nucleotide sequence database collaboration. *Nucleic Acids Res* 49, D121-D124. 10.1093/nar/gkaa967.

Armstrong, M.J., and Okun, M.S. (2020a). Choosing a Parkinson Disease Treatment. *JAMA* 323, 1420. 10.1001/jama.2020.1224.

Armstrong, M.J., and Okun, M.S. (2020b). Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* 323, 548-560. 10.1001/jama.2019.22360.

Armstrong, R. (2020). What causes neurodegenerative disease? *Folia Neuropathol* 58, 93-112. 10.5114/fn.2020.96707.

Bagheri, V., Khorramdelazad, H., Hassanshahi, G., Moghadam-Ahmadi, A., and Vakilian, A. (2018). CXCL12 and CXCR4 in the Peripheral Blood of Patients with Parkinson's Disease. *Neuroimmunomodulation* 25, 201-205. 10.1159/000494435.

Baizabal-Carvalho, J.F., and Alonso-Juarez, M. (2020). The Link between Gut Dysbiosis and Neuroinflammation in Parkinson's Disease. *Neuroscience* 432, 160-173. 10.1016/j.neuroscience.2020.02.030.

Baldereschi, M., Di Carlo, A., Rocca, W.A., Vanni, P., Maggi, S., Perissinotto, E., Grigoletto, F., Amaducci, L., and Inzitari, D. (2000). Parkinson's disease and parkinsonism in a longitudinal study: two-fold higher incidence in men. ILSA Working Group. Italian Longitudinal Study on Aging. *Neurology* 55, 1358-1363. 10.1212/wnl.55.9.1358.

Balestrino, R., and Schapira, A.H.V. (2020). Parkinson disease. *Eur J Neurol* 27, 27-42. 10.1111/ene.14108.

Barrett, T., Clark, K., Gevorgyan, R., Gorelenkov, V., Gribov, E., Karsch-Mizrachi, I., Kimelman, M., Pruitt, K.D., Resenchuk, S., Tatusova, T., et al. (2012). BioProject and BioSample databases

at NCBI: facilitating capture and organization of metadata. *Nucleic Acids Res* *40*, D57-63. 10.1093/nar/gkr1163.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testin. *Journal of the Royal Statistical Society* *57*, 289-300.

Berardelli, A., Rothwell, J.C., Thompson, P.D., and Hallett, M. (2001). Pathophysiology of bradykinesia in Parkinson's disease. *Brain* *124*, 2131-2146. 10.1093/brain/124.11.2131.

Bhidayasiri, R. (2005). Differential diagnosis of common tremor syndromes. *Postgrad Med J* *81*, 756-762. 10.1136/pgmj.2005.032979.

Bjorling-Poulsen, M., Andersen, H.R., and Grandjean, P. (2008). Potential developmental neurotoxicity of pesticides used in Europe. *Environ Health* *7*, 50. 10.1186/1476-069X-7-50.

Blauwendraat, C., Nalls, M.A., and Singleton, A.B. (2020). The genetic architecture of Parkinson's disease. *Lancet Neurol* *19*, 170-178. 10.1016/S1474-4422(19)30287-X.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114-2120. 10.1093/bioinformatics/btu170.

Brandt, R., Leger, J., and Lee, G. (1995). Interaction of tau with the neural plasma membrane mediated by tau's amino-terminal projection domain. *J Cell Biol* *131*, 1327-1340. 10.1083/jcb.131.5.1327.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* *34*, 525-527. 10.1038/nbt.3519.

Breckenridge, C.B., Berry, C., Chang, E.T., Sielken, R.L., Jr., and Mandel, J.S. (2016). Association between Parkinson's Disease and Cigarette Smoking, Rural Living, Well-Water Consumption, Farming and Pesticide Use: Systematic Review and Meta-Analysis. *PLoS One* *11*, e0151841. 10.1371/journal.pone.0151841.

Brimblecombe, K.R., Vietti-Michelina, S., Platt, N.J., Kastli, R., Hnieno, A., Gracie, C.J., and Cragg, S.J. (2019). Calbindin-D28K Limits Dopamine Release in Ventral but Not Dorsal Striatum by Regulating Ca(2+) Availability and Dopamine Transporter Function. *ACS Chem Neurosci* *10*, 3419-3426. 10.1021/acscemneuro.9b00325.

Brown, D.A., and London, E. (1998). Functions of lipid rafts in biological membranes. *Annu Rev Cell Dev Biol* *14*, 111-136. 10.1146/annurev.cellbio.14.1.111.

Buermans, H.P., and den Dunnen, J.T. (2014). Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta* *1842*, 1932-1941. 10.1016/j.bbadis.2014.06.015.

Caruso, R., Warner, N., Inohara, N., and Nunez, G. (2014). NOD1 and NOD2: signaling, host defense, and inflammatory disease. *Immunity* *41*, 898-908. 10.1016/j.immuni.2014.12.010.

Castren, E., and Kojima, M. (2017). Brain-derived neurotrophic factor in mood disorders and antidepressant treatments. *Neurobiol Dis* *97*, 119-126. 10.1016/j.nbd.2016.07.010.

Chan, N.C., Salazar, A.M., Pham, A.H., Sweredoski, M.J., Kolawa, N.J., Graham, R.L., Hess, S., and Chan, D.C. (2011). Broad activation of the ubiquitin-proteasome system by Parkin is critical for mitophagy. *Hum Mol Genet* *20*, 1726-1737. 10.1093/hmg/ddr048.

Chang, K.H., and Chen, C.M. (2020). The Role of Oxidative Stress in Parkinson's Disease. *Antioxidants (Basel)* *9*. 10.3390/antiox9070597.

Chang, K.H., Wu, Y.R., Chen, Y.C., Wu, H.C., and Chen, C.M. (2019). Association between CSF1 and CSF1R Polymorphisms and Parkinson's Disease in Taiwan. *J Clin Med* *8*. 10.3390/jcm8101529.

- Chen, J.J., and Marsh, L. (2014). Anxiety in Parkinson's disease: identification and management. *Ther Adv Neurol Disord* 7, 52-59. 10.1177/1756285613495723.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884-i890. 10.1093/bioinformatics/bty560.
- Clark, L.N. (2019). *Molecular Biology* 3rd Edition.
- Clayton, D.F., and George, J.M. (1998). The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. *Trends Neurosci* 21, 249-254. 10.1016/s0166-2236(97)01213-7.
- Corti, O., Lesage, S., and Brice, A. (2011). What genetics tells us about the causes and mechanisms of Parkinson's disease. *Physiol Rev* 91, 1161-1218. 10.1152/physrev.00022.2010.
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561-563. 10.1038/227561a0.
- Crick, F.H. (1958). On protein synthesis. *Symp Soc Exp Biol* 12, 138-163.
- Cuervo, A.M., Stefanis, L., Fredenburg, R., Lansbury, P.T., and Sulzer, D. (2004). Impaired degradation of mutant alpha-synuclein by chaperone-mediated autophagy. *Science* 305, 1292-1295. 10.1126/science.1101738.
- Cummins, C., Ahamed, A., Aslam, R., Burgin, J., Devraj, R., Edbali, O., Gupta, D., Harrison, P.W., Haseeb, M., Holt, S., et al. (2022). The European Nucleotide Archive in 2021. *Nucleic Acids Res* 50, D106-D110. 10.1093/nar/gkab1051.
- Dahodwala, N., Shah, K., He, Y., Wu, S.S., Schmidt, P., Cubillos, F., and Willis, A.W. (2018). Sex disparities in access to caregiving in Parkinson disease. *Neurology* 90, e48-e54. 10.1212/WNL.0000000000004764.
- Dekker, M., Bonifati, V., van Swieten, J., Leenders, N., Galjaard, R.J., Snijders, P., Horstink, M., Heutink, P., Oostra, B., and van Duijn, C. (2003). Clinical features and neuroimaging of PARK7-linked parkinsonism. *Mov Disord* 18, 751-757. 10.1002/mds.10422.
- Demkow, U., and Ploski, R., eds. (2015). *Clinical Applications for Next-Generation Sequencing* (Academic Press).
- Dexter, D.T., and Jenner, P. (2013). Parkinson disease: from pathology to molecular disease mechanisms. *Free Radic Biol Med* 62, 132-144. 10.1016/j.freeradbiomed.2013.01.018.
- Di Fonzo, A., Chien, H.F., Socal, M., Giraudo, S., Tassorelli, C., Iliceto, G., Fabbrini, G., Marconi, R., Fincati, E., Abbruzzese, G., et al. (2007). ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. *Neurology* 68, 1557-1562. 10.1212/01.wnl.0000260963.08711.08.
- Dirnberger, G., and Jahanshahi, M. (2013). Executive dysfunction in Parkinson's disease: a review. *J Neuropsychol* 7, 193-224. 10.1111/jnp.12028.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21. 10.1093/bioinformatics/bts635.
- Dolgacheva, L.P., Berezhnov, A.V., Fedotova, E.I., Zinchenko, V.P., and Abramov, A.Y. (2019). Role of DJ-1 in the mechanism of pathogenesis of Parkinson's disease. *J Bioenerg Biomembr* 51, 175-188. 10.1007/s10863-019-09798-4.
- Dong, X.X., Wang, Y., and Qin, Z.H. (2009). Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. *Acta Pharmacol Sin* 30, 379-387. 10.1038/aps.2009.24.

- Dumitriu, A., Golji, J., Labadorf, A.T., Gao, B., Beach, T.G., Myers, R.H., Longo, K.A., and Latourelle, J.C. (2016). Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease. *BMC Med Genomics* 9, 5. 10.1186/s12920-016-0164-y.
- Durif, F., Debilly, B., Galitzky, M., Morand, D., Viallet, F., Borg, M., Thobois, S., Broussolle, E., and Rascol, O. (2004). Clozapine improves dyskinesias in Parkinson disease: a double-blind, placebo-controlled study. *Neurology* 62, 381-388. 10.1212/01.wnl.0000110317.52453.6c.
- Erkkinen, M.G., Kim, M.O., and Geschwind, M.D. (2018). *Clinical Neurology and Epidemiology of the Major Neurodegenerative Diseases*. Cold Spring Harb Perspect Biol 10. 10.1101/cshperspect.a033118.
- Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047-3048. 10.1093/bioinformatics/btw354.
- Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8, 186-194.
- Gai, W.P., Power, J.H., Blumbergs, P.C., and Blessing, W.W. (1998). Multiple-system atrophy: a new alpha-synuclein disease? *Lancet* 352, 547-548. 10.1016/s0140-6736(05)79256-4.
- Gandhi, K.R., and Saadabadi, A. (2022). Levodopa (L-Dopa). In *StatPearls*.
- Garcia-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F., and Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678-2679. 10.1093/bioinformatics/bts503.
- Gene Ontology, C. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res* 43, D1049-1056. 10.1093/nar/gku1179.
- George, J.M. (2002). The synucleins. *Genome Biol* 3, REVIEWS3002. 10.1186/gb-2001-3-1-reviews3002.
- Georgiev, D., Hamberg, K., Hariz, M., Forsgren, L., and Hariz, G.M. (2017). Gender differences in Parkinson's disease: A clinical perspective. *Acta Neurol Scand* 136, 570-584. 10.1111/ane.12796.
- Goedert, M. (1999). Filamentous nerve cell inclusions in neurodegenerative diseases: tauopathies and alpha-synucleinopathies. *Philos Trans R Soc Lond B Biol Sci* 354, 1101-1118. 10.1098/rstb.1999.0466.
- Goetz, C.G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T., Counsell, C., Giladi, N., Holloway, R.G., Moore, C.G., Wenning, G.K., et al. (2004). Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord* 19, 1020-1028. 10.1002/mds.20213.
- Goldenberg, M.M. (2008). Medical management of Parkinson's disease. *P T* 33, 590-606.
- Grossberg, S. (2017). Acetylcholine Neuromodulation in Normal and Abnormal Learning and Memory: Vigilance Control in Waking, Sleep, Autism, Amnesia and Alzheimer's Disease. *Front Neural Circuits* 11, 82. 10.3389/fncir.2017.00082.
- Guimera, R.V. (2012). bcbio-nextgen: Automated, distributed next-gen sequencing pipeline. *2012* 17. 10.14806/ej.17.B.286.
- Harbers, M., and Carninci, P. (2005). Tag-based approaches for transcriptome research and genome annotation. *Nat Methods* 2, 495-502. 10.1038/nmeth768.

Hatano, T., Kubo, S., Imai, S., Maeda, M., Ishikawa, K., Mizuno, Y., and Hattori, N. (2007). Leucine-rich repeat kinase 2 associates with lipid rafts. *Hum Mol Genet* 16, 678-690. 10.1093/hmg/ddm013.

Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16, 169. 10.1186/s12859-015-0611-3.

Hedrich, K., Djarmati, A., Schafer, N., Hering, R., Wellenbrock, C., Weiss, P.H., Hilker, R., Vieregge, P., Ozelius, L.J., Heutink, P., et al. (2004). DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. *Neurology* 62, 389-394. 10.1212/01.wnl.0000113022.51739.88.

Hely, M.A., Reid, W.G., Adena, M.A., Halliday, G.M., and Morris, J.G. (2008). The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. *Mov Disord* 23, 837-844. 10.1002/mds.21956.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., et al. (2021). Ensembl 2021. *Nucleic Acids Res* 49, D884-D891. 10.1093/nar/gkaa942.

Hruska, K.S., LaMarca, M.E., Scott, C.R., and Sidransky, E. (2008). Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum Mutat* 29, 567-583. 10.1002/humu.20676.

The human Gene Database, Gene Cards. (2023).

Illumina Sequencing. <https://www.1010genome.com/illumina-sequencing/>.

Iwasawa, C., Kuzumaki, N., Suda, Y., Kagawa, R., Oka, Y., Hattori, N., Okano, H., and Narita, M. (2019). Reduced expression of somatostatin in GABAergic interneurons derived from induced pluripotent stem cells of patients with parkin mutations. *Mol Brain* 12, 5. 10.1186/s13041-019-0426-7.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry* 79, 368-376. 10.1136/jnnp.2007.131045.

Jiang, H., Lei, R., Ding, S.W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15, 182. 10.1186/1471-2105-15-182.

Jiang, L., Zhang, H., Wang, C., Ming, F., Shi, X., and Yang, M. (2019). Serum level of brain-derived neurotrophic factor in Parkinson's disease: a meta-analysis. *Prog Neuropsychopharmacol Biol Psychiatry* 88, 168-174. 10.1016/j.pnpbp.2018.07.010.

Jiang, Y., Fay, J.M., Poon, C.D., Vinod, N., Zhao, Y., Bullock, K., Qin, S., Manickam, D.S., Yi, X., Banks, W.A., and Kabanov, A.V. (2018). Nanoformulation of Brain-Derived Neurotrophic Factor with Target Receptor-Triggered-Release in the Central Nervous System. *Adv Funct Mater* 28. 10.1002/adfm.201703982.

Joshi, N., Raveendran, A., and Nagotu, S. (2020). Chaperones and Proteostasis: Role in Parkinson's Disease. *Diseases* 8. 10.3390/diseases8020024.

Kanehisa, M., Goto, S., Kawashima, S., and Nakaya, A. (2002). The KEGG databases at GenomeNet. *Nucleic Acids Res* 30, 42-46. 10.1093/nar/30.1.42.

Kano, O., Ikeda, K., Cridebring, D., Takazawa, T., Yoshii, Y., and Iwasaki, Y. (2011). Neurobiology of depression and anxiety in Parkinson's disease. *Parkinsons Dis* 2011, 143547. 10.4061/2011/143547.



Karsch-Mizrachi, I., Takagi, T., Cochrane, G., and International Nucleotide Sequence Database, C. (2018). The international nucleotide sequence database collaboration. *Nucleic Acids Res* 46, D48-D51. 10.1093/nar/gkx1097.

Khatri, P., Sirota, M., and Butte, A.J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* 8, e1002375. 10.1371/journal.pcbi.1002375.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37, 907-915. 10.1038/s41587-019-0201-4.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14, R36. 10.1186/gb-2013-14-4-r36.

King, M.E., Gamblin, T.C., Kuret, J., and Binder, L.I. (2000). Differential assembly of human tau isoforms in the presence of arachidonic acid. *J Neurochem* 74, 1749-1757. 10.1046/j.1471-4159.2000.0741749.x.

Klein, C., and Westenberger, A. (2012). Genetics of Parkinson's disease. *Cold Spring Harb Perspect Med* 2, a008888. 10.1101/cshperspect.a008888.

Koutmani, Y., Politis, P.K., Elkouris, M., Agrogiannis, G., Kemerli, M., Patsouris, E., Remboutsika, E., and Karalis, K.P. (2013). Corticotropin-releasing hormone exerts direct effects on neuronal progenitor cells: implications for neuroprotection. *Mol Psychiatry* 18, 300-307. 10.1038/mp.2012.198.

Krueger, F. (2012). Trim Galore!

Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E.L., Bohler, A., Melius, J., Waagmeester, A., Sinha, S.R., Miller, R., et al. (2016). WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Res* 44, D488-494. 10.1093/nar/gkv1024.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359. 10.1038/nmeth.1923.

Larsen, J.P., and Tandberg, E. (2001). Sleep disorders in patients with Parkinson's disease: epidemiology and management. *CNS Drugs* 15, 267-275. 10.2165/00023210-200115040-00002.

Lee, A., and Gilbert, R.M. (2016). Epidemiology of Parkinson Disease. *Neurol Clin* 34, 955-965. 10.1016/j.ncl.2016.06.012.

Lee, B.T., Barber, G.P., Benet-Pages, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J.N., Hinrichs, A.S., Lee, C.M., et al. (2022). The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res* 50, D1115-D1122. 10.1093/nar/gkab959.

Lee, D.J., Milosevic, L., Gramer, R., Sasikumar, S., Al-Ozzi, T.M., De Vloo, P., Dallapiazza, R.F., Elias, G.J.B., Cohn, M., Kalia, S.K., et al. (2019). Nucleus basalis of Meynert neuronal activity in Parkinson's disease. *J Neurosurg* 132, 574-582. 10.3171/2018.11.JNS182386.

Lesage, S., and Brice, A. (2009). Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum Mol Genet* 18, R48-59. 10.1093/hmg/ddp012.

Levine, B., and Klionsky, D.J. (2004). Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev Cell* 6, 463-477. 10.1016/s1534-5807(04)00099-1.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079. 10.1093/bioinformatics/btp352.

- Li, P., Ensink, E., Lang, S., Marshall, L., Schilthuis, M., Lamp, J., Vega, I., and Labrie, V. (2020). Hemispheric asymmetry in the human brain and in Parkinson's disease is linked to divergent epigenetic patterns in neurons. *Genome Biol* 21, 61. 10.1186/s13059-020-01960-1.
- Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* 5, 337. 10.1186/1756-0500-5-337.
- Liu, Y., Qiang, M., Wei, Y., and He, R. (2011). A novel molecular mechanism for nitrated {alpha}-synuclein-induced cell death. *J Mol Cell Biol* 3, 239-249. 10.1093/jmcb/mjr011.
- Liu, Y., Zhou, Q., Tang, M., Fu, N., Shao, W., Zhang, S., Yin, Y., Zeng, R., Wang, X., Hu, G., and Zhou, J. (2015). Upregulation of alphaB-crystallin expression in the substantia nigra of patients with Parkinson's disease. *Neurobiol Aging* 36, 1686-1691. 10.1016/j.neurobiolaging.2015.01.015.
- Liu, Z., Guo, J., Wang, Y., Li, K., Kang, J., Wei, Y., Sun, Q., Xu, Q., Xu, C., Yan, X., and Tang, B. (2016). Lack of association between IL-10 and IL-18 gene promoter polymorphisms and Parkinson's disease with cognitive impairment in a Chinese population. *Scientific Reports* 6, 19021. 10.1038/srep19021.
- Lohmann, E., Periquet, M., Bonifati, V., Wood, N.W., De Michele, G., Bonnet, A.M., Fraix, V., Broussolle, E., Horstink, M.W., Vidailhet, M., et al. (2003). How much phenotypic variation can be attributed to parkin genotype? *Ann Neurol* 54, 176-185. 10.1002/ana.10613.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. 10.1186/s13059-014-0550-8.
- Marco-Sola, S., Sammeth, M., Guigo, R., and Ribeca, P. (2012). The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods* 9, 1185-1188. 10.1038/nmeth.2221.
- Mardis, E.R. (2008). Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9, 387-402. 10.1146/annurev.genom.9.081307.164359.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. doi.org/10.14806/ej.17.1.200.
- Masato, A., Plotegher, N., Boassa, D., and Bubacco, L. (2019). Impaired dopamine metabolism in Parkinson's disease pathogenesis. *Mol Neurodegener* 14, 35. 10.1186/s13024-019-0332-6.
- Massano, J., and Bhatia, K.P. (2012). Clinical approach to Parkinson's disease: features, diagnosis, and principles of management. *Cold Spring Harb Perspect Med* 2, a008870. 10.1101/cshperspect.a008870.
- McGregor, M.M., and Nelson, A.B. (2019). Circuit Mechanisms of Parkinson's Disease. *Neuron* 101, 1042-1056. 10.1016/j.neuron.2019.03.004.
- McNaught, K.S., Olanow, C.W., Halliwell, B., Isacson, O., and Jenner, P. (2001). Failure of the ubiquitin-proteasome system in Parkinson's disease. *Nat Rev Neurosci* 2, 589-594. 10.1038/35086067.
- McNaught, K.S., Shashidharan, P., Perl, D.P., Jenner, P., and Olanow, C.W. (2002). Aggresome-related biogenesis of Lewy bodies. *Eur J Neurosci* 16, 2136-2148. 10.1046/j.1460-9568.2002.02301.x.
- Meireles, J., and Massano, J. (2012). Cognitive impairment and dementia in Parkinson's disease: clinical features, diagnosis, and management. *Front Neurol* 3, 88. 10.3389/fneur.2012.00088.
- Mosteller, F., and Bush, R.R. (1954). Selected quantitative techniques. In *Handbook of social psychology: Vol. 1. Theory and Method*, G. Lindzey, ed. (Addison-Wesley), pp. 289–334.

- Murman, D.L. (2015). The Impact of Age on Cognition. *Semin Hear* 36, 111-121. 10.1055/s-0035-1555115.
- Najim al-Din, A.S., Wriekat, A., Mubaidin, A., Dasouki, M., and Hiari, M. (1994). Pallido-pyramidal degeneration, supranuclear upgaze paresis and dementia: Kufor-Rakeb syndrome. *Acta Neurol Scand* 89, 347-352. 10.1111/j.1600-0404.1994.tb02645.x.
- Nalls, M.A., Pankratz, N., Lill, C.M., Do, C.B., Hernandez, D.G., Saad, M., DeStefano, A.L., Kara, E., Bras, J., Sharma, M., et al. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* 46, 989-993. 10.1038/ng.3043.
- Navarro-Zaragoza, J., Cuenca-Bermejo, L., Almela, P., Laorden, M.L., and Herrero, M.T. (2021). Could Small Heat Shock Protein HSP27 Be a First-Line Target for Preventing Protein Aggregation in Parkinson's Disease? *Int J Mol Sci* 22. 10.3390/ijms22063038.
- Newell, K.L., Boyer, P., Gomez-Tortosa, E., Hobbs, W., Hedley-Whyte, E.T., Vonsattel, J.P., and Hyman, B.T. (1999). Alpha-synuclein immunoreactivity is present in axonal swellings in neuroaxonal dystrophy and acute traumatic brain injury. *J Neuropathol Exp Neurol* 58, 1263-1268. 10.1097/00005072-199912000-00007.
- Nguyen, H.P., Van Broeckhoven, C., and van der Zee, J. (2018). ALS Genes in the Genomic Era and their Implications for FTD. *Trends in genetics : TIG* 34, 404-423. 10.1016/j.tig.2018.03.001.
- Noyce, A.J., Bestwick, J.P., Silveira-Moriyama, L., Hawkes, C.H., Giovannoni, G., Lees, A.J., and Schrag, A. (2012). Meta-analysis of early nonmotor features and risk factors for Parkinson disease. *Ann Neurol* 72, 893-901. 10.1002/ana.23687.
- Okido, T., Kodama, Y., Mashima, J., Kosuge, T., Fujisawa, T., and Ogasawara, O. (2022). DNA Data Bank of Japan (DDBJ) update report 2021. *Nucleic Acids Res* 50, D102-D105. 10.1093/nar/gkab995.
- Olson, E.J., Boeve, B.F., and Silber, M.H. (2000). Rapid eye movement sleep behaviour disorder: demographic, clinical and laboratory findings in 93 cases. *Brain* 123 ( Pt 2), 331-339. 10.1093/brain/123.2.331.
- Palasz, E., Wysocka, A., Gasiorowska, A., Chalimoniuk, M., Niewiadomski, W., and Niewiadomska, G. (2020). BDNF as a Promising Therapeutic Agent in Parkinson's Disease. *Int J Mol Sci* 21. 10.3390/ijms21031170.
- Pan, T., Kondo, S., Le, W., and Jankovic, J. (2008). The role of autophagy-lysosome pathway in neurodegeneration associated with Parkinson's disease. *Brain* 131, 1969-1978. 10.1093/brain/awm318.
- Park, J.S., Blair, N.F., and Sue, C.M. (2015). The role of ATP13A2 in Parkinson's disease: Clinical phenotypes and molecular mechanisms. *Mov Disord* 30, 770-779. 10.1002/mds.26243.
- Parkinson, J. (1817). An essay on the shaking palsy (Whittingham and Rowland for Sherwood, Neely and Jones).
- Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14, 417-419. 10.1038/nmeth.4197.
- Pelechano, V., and Steinmetz, L.M. (2013). Gene regulation by antisense transcription. *Nat Rev Genet* 14, 880-893. 10.1038/nrg3594.
- Pinero, J., Ramirez-Anguita, J.M., Sauch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., and Furlong, L.I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* 48, D845-D855. 10.1093/nar/gkz1021.

Polissidis, A., Petropoulou-Vathi, L., Nakos-Bimpos, M., and Rideout, H.J. (2020). The Future of Targeted Gene-Based Treatment Strategies and Biomarkers in Parkinson's Disease. *Biomolecules* *10*. 10.3390/biom10060912.

Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., et al. (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* *276*, 2045-2047. 10.1126/science.276.5321.2045.

Rajawat, Y.S., Hilioti, Z., and Bossis, I. (2009). Aging: central role for autophagy and the lysosomal degradative system. *Ageing Res Rev* *8*, 199-213. 10.1016/j.arr.2009.05.001.

Reese, M.G., Moore, B., Batchelor, C., Salas, F., Cunningham, F., Marth, G.T., Stein, L., Flicek, P., Yandell, M., and Eilbeck, K. (2010). A standard variation file format for human genome sequences. *Genome Biology* *11*, R88. 10.1186/gb-2010-11-8-r88.

Reich, S.G., and Savitt, J.M. (2019). Parkinson's Disease. *Med Clin North Am* *103*, 337-350. 10.1016/j.mcna.2018.10.014.

Reinartz, J., Bruyns, E., Lin, J.Z., Burcham, T., Brenner, S., Bowen, B., Kramer, M., and Woychik, R. (2002). Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. *Brief Funct Genomic Proteomic* *1*, 95-104. 10.1093/bfgp/1.1.95.

Ross, O.A., Braithwaite, A.T., Skipper, L.M., Kachergus, J., Hulihan, M.M., Middleton, F.A., Nishioka, K., Fuchs, J., Gasser, T., Maraganore, D.M., et al. (2008). Genomic investigation of alpha-synuclein multiplication and parkinsonism. *Ann Neurol* *63*, 743-750. 10.1002/ana.21380.

Rubinsztein, D.C. (2006). The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* *443*, 780-786. 10.1038/nature05291.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* *74*, 5463-5467. 10.1073/pnas.74.12.5463.

Sarsour, E.H., Goswami, M., Kalen, A.L., and Goswami, P.C. (2010). MnSOD activity protects mitochondrial morphology of quiescent fibroblasts from age associated abnormalities. *Mitochondrion* *10*, 342-349. 10.1016/j.mito.2010.02.004.

Sayers, E.W., Cavanaugh, M., Clark, K., Pruitt, K.D., Schoch, C.L., Sherry, S.T., and Karsch-Mizrachi, I. (2022). GenBank. *Nucleic Acids Res* *50*, D161-D164. 10.1093/nar/gkab1135.

Scheiblich, H., Bousset, L., Schwartz, S., Griep, A., Latz, E., Melki, R., and Heneka, M.T. (2021). Microglial NLRP3 Inflammasome Activation upon TLR2 and TLR5 Ligation by Distinct alpha-Synuclein Assemblies. *J Immunol* *207*, 2143-2154. 10.4049/jimmunol.2100035.

Schrempf, W., Brandt, M.D., Storch, A., and Reichmann, H. (2014). Sleep disorders in Parkinson's disease. *J Parkinsons Dis* *4*, 211-221. 10.3233/JPD-130301.

Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K., and Suzuki, T. (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet* *25*, 302-305. 10.1038/77060.

Shin, R.W., Iwaki, T., Kitamoto, T., and Tateishi, J. (1991). Hydrated autoclave pretreatment enhances tau immunoreactivity in formalin-fixed normal and Alzheimer's disease brain tissues. *Lab Invest* *64*, 693-702.

Shokralla, S., Spall, J.L., Gibson, J.F., and Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* *21*, 1794-1805. 10.1111/j.1365-294X.2012.05538.x.

- Singh, A., Kukreti, R., Saso, L., and Kukreti, S. (2019). Oxidative Stress: A Key Modulator in Neurodegenerative Diseases. *Molecules* 24. 10.3390/molecules24081583.
- Singleton, A., and Hardy, J. (2011). A generalizable hypothesis for the genetic architecture of disease: pleomorphic risk loci. *Hum Mol Genet* 20, R158-162. 10.1093/hmg/ddr358.
- Smajić, S., Prada-Medina, C.A., Landoulsi, Z., Dietrich, C., Jarazo, J., Henck, J., Balachandran, S., Pachchek, S., Morris, C.M., Antony, P., et al. (2020). Single-cell sequencing of the human midbrain reveals glial activation and a neuronal state specific to Parkinson's disease. medRxiv, 2020.2009.2028.20202812. 10.1101/2020.09.28.20202812.
- Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., and Kasprzyk, A. (2009). BioMart--biological queries made easy. *BMC Genomics* 10, 22. 10.1186/1471-2164-10-22.
- Song, Y.-H., Yoon, J., and Lee, S.-H. (2021). The role of neuropeptide somatostatin in the brain and its application in treating neurological disorders. *Experimental & Molecular Medicine* 53, 328-338. 10.1038/s12276-021-00580-4.
- Spillantini, M.G., Crowther, R.A., Jakes, R., Hasegawa, M., and Goedert, M. (1998). alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proc Natl Acad Sci U S A* 95, 6469-6473. 10.1073/pnas.95.11.6469.
- Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R., and Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature* 388, 839-840. 10.1038/42166.
- Spira, P.J., Sharpe, D.M., Halliday, G., Cavanagh, J., and Nicholson, G.A. (2001). Clinical and pathological features of a Parkinsonian syndrome in a family with an Ala53Thr alpha-synuclein mutation. *Ann Neurol* 49, 313-319.
- Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol* 9, 219-230. 10.1038/nrm2347.
- Stefanis, L. (2012). alpha-Synuclein in Parkinson's disease. *Cold Spring Harb Perspect Med* 2, a009399. 10.1101/cshperspect.a009399.
- Stephenson, J., Nutma, E., van der Valk, P., and Amor, S. (2018). Inflammation in CNS neurodegenerative diseases. *Immunology* 154, 204-219. 10.1111/imm.12922.
- Stolzenberg, E., Berry, D., Yang, D., Lee, E.Y., Kroemer, A., Kaufman, S., Wong, G.C.L., Oppenheim, J.J., Sen, S., Fishbein, T., et al. (2017). A Role for Neuronal Alpha-Synuclein in Gastrointestinal Immunity. *J Innate Immun* 9, 456-463. 10.1159/000477990.
- Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A., and Williams, R.M. (1949). *The American Soldier: Adjustment During Army Life* (Princeton University Press).
- Stuss, D.T. (2011). Functions of the frontal lobes: relation to executive functions. *J Int Neuropsychol Soc* 17, 759-765. 10.1017/S1355617711000695.
- Suzuki, K., Miyamoto, M., Miyamoto, T., Iwanami, M., and Hirata, K. (2011). Sleep disturbances associated with Parkinson's disease. *Parkinsons Dis* 2011, 219056. 10.4061/2011/219056.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43, D447-452. 10.1093/nar/gku1003.
- Toffoli, M., Vieira, S.R.L., and Schapira, A.H.V. (2020). Genetic causes of PD: A pathway to disease modification. *Neuropharmacology* 170, 108022. 10.1016/j.neuropharm.2020.108022.
- Trinh, J., and Farrer, M. (2013). Advances in the genetics of Parkinson disease. *Nat Rev Neurol* 9, 445-454. 10.1038/nrneurol.2013.132.

Valente, E.M., Abou-Sleiman, P.M., Caputo, V., Muqit, M.M., Harvey, K., Gispert, S., Ali, Z., Del Turco, D., Bentivoglio, A.R., Healy, D.G., et al. (2004). Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* 304, 1158-1160. 10.1126/science.1096284.

Valente, E.M., Bentivoglio, A.R., Dixon, P.H., Ferraris, A., Ialongo, T., Frontali, M., Albanese, A., and Wood, N.W. (2001). Localization of a novel locus for autosomal recessive early-onset parkinsonism, PARK6, on human chromosome 1p35-p36. *Am J Hum Genet* 68, 895-900. 10.1086/319522.

Vekrellis, K., Rideout, H.J., and Stefanis, L. (2004). Neurobiology of alpha-synuclein. *Mol Neurobiol* 30, 1-21. 10.1385/MN:30:1:001.

Velayati, A., Yu, W.H., and Sidransky, E. (2010). The role of glucocerebrosidase mutations in Parkinson disease and Lewy body disorders. *Curr Neurol Neurosci Rep* 10, 190-198. 10.1007/s11910-010-0102-x.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. *Science* 270, 484-487. 10.1126/science.270.5235.484.

Wakabayashi, K., Tanji, K., Odagiri, S., Miki, Y., Mori, F., and Takahashi, H. (2013). The Lewy body in Parkinson's disease and related neurodegenerative disorders. *Mol Neurobiol* 47, 495-508. 10.1007/s12035-012-8280-y.

Wang, X., and Michaelis, E.K. (2010). Selective neuronal vulnerability to oxidative stress in the brain. *Front Aging Neurosci* 2, 12. 10.3389/fnagi.2010.00012.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10, 57-63. 10.1038/nrg2484.

Webb, J.L., Ravikumar, B., Atkins, J., Skepper, J.N., and Rubinsztein, D.C. (2003). Alpha-Synuclein is degraded by both autophagy and the proteasome. *J Biol Chem* 278, 25009-25013. 10.1074/jbc.M300227200.

Weingarten, M.D., Lockwood, A.H., Hwo, S.Y., and Kirschner, M.W. (1975). A protein factor essential for microtubule assembly. *Proc Natl Acad Sci U S A* 72, 1858-1862. 10.1073/pnas.72.5.1858.

Whiffin, N., Armean, I.M., Kleinman, A., Marshall, J.L., Minikel, E.V., Goodrich, J.K., Quaipe, N.M., Cole, J.B., Wang, Q., Karczewski, K.J., et al. (2020). The effect of LRRK2 loss-of-function variants in humans. *Nat Med* 26, 869-877. 10.1038/s41591-020-0893-5.

Wirdefeldt, K., Adami, H.O., Cole, P., Trichopoulos, D., and Mandel, J. (2011). Epidemiology and etiology of Parkinson's disease: a review of the evidence. *Eur J Epidemiol* 26 Suppl 1, S1-58. 10.1007/s10654-011-9581-6.

Xicoy, H., Brouwers, J.F., Wieringa, B., and Martens, G.J.M. (2020). Explorative Combined Lipid and Transcriptomic Profiling of Substantia Nigra and Putamen in Parkinson's Disease. *Cells* 9. 10.3390/cells9091966.

Yoshino, H., Hirano, M., Stoessl, A.J., Imamichi, Y., Ikeda, A., Li, Y., Funayama, M., Yamada, I., Nakamura, Y., Sossi, V., et al. (2017). Homozygous alpha-synuclein p.A53V in familial Parkinson's disease. *Neurobiol Aging* 57, 248 e247-248 e212. 10.1016/j.neurobiolaging.2017.05.022.

Youle, R.J., and Narendra, D.P. (2011). Mechanisms of mitophagy. *Nat Rev Mol Cell Biol* 12, 9-14. 10.1038/nrm3028.

Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 33, W741-748. 10.1093/nar/gki475.

Zhang, Z., Liu, Z., Lv, A., and Fan, C. (2023). How Toll-like receptors influence Parkinson's disease in the microbiome-gut-brain axis. *Frontiers in immunology* *14*, 1154626. 10.3389/fimmu.2023.1154626.

Zheng, Y., Zhang, L., Xie, J., and Shi, L. (2021). The Emerging Role of Neuropeptides in Parkinson's Disease. *Front Aging Neurosci* *13*, 646726. 10.3389/fnagi.2021.646726.

Zogopoulos, V.L., Malatras, A., Kyriakidis, K., Charalampous, C., Makrygianni, E.A., Duguez, S., Koutsi, M.A., Pouliou, M., Vasileiou, C., Duddy, W.J., et al. (2023). HGCA2.0: An RNA-Seq Based Webtool for Gene Coexpression Analysis in Homo sapiens. *Cells* *12*. 10.3390/cells12030388.

Νικολάου, Χ., and Χουβαρδάς, Π. (2015). Υπολογιστική Βιολογία (Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις).