

**ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**  
**ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ**



ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ AGRICULTURAL UNIVERSITY OF ATHENS

ΔΙΑΤΡΙΒΗ ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ

**ΒΙΟΛΟΓΙΑΣ ΣΥΣΤΗΜΑΤΩΝ**

**ΤΙΤΛΟΣ**

**ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ ΑΝΑΛΥΣΗ ΠΕΝΗΝΤΑ ΕΝΝΕΑ ΠΡΩΤΕΪΝΩΝ ΜΕ  
ΣΤΟΧΟ ΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΤΟΥΣ ΩΣ ΔΥΝΗΤΙΚΟΙ ΒΙΟΔΕΙΚΤΕΣ ΣΤΟ  
ΠΟΛΛΑΠΛΟ ΜΥΕΛΩΜΑ.**

Όνοματεπώνυμο: Λουκέλης Γ. Κωνσταντίνος

Τόπος ολοκλήρωσης διπλωματικής εργασίας: Αθήνα

Τόποθεσία εργαστηρίου: Εργαστήριο Πρωτεωμικής ΙΙΒΕΑΑ

Έτος ολοκλήρωσης διπλωματικής εργασίας: 2020

**Επίβλεψη:** Ιερώνυμος Ζωιδάκης, μεταδιδακτορικός ερευνητής εργαστηρίου  
Πρωτεομικής, ΙΙΒΕΑΑ (BRFAA)

ΔΙΑΤΡΙΒΗ ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ

ΒΙΟΛΟΓΙΑΣ ΣΥΣΤΗΜΑΤΩΝ

ΤΙΤΛΟΣ

**ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ ΑΝΑΛΥΣΗ ΠΕΝΗΝΤΑ ΕΝΝΕΑ ΠΡΩΤΕΪΝΩΝ ΜΕ ΣΤΟΧΟ ΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΤΟΥΣ ΩΣ ΔΥΝΗΤΙΚΟΙ ΒΙΟΔΕΙΚΤΕΣ ΣΤΟ ΠΟΛΛΑΠΛΟ ΜΥΕΛΩΜΑ.**

TITLE

**BIOINFORMATICS ANALYSIS ON 59 PROTEINS IN ORDER TO VALIDATE THEIR IMPACT AS POTENTIAL BIOMARKERS FOR MULTIPLE MYELOMA IDENTIFICATION**

Όνοματεπώνυμο: Λουκέλης Γ. Κωνσταντίνος

**Επίβλεψη:** Ιερόνυμος Ζωιδάκης, μεταδιδακτορικός ερευνητής εργαστηρίου Πρωτεομικής, ΙΙΒΕΑΑ (BRFAA)

**Τριμελής επιτροπή:**

- Γεράσιμος Δάρας, Επίκουρος Καθηγητής τμήματος Βιοτεχνολογίας ΓΠΑ
- Πολυδεύκης Χατζόπουλος, Καθηγητής τμήματος Βιοτεχνολογίας ΓΠΑ
- Σταμάτης Ρήγας, Επίκουρος Καθηγητής τμήματος Βιοτεχνολογίας ΓΠΑ

## **Περίληψη**

Καθώς η ιατρική επιστήμη σημειώνει καθημερινά σημαντικές εξελίξεις όσον αφορά στην ανάπτυξη διαγνωστικών και θεραπευτικών τεχνικών που βοηθούν στην πρόληψη και αντιμετώπιση προβλημάτων υγείας, ο καρκίνος συνεχίζει για δεκαετίες να μονοπωλεί το ενδιαφέρον των ερευνητών λόγω του δυσνόητου και πολύπλοκου βιολογικά χαρακτήρα του. Επιπλέον, η επιμόλυνση του περιβάλλοντος, η βιομηχανοποίηση της διατροφής του σύγχρονου ανθρώπου και οι διάφορες επιβλαβείς συνήθειες όπως το κάπνισμα, έχουν οδηγήσει τα τελευταία πενήντα χρόνια σε εκθετική αύξηση τη συχνότητα εμφάνισης καρκίνου. Ειδικότερα, το πολλαπλό μύελωμα είναι ένας τύπος καρκίνου που αφορά σε μια σχετικά σπάνια μορφή κακοήθειας και προσβάλλει κυρίως το μυελό των οστών ενώ η δράση του συνδέεται με εξάπλωση σε άλλα όργανα σε πάνω από 90% των περιπτώσεων οδηγώντας τελικά στο θάνατο. Πρόκειται για ένα είδος καρκίνου το οποίο τις περισσότερες φορές δεν συνοδεύεται από εκδήλωση συμπτωμάτων ενώ όταν αυτά αρχίζουν να παρουσιάζονται, συνήθως έχει ήδη προηγηθεί μετάσταση και πλέον ο καρκίνος είναι μη χειρουργήσιμος και μη ανταποκρινόμενος σε φαρμακευτική αγωγή. Μια εναλλακτική προσέγγιση είναι η όσο το δυνατόν πιο έγκαιρη διάγνωση, κάτι που μπορούν να προσφέρουν οι βιοδείκτες, βιομόρια δηλαδή τα οποία είναι ενδεικτικά για την ύπαρξη μιας ασθένειας ενώ συνάμα χαρακτηρίζονται από υψηλή ειδικότητα και ευαισθησία για αυτήν. Ο σχεδιασμός αλλά και οι κλινικές δοκιμές που απαιτούνται για την αναγνώριση της αξίας και της εγκυρότητας των εκάστοτε βιοδεικτών αποτελεί ένα δύσκολο πόνημα, ιδίως μάλιστα όταν υπάρχουν αναρίθμητοι τροποποιητικοί παράγοντες όπως το φύλο, η ηλικία και το γενετικό υπόβαθρο, που θα μπορούσαν να είναι υπεύθυνοι για την υποβάθμιση της σημασίας τους. Σε αυτήν τη διπλωματική εργασία παρουσιάζονται σε λεπτομέρεια οι διάφορες πτυχές του πολλαπλού μυελώματος με εξέχοντα ρόλο να καταλαμβάνουν η γενετική και η παθοφυσιολογία της ασθένειας, καθώς και η στατιστική αξία των επιδημιολογικών ερευνών. Επίσης, συζητούνται οι διαφορετικές μορφές προσεγγίσεων και θεραπείας του προβλήματος καθώς επίσης και η αξία των βιοδεικτών στη σύγχρονη ιατρική εν γένει και ιδίως στον καρκίνο. Ο κύριος στόχος της παρούσας εργασίας είναι να συγκρίνει ένα δείγμα 59 πρωτεϊνών ως προς την αξία τους ως δυνητικοί βιοδείκτες, οι οποίες προέκυψαν από ένα δείγμα ασθενών με πολλαπλό μύελωμα και στους οποίους είχε χορηγηθεί το φάρμακο εκλογής μπορτεζομίμη όπου πρόκειται για αναστολέα πρωτεασώματος. Η σύγκριση διενεργείται με τη χρήση μιας σειράς διαφορετικών βιοπληροφορικών προγραμμάτων (Cluego, PCA, STRING, Heatmap) με στόχο την ποσοτικοποίηση αλλά και οπτικοποίηση των δεδομένων μας και ως εκ τούτου, την πιθανή ανάδειξη ορισμένων πρωτεϊνών από αυτές ως κλινικά χρήσιμους βιοδείκτες.

**Επιστημονικό πεδίο: Βιοτεχνολογία/Βιοπληροφορική/Πρωτεωμική**

**Λέξεις κλειδιά: Πολλαπλό μύελωμα, καρκίνος, βιοπληροφορική, βιοτεχνολογία, βιοδείκτες, σύγχρονη ιατρική διαγνωστική, πρωτεωμική, πρωτεΐνες**

# **BIOINFORMATICS ANALYSIS ON 59 PROTEINS IN ORDER TO VALIDATE THEIR IMPACT AS POTENTIAL BIOMARKERS FOR MULTIPLE MYELOMA IDENTIFICATION**

## **Abstract**

As medical science keeps on expanding upon the development of new diagnostic and therapeutic techniques that allow to prevent and treat health problems, cancer has been monopolizing researchers' interest for decades due to its difficult and complex biological nature. In addition, environmental pollution, the industrialization of the modern human diet and various other harmful habits such as smoking have led to an increase in the incidence of cancer in the last fifty years. In particular, multiple myeloma is a type of cancer that involves a relatively rare form of malignancy and mainly affects the bone marrow while it continues to spread to other organs in more than 90% of cases, eventually leading to mortality. It is a type of cancer that is often not accompanied by symptoms, and when that occurs, it is usually preceded by metastasis and the cancer is now non-surgical and unresponsive to medication. An alternative approach is to diagnose the cancer as early as possible, which can be done with the use of biomarkers, biomolecules that are indicative of a disease and at the same time are characterized by high specificity and sensitivity for that particular disease. The design and clinical trials that are required to identify the value and validity of each biomarker are both a difficult task, especially when there are countless modifying factors such as gender, age and genetic background, which could be responsible for degrading their importance. In this dissertation, the various aspects of multiple myeloma are presented in detail, with genetics and pathophysiology information about the disease, as well as the statistical value of epidemiological research. In addition, the different forms of approach and treatment practices are discussed meticulously as well as the value of biomarkers in modern medicine in general and cancer in particular. The main aim of this study is to compare a sample of 59 proteins in terms of their value as potential biomarkers, which emerged from a sample of patients with multiple myeloma who were treated with the drug of choice bortezomib, a protease inhibitor. The comparison is conducted using a series of different bioinformatics programs (Cluego, PCA, STRING, Heatmap) with the aim of quantifying and visualizing our data and therefore, measuring the impact of some of the studied proteins as biomarkers for multiple myeloma.

Research Field: **Biotechnology/Bioinformatics/Proteomics**

Keywords: **Multiple myeloma, cancer, bioinformatics, biotechnology, biomarkers, modern medical diagnostics, proteomics, proteins**

## ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1 – ΠΟΛΛΑΠΛΟ ΜΥΕΛΩΜΑ.....	5
1.1 Εισαγωγή – Βασικά στοιχεία καρκίνου.....	5
1.2 Στατιστικά στοιχεία εμφάνισης καρκίνου.....	9
1.3 Πολλαπλό μυέλωμα – Βασικές Αρχές.....	11
1.4 Επιδημιολογικά στοιχεία πολλαπλού μυελώματος.....	12
1.5 Παθοφυσιολογία και μοριακή βάση πολλαπλού μυελώματος.....	15
1.6 Γενετική του πολλαπλού μυελώματος.....	18
1.7 Συμπτώματα, κριτήρια διαφοροδιάγνωσης και θεραπεία.....	25
1.8 Βιοδείκτες στην Πρωτεομική του καρκίνου.....	30
ΚΕΦΑΛΑΙΟ 2 - ΠΡΟΓΡΑΜΜΑΤΑ ΤΗΣ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ ΑΝΑΛΥΣΗΣ.....	33
2.1 Βιοπληροφορική και Πρωτεομική.....	33
2.2 Cytoscape.....	38
2.3 Cluego.....	42
2.4 STRING.....	44
2.5 Heatmap.....	47
2.6 PCA.....	51
ΚΕΦΑΛΑΙΟ 3 – ΑΠΟΤΕΛΕΣΜΑΤΑ.....	54
3.1 Excel Datase .....	54
3.2 Λειτουργίες πρωτεϊνών.....	56
3.3 Cluego Αποτελέσματα.....	62

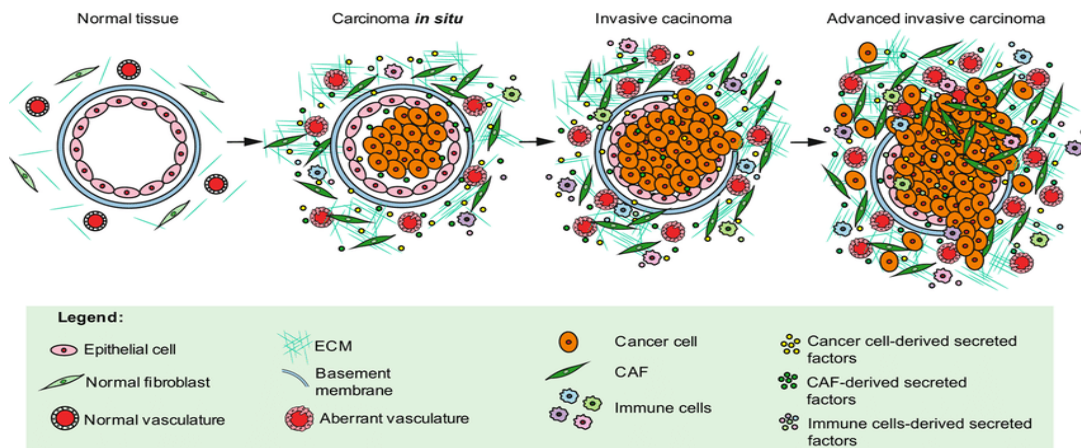
<b>3.4 STRING Αποτελέσματα.....</b>	<b>67</b>
<b>3.5 Heatmap Αποτελέσματα.....</b>	<b>71</b>
<b>3.6 PCA Αποτελέσματα.....</b>	<b>72</b>
<b>ΚΕΦΑΛΑΙΟ 4 – ΣΥΜΠΕΡΑΣΜΑΤΑ.....</b>	<b>76</b>
<b>ΚΕΦΑΛΑΙΟ 5 – ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ.....</b>	<b>77</b>

## **ΚΕΦΑΛΑΙΟ 1                      ΠΟΛΛΑΠΛΟ ΜΥΕΛΩΜΑ**

### **1.1 Εισαγωγή - Βασικά στοιχεία καρκινικού όγκου**

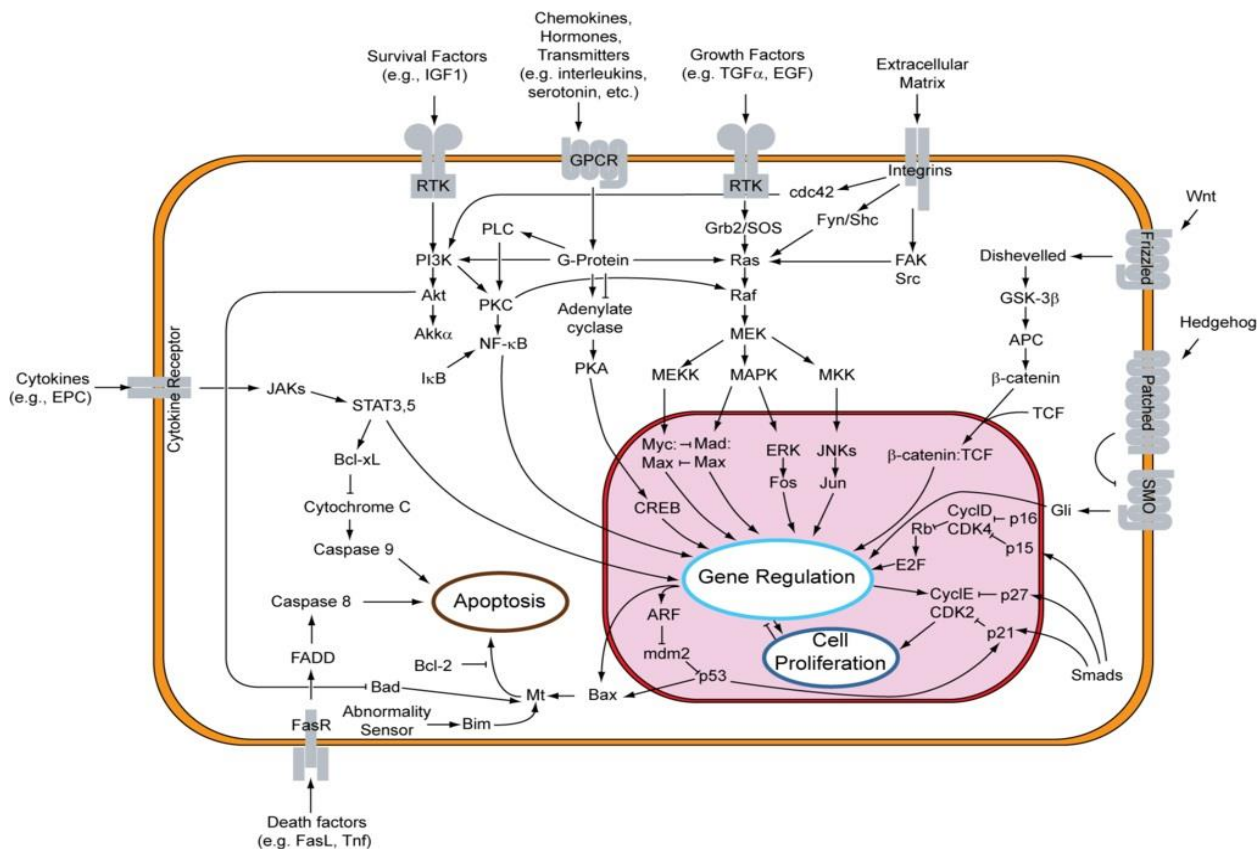
Ο καρκίνος θεωρείται αναμφισβήτητα η κύρια θανάσιμη ασθένεια της εποχής που διατρέχουμε και όχι μη δικαιολογημένα. Τεράστιος αριθμός ερευνών διενεργείται αυτή τη στιγμή προκειμένου να διερευνηθούν τα διάφορα πιθανά αίτια τα οποία σχετίζονται με την εμφάνιση του. Μεταξύ αυτών κύριο ρόλο διαδραματίζουν η διατροφή, το γενετικό υπόβαθρο, οι συνθήκες διαβίωσης, το κάπνισμα καθώς επίσης και το γηγενές περιβάλλον. Και ενώ έχουν γίνει σημαντικά βήματα τόσο στην έγκαιρη ανίχνευση ορισμένων ειδών καρκίνου όσο και στην αντιμετώπιση αυτών, συνεχίζει να δεσπόζει σε πρωταρχική θέση στα ποσοστά θνησιμότητας παγκοσμίως και ο λόγος είναι ότι πρόκειται για νόσημα πολυπαραγοντικό και ταχέως μεταβαλλόμενο, πολλές φορές δυσκολεύοντας την εξάλειψη του με τα σύγχρονα συμβατικά κλινικά μέσα. Αυτό οφείλεται στο γεγονός πως ο καρκίνος χαρακτηρίζεται από μεγάλη τυχαιότητα στην εμφάνιση και την εξέλιξη του ενώ ακόμη και οι πιο παγιωμένες θεραπευτικές οδοί δεν πετυχαίνουν πάντα την πλήρη εξυγίανση του οργανισμού. Έτσι το σύνθετο και διαπλέκον περιβάλλον εμφάνισης της ασθένειας και η ανταπόκριση των νοσούντων οργανισμών κατά τρόπο μοναδικό στα υπάρχοντα φαρμακολογικά σχήματα, δυσχεραίνει το έργο των εκάστοτε ιατρικών ομάδων.

Για να καταλάβει κανείς το πολυδαίδαλο του καρκίνου, θα πρέπει να κατανοήσει τα βασικά στάδια μέσα από τα οποία ο καρκίνος γεννάται και τελικά οδηγείται στην εξάπλωση του στο σώμα.



**Εικόνα 1.1** *Cancer Associated Fibroblasts: The Architects of Stroma Remodeling. PROTEOMICS, 18(5-6) Βήματα ανάπτυξης όγκου:* Αρχικά ο όγκος περιορίζεται γεωγραφικά εντός του κυττάρου. Καθώς αποκτά ολοένα και μεγαλύτερο μέγεθος, η μηχανική πίεση που ασκεί εκ των έσω προκαλεί ρήξη της κυτταρικής μεμβράνης, λύση του κυττάρου και διασπορά του όγκου μέσω στοιχείων του αίματος και της λέμφου σε απομακρυσμένα σημεία. [1]

Από την εικόνα 1.1 μπορούμε να εξάγουμε κάποια βασικά συμπεράσματα για το πως πραγματοποιείται η ανάπτυξη και η διασπορά του όγκου στο σώμα. Υπάρχουν δύο κατηγορίες όγκων με βάση την επικινδυνότητά τους, οι καλοήθεις και οι κακοήθεις. Ως καλοήθης χαρακτηρίζεται ένας όγκος όταν αυτός παραμένει στο σημείο γέννησής του, δεν έχει δηλαδή μεταναστεύσει μέσω της λέμφου ή μέσω του αίματος σε κοντινούς ή μακρινούς στόχους. Σε αυτό το στάδιο, ο καρκίνος θεωρείται γενικά ακίνδυνος και η εκτομή αυτού, δεδομένου βέβαια ότι πληροί τις προϋποθέσεις για χειρουργείο, με παράλληλη συντηρητική μεταχειρουργική φαρμακευτική περίθαλψη και τακτικό έλεγχο, συνήθως έχει θετική έκφραση για τον ασθενή. Από την άλλη, η κατάσταση περιπλέκεται όταν ο όγκος αρχίζει να μετακινείται και να διασπείρεται στους γύρω ιστούς όπου και χαρακτηρίζεται πλέον ως κακοήθης. Προκειμένου να συμβεί αυτό, οφείλουν να προηγηθούν ορισμένα γεγονότα.[2] Κάθε κύτταρο προκειμένου να επιβιώσει και να αναπτυχθεί, οφείλει να συναισθάνεται την περιρρέουσα ατμόσφαιρά του μέσω κάποιων μηχανισμών επικοινωνίας. Υπάρχουν πάρα πολλοί τέτοιοι μηχανισμοί με ορισμένους από αυτούς να παρουσιάζονται στην εικόνα 1.2:



**Εικόνα 1.2** *Tumor viruses and cancer biology: Modulating signaling pathways for therapeutic intervention*  
 Τεχνικές κυτταρικής σηματοδότησης: Από την εικόνα διαφαίνεται η πολυπλοκότητα των σηματοδοτικών μονπατιών και το πλήθος των συνδυασμών τους. [2]

Από την εικόνα καθίσταται πασιφανής η πολυπλοκότητα των σημάτων που δέχεται ένα κύτταρο τόσο σε αριθμό όσο και ως προς το μέγεθος των δυνητικά διαφορετικών συνδυασμών αυτών. Αυτό είναι ένα πολύ λεπτό σημείο κατανόησης το οποίο από μόνο του περιπλέκει τη διερεύνηση των αιτιών της νόσου καθώς η μελέτη ενός μονοπατιού σηματοδότησης και μόνο πολλές φορές δεν επαρκεί για την κατανόηση του ολιστικού μηχανισμού που υποκρύπτεται. Κάθε ένα από αυτά τα σηματοδοτικά μονοπάτια εμπλέκεται σε σειρά λειτουργιών απαραίτητων για τη λειτουργία του κυττάρου και το διαπλέξιμο αυτών συνεισφέρει σε μοναδικούς μηχανισμούς.[2] Αντίστοιχα, τεράστια είναι και η ποικιλία των διαφορετικών υποδοχέων για όλα αυτά τα σήματα που διαθέτει στη φαρέτρα του το κάθε είδος κυττάρου. Μερικοί από αυτούς τους μηχανισμούς σχετίζονται με την επιβίωση, τη διαίρεση, τον πολλαπλασιασμό, τη μετανάστευση αλλά και την αντοχή του κυττάρου σε δυσμενές περιβάλλον. Παράλληλα, υπάρχει μια σειρά στοιχείων εξωτερικά του κυττάρου απαρτιζόμενων κυρίως από κολλαγόνο, γλυκοζαμινογλυκάνες, πρωτεογλυκάνες και γλυκοπρωτεΐνες τα οποία συνιστούν τη βασική στιβάδα που με τη σειρά της επιτρέπει στο κύτταρο τόσο την επικοινωνία όσο και την ανάπτυξη μηχανικής αντοχής στο περιορισμένο περιβάλλον του καθώς επίσης και δομές που βρίσκονται εσωτερικά όπως οι χασματοσύνδεσμοι που προσφέρουν τη δυνατότητα δημιουργίας συμβιωτικής



σχέσης μεταξύ δύο κυττάρων μέσω δομών που αναφέρονται ως κοννεξόνια, δίνοντας τους την ευκαιρία να συζεύξουν τις βιοχημείες τους. Οποιαδήποτε παρέκκλιση από τη σωστή λειτουργία έστω και ενός από όλα τα επιμέρους στοιχεία εγκυμονεί κινδύνους για ανάπτυξη καρκίνου διακυμεινόμενης σοβαρότητας, ανάλογα με το στάδιο που πλήττεται κάθε φορά. Συνεπώς ο καρκίνος τις περισσότερες φορές αποτελεί απότοκο είτε ανωμαλιών κατά την κυτταρική σηματοδότηση είτε της ελαττωματικής λειτουργίας των διαφόρων δομών του.

Όσον αφορά στους κακοήθεις όγκους, όπως αναφέρθηκε, βασική προϋπόθεση ανάπτυξής τους αποτελεί η μετανάστευση του όγκου. Όπως φαίνεται στην εικόνα 1.1, ύστερα από την ανάπτυξη του όγκου που παρουσιάζεται στο βήμα 2[1] που ακόμα χαρακτηρίζεται ως καλοήθης και αφού λάβουν χώρα ορισμένα γεγονότα που θα εκτρέψουν τη φυσιολογική σηματοδότηση, τελικά ο όγκος μετά τη μηχανική διάσπαση του κυττάρου στο οποίο γεννήθηκε είναι ελεύθερος να πλοηγηθεί μέσω των στοιχείων του αίματος ή της λέμφου σε άλλους στόχους και να δημιουργήσει τις πρώτες εστίες εξάπλωσης του καρκίνου. Καθώς νέοι όγκοι αναπτύσσονται, συνήθως υπάρχει και αύξηση αιματικής ροής προκειμένου να διασφαλίζεται η απαραίτητη οξυγόνωση των εστιών με αποτέλεσμα την υπέρμετρη αγγειογένεση. Σε αυτό το σημείο η κατάσταση αρχίζει να δυσκολεύει το θεραπευτικό σχήμα και αυτό είναι το στάδιο που μπορεί να οδηγήσει στο θάνατο.

Στη συνέχεια ακολουθεί μια σειρά πινάκων όπου παρουσιάζονται τα βασικά στοιχεία που χαρακτηρίζουν έναν όγκο ως προς το μέγεθος και το βαθμό εξάπλωσης του και που εν γένει χρησιμοποιούνται ως γνώμονας για την ταξινόμηση του σε στάδια:

**Πίνακας 1.1** MD.Holland-Frei Cancer Medicine. 6th edition. Κεφάλαιο αξιολόγησης σταδίου καρκινικού όγκου [3]

### Ταξινόμηση όγκου

• TX: Ο πρωτοπαθής όγκος δεν μπορεί να εκτιμηθεί
• T0: Δεν υπάρχουν ενδείξεις πρωτοπαθούς όγκου
• Tis: Καρκίνωμα in situ
• Tis (DCIS): Καρκίνωμα Ductal(πορώδες)
• Tis (LCIS): λοβιακό καρκίνωμα in situ
• T1: Όγκος 2,0 cm ή λιγότερο
• T1mic:Μικρή διήθηση όγκου 0,1 cm ή λιγότερο
• T1a: Όγκος μεγαλύτερος από 0,1 αλλά όχι μεγαλύτερος από 0,5 cm
• T1b: Όγκος μεγαλύτερος από 0,5 cm αλλά όχι μεγαλύτερος από 1,0 cm
• T1c: Όγκος μεγαλύτερος από 1,0 cm αλλά όχι μεγαλύτερος από 2,0 cm
• T2: Όγκος μεγαλύτερος από 2,0 cm αλλά όχι μεγαλύτερος από 5,0 cm
• T3: Όγκος μεγαλύτερος από 5,0 cm
• T4: Όγκος οποιουδήποτε μεγέθους με άμεση επέκταση σε θώρακα ή δέρμα
• T4a: Επέκταση στο θωρακικό τοίχωμα, μη συμπεριλαμβανομένου του θωρακικού μυός
• T4b: Οίδημα (συμπεριλαμβανομένου peau d'orange) ή εξέλκωση του δέρματος
• T4c: Τόσο T4a όσο και T4b
• T4d: Φλεγμονώδες καρκίνωμα

**Πίνακας 1.2** MD.Holland-Frei Cancer Medicine. 6th edition. Κεφάλαιο αξιολόγησης καρκινικού όγκου[3]

### Ταξινόμηση λεμφαδένων

• NX: Οι περιφερειακοί λεμφαδένες δεν μπορούν να εκτιμηθούν.(π.χ., έχουν αφαιρεθεί προηγουμένως)
• N0: Δεν υπάρχουν περιφερειακές μεταστάσεις λεμφαδένων.
• N1: Μεταστάσεις σε κινητά μασχαλιαία λεμφογάγγλια ομοπλευρικά.
• N2: Μεταστάσεις σε μασχαλιαία λεμφογάγγλια που είναι σταθεροποιημένα και βρίσκονται ομοπλευρικά ή σε κλινικά εμφανή λεμφογάγγλια έσω μαζικού αδένα ομοπλευρικά με ταυτόχρονη απουσία σε μασχαλιαία.
• N3:A) Μεταστάσεις σε υποκλείδιους λεμφαδένες ομοπλευρικά με ή χωρίς μεταστάση σε μασχαλιαίους λεμφαδένες ή Β)σε κλινικώς εμφανή εσωτερικά μαστικά λεμφογάγγλια ομοπλευρικά με ταυτόχρονη παρουσία κλινικώς εμφανώς μεταστάσεων σε μασχαλιαίους λεμφαδένες.Γ)Μετάσταση σε ομόπλευρους υπερκλείδιους λεμφαδένες με ή χωρίς εμπλοκή μασχαλιαίας ή μαστικής λεμφαδενικής μετάστασης.

**Πίνακας 1.3** MD.Holland-Frei Cancer Medicine. 6th edition. Κεφάλαιο αξιολόγησης σταδίου καρκινικού όγκου [3]

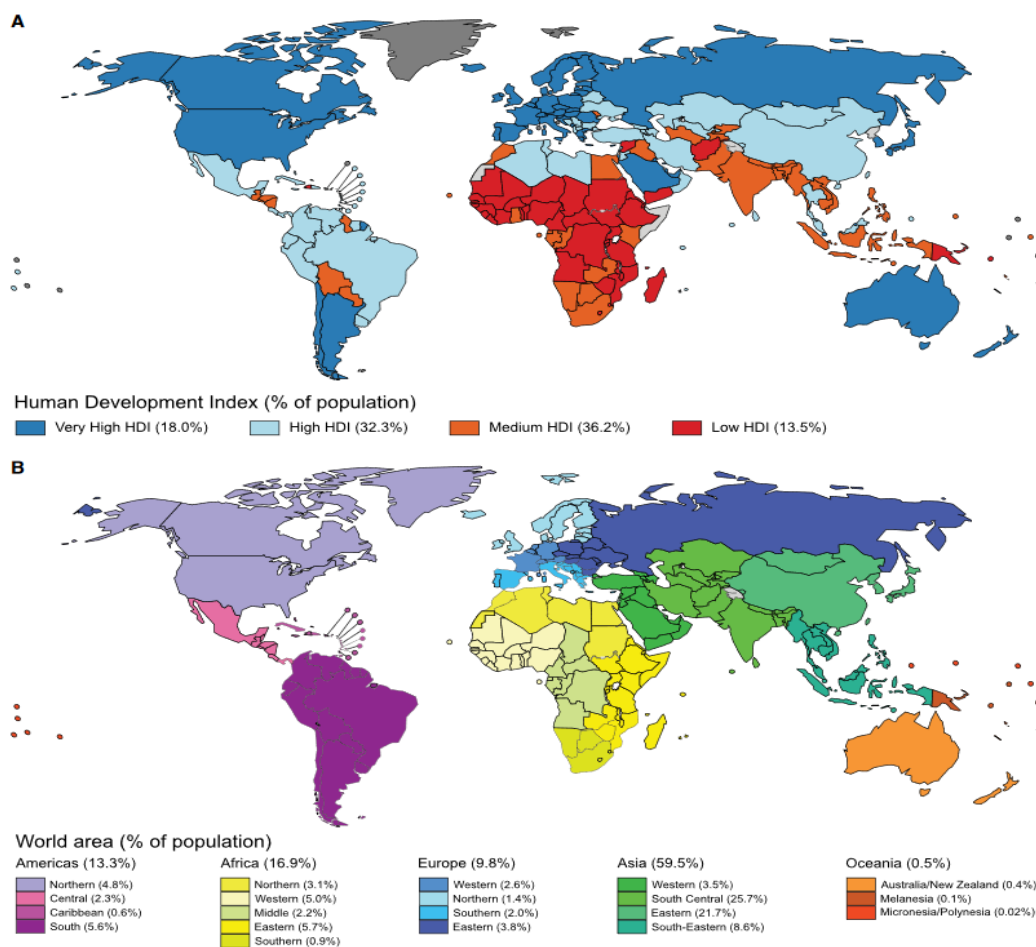
### Ταξινόμηση μεταστάσεων

• MX: Η ύπαρξη μακρινών μεταστάσεων δεν μπορεί να εκτιμηθεί
• M0: Δεν υπάρχει μακρινή μετάσταση
• M1: Υπάρχει μακρινή μετάσταση

Με βάση τα τρία αυτά στοιχεία ο όγκος κατηγοριοποιείται ως TNM(X,Y,Z) όπου X,Y,Z ο βαθμός επέκτασης όγκου, Y η λεμφαδενική εξάπλωση και Z η μεταστατική κατάσταση και ξεκινά στη συνέχεια η διαδικασία αντιμετώπισης του.

## 1.2 Στατιστικά στοιχεία εμφάνισης καρκίνου

Ακολουθεί μια εικόνα στην οποία παρουσιάζεται η γενική κατάσταση της εξάπλωσης της ασθένειας στις διάφορες ηπείρους:



**Εικόνα 1.3** *Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, CA: A Cancer Journal for Clinicians.* Στο τμήμα A παρουσιάζεται η διακύμανση του καρκίνου με βάση το βιοτικό επίπεδο ενώ το τμήμα B προκύπτει από στοιχεία με μοναδικό γνώμονα τη γεωγραφική διασπορά του πληθυσμού. [4]

Από την εικόνα 1.3 προκύπτουν κάποια σημαντικά συμπεράσματα. Εν πρώτοις, όσον αφορά στην εικόνα 1.3A, [4] βλέπουμε πως οι γεωγραφικές περιοχές μέγιστου βιοτικού επιπέδου όπως η Β. Αμερική, η Ρωσία και η ήπειρος της Ωκεανίας παρουσιάζουν ποσοστά εμφάνισης (18.0%) αρκετά κοντινά με αυτά των πιο υποβαθμισμένων οικονομικά χωρών (13.5%) όπως εκείνων της Κεντρικής Αφρικής. Επιπλέον, περιοχές υψηλής ανάπτυξης όπως η Β. Λατινική Αμερική, η Μογγολία και η Κίνα, έχουν ποσοστό καρκίνου (32.3%) που προσεγγίζει αυτό των περιοχών μετριας ανάπτυξης (36.2%). Από τα παραπάνω μπορούμε να καταλήξουμε πως το βιοτικό επίπεδο δεν φαίνεται να έχει σημαντική στατιστική ισχύ στην εμφάνιση του καρκίνου, παρότι η διατροφή και οι συνθήκες διαβίωσης έχουν πολλάκις θεωρηθεί στις διάφορες έρευνες ως πρωτεύουσες αιτίες της ασθένειας. Τα στοιχεία αυτά από μόνα τους βέβαια δεν είναι ικανά για να οδηγήσουν σε ασφαλές συμπέρασμα λόγω της διαφορετικής κατανομής του πληθυσμού ανά γεωγραφικό διαμέρισμα. Εν δεύτεροις λοιπόν, η εικόνα

1.3B[4] έρχεται να λειτουργήσει επικουρικά στα παραπάνω, παρέχοντάς μας πληροφορίες για τον καρκίνο με βάση τη γεωγραφία. Από τα ποσοστά και πάλι δεν μπορούμε να εξάγουμε κάποιο συγκεκριμένο αποτέλεσμα καθώς από ό,τι φαίνεται, η εμφάνιση του καρκίνου έστω και αδρομερώς είναι συνηθισμένη με το ποσοστό του πληθυσμού της κάθε γεωγραφικής περιοχής, δηλαδή η κάθε ήπειρος εμφανίζει παρόμοια ποσοστά καρκίνου αναλογικά με τον πληθυσμό της. Για παράδειγμα Η Β.Αμερική με ποσοστό 4.8% επί του συνολικού ποσοστού καρκίνου και πληθυσμό περίπου 550 εκατομμύρια , η Κίνα με τη Μογγολία με ποσοστό 21.7% και πληθυσμό περίπου 1,4 δισεκατομμύρια και η Αφρική με ποσοστό 16.9% και πληθυσμό περίπου 1 δισεκατομμύριο φαίνεται να μην έχουν σημαντικές διακυμάνσεις με βάση το ποσοστό τους και αναλογικά με τον πληθυσμό τους. Ωστόσο οι λιγότερο ανεπτυγμένες χώρες παρουσιάζουν ελαφρώς περισσότερα κρούσματα σε σχέση με τις περισσότερες ανεπτυγμένες. Και πάλι όμως και ενώ δεν υπάρχει μια απόλυτη κατανομή του καρκίνου με βάση το πληθυσμιακό προφίλ του πλανήτη, αυτό που εύκολα μπορούμε να κατανοήσουμε είναι ότι ο καρκίνος αποτελεί ένα νόσημα πλήρως διαδεμένο ανά τη γη, χωρίς ιδιαίτερες εξαιρέσεις.

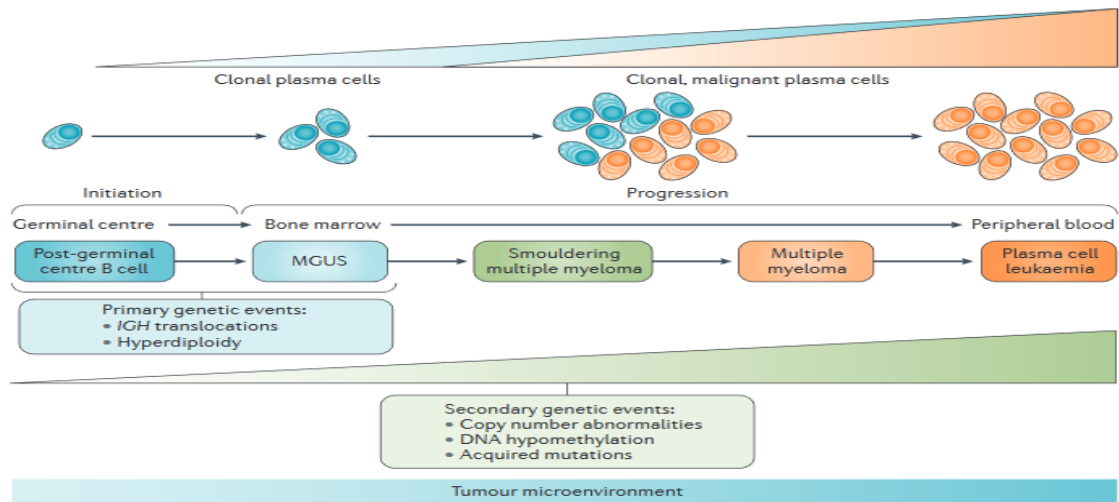
### 1.3 Πολλαπλό μύελωμα – Βασικές αρχές

Το πολλαπλό μύελωμα είναι μια μορφή καρκίνου του αίματος με ολοένα και περισσότερο αισθητή παρουσία τις τελευταίες τρεις δεκαετίες. Πρόκειται για όγκο ο οποίος προκύπτει από πλασματοκύτταρα, κύτταρα δηλαδή τα οποία έχουν προκύψει από τη διαφοροποίηση β-λεμφοκυττάρων και έχουν εξειδικευθεί στην παραγωγή συγκεκριμένων αντισωμάτων. Τα συμπτώματα είναι κυμαινόμενα και αλληλένδετα με την ένταση και το στάδιο της ασθένειας και μπορεί να περιλαμβάνουν αναιμία, οστικό πόνο, εκτεταμένη εφίδρωση και κόπωση, πυρετό ακόμη και αιμορραγία. Όσον αφορά στους αιματολογικούς καρκίνους, μετά από το λέμφωμα τύπου Hodgkin, αποτελεί τη δεύτερη συνηθέστερη αιτία καρκίνου αίματος.[5] Ως ασθένεια το πολλαπλό μύελωμα ανήκει στην κατηγορία των μονοκλωνικών γαμμοπαθειών. Πιο συγκεκριμένα, κατά την εκδήλωση της ασθένειας, λεμφοκύτταρα τύπου Β τα οποία έχουν διαφοροποιηθεί σε πλασματοκύτταρα παραγωγής αντισωμάτων τύπου IgM, (αντισώματα Μ) αρχίζουν για άγνωστο λόγο να πολλαπλασιάζονται υπέρμετρα και δίχως έλεγχο με αποτέλεσμα την αύξηση και συγκέντρωση τους στο μυελό των οστών, το κέντρο της αιμοποίησης κατά την ενήλικη ζωή. Γι'αυτό και τα συμπτώματα τόσο της αιμορραγίας όσο και της ελάττωσης των υπόλοιπων στοιχείων του αίματος είναι συνήθως συνοδευόμενα της ασθένειας.[6] Επιπλέον σε περίπου 15-20% των περιπτώσεων εμφάνισης του πολλαπλού μυελώματος, παράλληλα με την αυξημένη παραγωγή IgM, σημαντική άνοδο σημειώνει και η παραγωγή της πρωτεΐνης κ ελαφριάς αλυσού, απαραίτητη κατά το ανασυνδυασμό της με τις βαριές ανοσοσφαιρικές αλυσίδες για το σχηματισμό των αντισωμάτων. Συνεπώς και τα δύο ευρήματα αυτά σε βιοχημικές εξετάσεις αποτελούν έναν αξιόπιστο δείκτη για τον εντοπισμό ενός ενδεχόμενου καρκίνου.

Ένα πολύ σημαντικό πρόβλημα του πολλαπλού μυελώματος αποτελεί η συνήθως μη συνοδεία συμπτωμάτων με τελική απότοκο την ανίχνευση της ασθένειας σε προχωρημένο στάδιο και τη δυστοκία εξέλιξης της. Επίσης μεγάλη δυσκολία εμφανίζεται και στο σημείο καθορισμού της μετάβασης από την απλή γενική μονόκλωνη γαμμοπάθεια στο πολλαπλό μύελωμα διότι πολλές φορές

οι βιοχημικές τιμές αυτές καθαυτές δεν μπορούν να είναι απόλυτα κατατοπιστικές για να γνωρίζουμε σε ποιο στάδιο έχει φτάσει.[5] Σε αυτό συνεισφέρει σε μεγάλο βαθμό και το γεγονός ότι δε γνωρίζουμε με σαφήνεια τα αίτια που την προκαλούν. Η αντιμετώπιση μέχρι στιγμής πραγματοποιείται με κλασικά αντικαρκινικά φάρμακα όπως οι ανθρακυκλίνες(οι οποίες εμφανίζουν όμως αυξημένα επίπεδα καρδιοτοξικότητας), αλκυλιωτικοί παράγοντες και κορτικοστεροειδή για την καταπράυνση των συμπτωμάτων σε περιόδους έντασης καθώς επίσης και μια σειρά φαρμάκων νέας γενιάς με δραστικότητα αναστολέα πρωτεοσώματος.[5]

Η επόμενη εικόνα παρουσιάζει τα βασικά στάδια παραγωγής του τελικού όγκου:



**Εικόνα 1.4** *Multiple Myeloma. Nature Reviews Disease Primers, 3, 17046.* Ξεκινώντας από μοναδιαία κύτταρα σχηματίζονται οι πρώτες αποικίες πρώιμης μονόκλωνικής γαμμοπάθειας οι οποίες αφού αυξηθούν και καφέρουν να εισαχθούν στην κυκλοφορία του αίματος, είναι έτοιμες να μεταναστεύσουν και να δημιουργήσουν νέες εστίες και τελικά πολλαπλό μυέλωμα. Παράλληλα με την εξέλιξη της νόσου, έχουμε και γενετικές αλλαγές όπως μεταβολή στη μεθυλίωση του DNA και τροποποίηση της συμπύκνωσης της χρωματίνης. [5]

## 1.4 Επιδημιολογικά στοιχεία πολλαπλού μυελώματος

Αν και όπως αναφέρθηκε δεν υπάρχει μια σαφής θέση για το πώς το πολλαπλό μυέλωμα ξεκινά, μια σειρά παραγόντων έχουν χαρακτηριστεί ως υψηλού κινδύνου για την εμφάνιση του. Μια επιπλέον δυσκολία στη διαλεύκανση της ασθένειας αποτελεί το γεγονός πως αν και από τις πιο συχνές ανάμεσα στους αιματολογικούς καρκίνους, γενικότερα ως προς το συνολικό ποσοστό των καρκίνων κατέχει μικρά ποσοστά εμφάνισης με αποτέλεσμα την όχι και τόσο άπλετη παροχή στατιστικών δεδομένων.

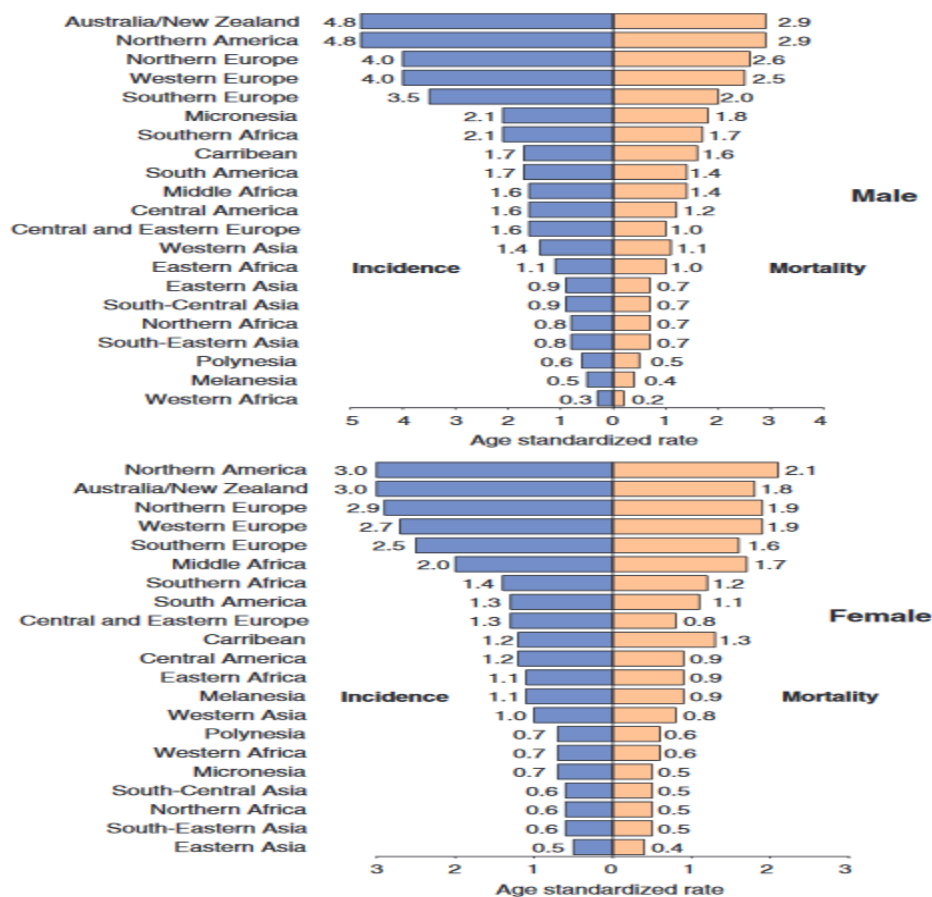
Το κάπνισμα είναι μια ανθυγιεινή συνήθεια η οποία έχει συνδεθεί με αρκετές μορφές καρκίνου, με κύριες των πνευμόνων και της ουροδόχου κύστης. Όσον αφορά στο πολλαπλό μυέλωμα πάντως, δεν υπάρχουν σημαντικές αποδείξεις που να το κατηγοριοποιούν ως παράγοντα εκκίνησης του νοσήματος. Επίσης το αλκοόλ δε φαίνεται μέχρι στιγμής να έχει αποτελέσει στοιχείο επιβαρυντικό για την

εκδήλωση της νόσου ενώ και οι έρευνες στη διατροφή των πάσχοντων δεν έχουν αποδώσει ιδιαίτερους καρπούς παρά μόνο κάποιες αναφορές που υποστηρίζουν πως τα φρούτα και λαχανικά όπως το πορτοκάλι και το λεμόνι έχουν βρεθεί σε ορισμένους ασθενείς να προσδίδουν κάποια μικρή αντίσταση στην εξέλιξη της.[7] Τα ίδια ακριβώς ισχύουν και για την παχυσαρκία και τη φυσική δραστηριότητα, ότι δηλαδή μειωμένα επίπεδα παχυσαρκίας και υψηλά άθλησης ίσως παίζουν κάποιο προστατευτικό ρόλο.

Μια από τις σοβαρότερες αιτίες καρκίνου εν γένει αποτελεί η ιονίζουσα ακτινοβολία. Αν και δεν υπάρχουν και σε αυτήν την περίπτωση πολύ κραταιά δεδομένα που να υποδεικνύουν την άμεση σύνδεση του πολλαπλού μυελώματος με την ακτινοβολία, φαίνεται να υπάρχει μια σχετική αύξηση των κρουσμάτων ύστερα από έκθεση ορισμένων ανθρώπων σε αυτήν. Τα δεδομένα αυτά προέκυψαν κυρίως από εργαζόμενους σε νοσοκομειακές μονάδες ακτινισκόπησης, είτε για διαγνωστικούς είτε για θεραπευτικούς σκοπούς, όπου και εμφάνιζαν μια υψηλότερη συχνότητα εμφάνισης έναντι του υπόλοιπου πληθυσμού. Ένας άλλος δυνητικά επιβαρυντικός παράγοντας που θα έπρεπε να εξεταστεί είναι οι διάφοροι ιοί καθώς αρκετά λεμφώματα, συμπεριλαμβανομένου και του λεμφώματος τύπου Hodgkin, έχουν συνδεθεί με ιούς όπως ο HIV και ο ιός Epstein-Barr. Πράγματι ασθενείς με HIV και νοσούντες ηπατίτιδας C βρέθηκαν να έχουν σημαντικά αυξημένη συχνότητα πολλαπλού μυελώματος σε σχέση με τον υπόλοιπο πληθυσμό. Αυτό δε θα πρέπει να κάνει ιδιαίτερη εντύπωση καθώς οι περισσότεροι εξασθενημένοι οργανισμοί στην προσπάθεια να αντιμετωπίσουν τη δυσχερή τους κατάσταση, υπερχρησιμοποιούν τους αμυντικούς τους μηχανισμούς γεγονός που ενέχει κίνδυνο για την εκδήλωση καρκίνου.[7]

Αδιαμφισβήτητος βασιλιάς μεταξύ των κινδυνόφρων παραγόντων αποτελεί το γενετικό υπόβαθρο. Ασθενείς με πολλαπλό μυέλωμα βρέθηκε πως συχνά είχαν συγγενείς με είτε πολλαπλό μυέλωμα είτε συγγενείς ανωμαλίες του αίματος.[7][8] Όπως και οι περισσότεροι καρκίνοι έτσι και το πολλαπλό μυέλωμα φαίνεται να είναι το αποτέλεσμα άγνωστων γονιδιακών συνδυασμών και συγκυριών που ακόμα δεν έχουν διευλευκανθεί πλήρως.

Ακολουθεί μια στατιστική μελέτη εμφάνισης και θνητότητας από πολλαπλό μυέλωμα καθώς επίσης και η σύγκριση των δεδομένων με γνώμονα το φύλο:



**Εικόνα 1.5** *Epidimiology of Multiple Myeloma, Multiple Myeloma, 25-35* Παρουσιάζονται και για τα δύο φύλα τόσο η πιθανότητα εμφάνισης όσο και τα ποσοστά θνητότητας σε συνάρτηση με την ηλικιακή ομαλοποίηση. Παρατηρούμε τη μεγαλύτερη συχνότητα πολλαπλού μυελώματος στους άνδρες και τα αυξημένα επίπεδα θνητότητας που συνδέονται με την ασθένεια. [7]

Από την εικόνα μπορούμε να εξάγουμε κάποια σχετικά ασφαλή συμπεράσματα. Αφενός η πιθανότητα εμφάνισης είναι μεγαλύτερη στους άνδρες σε σχέση με τις γυναίκες και αφετέρου ότι η ασθένεια εμφανίζει πολύ υψηλά επίπεδα θνητότητας. Ένα ακόμη σημαντικό στοιχείο είναι ότι η ασθένεια κατά βάση εμφανίζεται με μεγαλύτερη συχνότητα στις περισσότερες έναντι των λιγότερο προηγμένων χωρών. Τέλος, παρατηρούμε πως η εμφάνιση αλλά και η θνητότητα είναι πλήρως συσφιασμένη με την αύξηση της ηλικίας.[7]

## 1.5 Παθοφυσιολογία και μοριακή βάση πολλαπλού μυελώματος

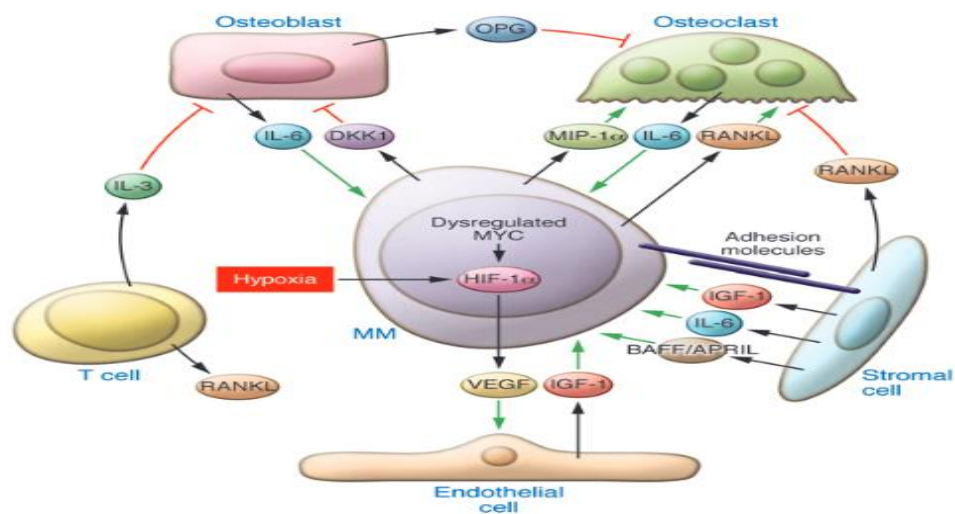
Όπως αναφέρθηκε, το πολλαπλό μυέλωμα αποτελεί ασθένεια που προκαλείται από την ανεξέλεγκτη παραγωγή πλασματοκυττάρων υπεύθυνων για την παραγωγή αντισωμάτων κυρίως τύπου M. Το πρώιμο στάδιο της ασθένειας ονομάζεται μονόκλωνη γαμμοπάθεια ανεξιχνίαστης σημασίας το οποίο διακρίνεται με τη σειρά του σε δύο κατηγορίες, τα λεμφοειδή όπου πρόκειται για διογκωμένους λεμφαδένες χωρίς όμως τον χαρακτηρισμό τους ως ογκογενή και τις μονόκλωνες γαμμοπάθειες ανεξιχνίαστης σημασίας που σχετίζονται με πλασματοκύτταρα(PC MGUS). Το πολλαπλό μυέλωμα αποτελεί συνέχεια της δεύτερης κατηγορίας.[9] Η διαφορά μεταξύ του PC MGUS και του πολλαπλού μυελώματος έγκειται στον εντοπισμό ή όχι βλάβης τελικού σταδίου σε κάποιο όργανο καθώς επίσης και σε ορισμένες βιοχημικές τιμές όπως για παράδειγμα αλλαγές στις τιμές λευκωματίνης του αίματος(αλβουμίνη) με πτώση κάτω από 3g/dl όπου το φυσιολογικό όριο κυμαίνεται μεταξύ 3.5-5g/dl αλλά και αλλαγές στη σύσταση του μυελού των οστών σε πλασματοκύτταρα με τον αριθμό τους να είναι μικρότερος από το 10% της περιεκτικότητάς του σε μονοκύτταρα τα οποία προέρχονται από διαφορετική κυτταρική σειρά. Εκτός αυτών των κριτηρίων, υπάρχει ένα ακρωνύμιο που καλείται CRAB και περιλαμβάνει μια σειρά διαφορετικών δυσλειτουργιών και με βάση το οποίο κρίνεται η ένταση του πολλαπλού μυελώματος. Αναφορικά με αυτό έχουμε C:αύξηση ασβεστίου R:νεφρική ανεπάρκεια A:αναιμία και B:οστικές κακώσεις.Επιπλέον η ασθένεια εμφανίζεται συνήθως σε μεγαλύτερες ηλικίες ενώ το προσδόκιμο ζωής σε πλήρως εξελισσόμενο πολλαπλό μυέλωμα δε ξεπερνάει τα πέντε έτη.[10],[13]

Ενώ ακόμη δεν έχει διευκρινιστεί ακριβώς η αιτία που δημιουργεί το πολλαπλό μυέλωμα, είναι αρκετά πιθανόν να εμπλέκεται κάποιος μηχανισμός ο οποίος καταστέλλει την απόπτωση. Ωστόσο η ανώμαλη και ακατάσχετη πολλαπλασιασση των πλασματοκυττάρων στο μυελό των οστών δεν είναι αρκετή για να προκαλέσει πολλαπλό μυέλωμα. Η πρώτη σοβαρή ένδειξη για την ύπαρξη της ασθένειας είναι η διατάραξη της ισορροπίας του μικροπεριβάλλοντος του μυελού των οστών και κυρίως μεταξύ της αλληλεπίδρασης των οστεοκυττάρων. Τα οστεοκύτταρα είναι μια κατηγορία κυττάρων στην οποία εμπεριέχονται σε ποσοστό 95% δύο είδη, οι οστεοβλάστες οι οποίοι αναλαμβάνουν την αναδόμηση των ήδη υπάρχοντων οστών αλλά και την επέκτασή τους κατά την ανάπτυξη των οργανισμών και οι οστεοκλάστες που έχουν ως αρμοδιότητα την πέψη των οστών είτε λόγω παλαίωσης είτε λόγω ανάγκης εξασφάλισης νέων οδών μηχανικής επέκτασης. Στην περίπτωση του πολλαπλού μυελώματος φαίνεται ότι υπάρχει διαταραχή στο ισοζύγιο μεταξύ των δύο αυτών κυτταρικών τύπων. Πιο συγκεκριμένα, οι περισσότεροι ασθενείς παρουσιάζουν μειωμένα επίπεδα οστεοβλαστών και φυσιολογικά ή ελαφρώς αυξημένα επίπεδα οστεοκλαστών. Έχουν σημειωθεί επίσης περιπτώσεις που τόσο οι οστεοβλάστες όσο και οι οστεοκλάστες μειώνονται αλλά με εμφατικότερη τη μείωση των οστεοβλαστών και τελικά παρόμοια κλινική εικόνα. Αποτέλεσμα αυτής της ανισορροπίας είναι η προοδευτικά φθίνουσα παραγωγή νέων οστών και η εμφάνιση συμπτωμάτων οστεοπόρωσης και ροπής καταγμάτων.[10][11][12] Η εκτεταμένη πέψη των οστών οδηγεί με τη σειρά της στην παραγωγή σωρείας διαμεσολαβητικών παραγόντων όπως ιντερλευκινών, οι οποίοι ακολούθως μεταναστεύουν σε διαφορετικά σημεία του μυελού και αναπαραγάγουν παρόμοια συμπτώματα με την εστία γέννησης τους δηλαδή τον αρχικό όγκο. Έτσι η διαδικασία αυτή συνεχίζεται και αυτό χαρακτηρίζεται πλέον ως πολλαπλό μυέλωμα εν εξελίξει. Επιπλέον, η αλληλεπίδραση των καρκινικών πλασματοκυτταρικών



κλώνων με το περιρρέον περιβάλλον των οστεοκυττάρων, παράγει σήματα τα οποία μεταφέρονται στο στη μεσεγχυματική στιβάδα του οστού από το οποίο αντλούνται τα πρόδρομα οστεοκύτταρα πριν διαφοροποιηθούν στις επιμέρους κατηγορίες. Έρευνες έχουν δείξει πως αυτή η αλληλεπίδραση καταλήγει μέσα από διάφορα βήματα στην έκλυση διαφόρων παραγόντων ιντερλευκινών ή αυξητικών παραγόντων όπως των TGF-1, MIP-1 και IL-7 οι οποίοι δημιουργούν ένα περιβάλλον θετικής ανατροφοδότησης, παρόμοιο με αυτό που προκαλείται από την εκτεταμένη πέψη των οστών, καθώς τα παραπάνω στοιχεία παρέχουν στον αναπτυσσόμενο όγκο ένα αρκετά σταθερό χημικά περιβάλλον και παράλληλα τον προστατεύουν από την εξουδετέρωση του από τους μηχανισμούς άμυνας του οργανισμού.[10]

Το κλειδί για την κατανόηση της ασθένειας αδιαμφισβήτητα αποτελεί ο τρόπος με τον οποίο τα καρκινικά πλασματικά κύτταρα αλληλεπιδρούν με τα οστεοκύτταρα και το συγγενές τους περιβάλλον ούτως ώστε να ξεκινήσει αυτός ο καταρράκτης καταστροφικών εξελίξεων. Ακολουθεί μια εικόνα που παρουσιάζει αδρομερώς ορισμένα από τα μονοπάτια και τα διαμεσολαβητικά μόρια που εμπλέκονται στην ογκογένεση του πολλαπλού μυελώματος:



**Εικόνα 1.6** Interactions of MM tumor cells with the BM microenvironment. Molecular pathogenesis of multiple myeloma and its premalignant precursor. *Journal of Clinical Investigation*, 122(10), 3456–3463 Παρουσίαση της αλληλεπίδρασης του βιολογικού μικροπεριβάλλοντος και συμβολή αυτού στο πολλαπλό μυέλωμα. Η εικόνα αυτή αποδεικνύει την πολυκλότητα της ασθένειας καθώς περιλαμβάνει τόσο τα σηματοδοτικά μόρια που παράγονται από τον ίδιο τον όγκο, όσο και τον τρόπο με τον οποίο αυτές οι μεταβολές οδηγούν σε εκτροπή του γηγενούς εξωκυττάρου περιβάλλοντος με συνέπεια την εκδήλωση καρκίνου.[9]

Γίνεται λοιπόν αντιληπτό το πόσο πολύπλοκα το μικροπεριβάλλον του μυελού των οστών μπορεί να συμβάλει στην εμφάνιση της ασθένειας. Ορισμένα από αυτά τα στοιχεία έχουν μελετηθεί περαιτέρω και παρέχουν εξηγήσεις για την εμπλοκή τους στο νόσημα. Από την εικόνα 1.6 παρατηρούμε

πως ένα από τα γονίδια που σχετίζονται με το πολλαπλό μύελωμα, το MYC, φαίνεται να παίζει σημαντικό ρόλο στην απορρύθμιση του πολλαπλασιασμού των πλασματοκυττάρων. Αυτή η απορρύθμιση επιφέρει υποξία στα κύτταρα λόγω του αυξημένου μεταβολισμού αυτών με αποτέλεσμα να ενεργοποιείται ο παράγοντας υποξίας τύπου HIF-1α, ο οποίος όταν εισαχθεί στο κύτταρο διαμεσολαβεί στην παραγωγή του αγγειοδραστικού παράγοντα VEGF. Με τη σειρά του εκείνος αφού δράσει σε παρακρινή ενδοθηλιακά κύτταρα, ενεργοποιεί έναν αυξητικό παράγοντα, τον IGF-1 (insulin growth factor), ο οποίος έχει συσχετιστεί με πολλά είδη καρκίνου στοχεύοντας απευθείας στην αύξηση της καρκινικής μάζας. Το παραπάνω αποτελεί ένα πολύ καλό παράδειγμα για το πως δουλεύει ένα από τα συστήματα ανατροφοδότης που συντηρούν και διασπείρουν το καρκίνωμα. Άλλοι καίριοι διαμεσολαβητές οι οποίοι αποτελούν απότοκο του τροποποιημένου χημικά περιβάλλοντος του στρώματος του μυελού αποτελούν οι IL-6, BAFF/APRIL και IGF-1 όπου συμμετέχουν στην ανοσοαντοχή του καρκίνου, στην αγγειογένεση, τη μεταβολή στο προφίλ οστεογένεσης και στη γενικότερη επιβίωση των καρκινικών κυττάρων.[9]

Πέραν όμως των παραπάνω στοιχείων, τα οποία προκύπτουν ως απόρροια της δράσης των ανώμαλα παραγόμενων πλασματοκυττάρων, είναι απαραίτητη και η αλληλεπίδραση τους με τα οστεοκύτταρα. Ένα από τα σημαντικότερα μονοπάτια αυτής της κατηγορίας περιλαμβάνει τα μόρια RANK, RANKL και OPG. Το συγκεκριμένο μονοπάτι βρίσκεται στην καρδιά του συστήματος ανακατασκευής και πέψης οστών.[10] Η RANKL είναι μια ομοτριμερής τύπου 2 πρωτεΐνη η οποία εντοπίζεται στην επιφάνεια της κυτταρικής μεμβράνης. Εκφράζεται περισσότερο σε όργανα όπως οι πνεύμονες, οι λεμφαδένες και ο θύμος αδένας ενώ σε μικρότερη ποσότητα στο σπλήνα και στο μυελό των οστών και εμφανίζεται τόσο σε οστεοβλάστες όσο και σε άλλα κύτταρα όπως τα επιθηλιακά. Από μόνη της δεν παρουσιάζει κάποια δραστηριότητα ενώ σαν κατηγορία ανήκει στην υπεροικογένεια των TNF (tumor necrosis factors). Ωστόσο η συγκεκριμένη πρωτεΐνη είναι σε θέση να αλληλεπιδρά με ένα πλήθος διαφορετικών στοιχείων τόσο σημάτων όσο και διαφορετικών κυττάρων.[12][14] Από την άλλη πλευρά, υπάρχει μια άλλη ομοτριμερής πρωτεΐνη τύπου 1 που ονομάζεται RANK. Η συγκεκριμένη εντοπίζεται σε οστεοκλάστες αλλά και σε δενδριτικά κύτταρα. Όπως και η RANKL, έτσι και η RANK δεν παρουσιάζει αυτόνομη δράση. Να σημειωθεί πως η RANK έχει εντοπιστεί να αυξάνεται σε περιπτώσεις καρκίνου, παραγόμενη κυρίως από T-λεμφοκύτταρα τραυματισμένου ιστού.[14] Τέλος, υπάρχει μια τρίτη πρωτεΐνη ονόματι οστεοπροτεγερίνη (OPG) η οποία ρυθμίζεται γονοδιακά από κοινά μονοπάτια ρύθμισης με την RANKL. Μεταξύ των τριών αυτών πρωτεϊνών υπάρχει μια πολύ σημαντική σύνδεση. Προκειμένου οι οστεοκλάστες να διαφοροποιηθούν από την πρόωμη μορφή τους στην ώριμη, είναι απαραίτητη η αλληλεπίδραση της πρωτεΐνης RANKL της επιφάνειας των οστεοβλαστών με την RANK των πρόωμων οστεοκλαστών. Όταν τα δύο μόρια αυτά συνδεθούν, οι οστεοκλάστες ωριμάζουν και είναι σε θέση να επιτελέσουν την οστεοκλαστική τους δραστηριότητα. Από την άλλη, όταν η πρωτεΐνη OPG ενεργοποιείται, δύναται να συνδεθεί με την πρωτεΐνη RANKL και από αυτήν την αλληλοσυμπλήρωση προκύπτει ένα πρωτεϊνικό σύμπλοκο το οποίο εμποδίζει την RANK να προσδεθεί με την RANKL και τελικά σταματά την οστεοκλαστογένεση.[11][14] Όπως παρουσιάζεται και στην εικόνα 1.6, στην περίπτωση του πολλαπλού μυελώματος έχουμε παραγωγή IL-3, μιας από τις ιντερλευκίνες που μεσολαβούν στην απενεργοποίηση της παραγωγής οστεοβλαστών ενώ παράλληλα η σύνδεση τους με τους οστεοκλάστες αυξάνει τη δράση τους. Όπως αναφέρθηκε, η μείωση των οστεοβλαστών οδηγεί σε μείωση της διαθέσιμης ποσότητας πρωτεΐνης OPG αφού ελέγχονται από κοινό γονιδιακό μονοπάτι και

ως εκ τούτου έχουμε μείωση των μοριακών φρένων παραγωγής οστεοκλαστών. Παράλληλα με τα άνωθι, τα καρκινικά κύτταρα αυξάνουν την ποσότητα RANKL και έτσι πλέον οι οστεοκλάστες λαμβάνουν RANKL όχι από τους οστεοβλάστες που φυσιολογικά θα έπρεπε αλλά από τους καρκινικούς κλώνους. Ως τελικό αποτέλεσμα λοιπόν έχουμε μείωση δράσης οστεοβλαστών και αύξηση δράσης οστεοκλαστών. Το συγκεκριμένο μεταβολικό πλέγμα είναι το ισχυρότερο μέχρι στιγμής εύρημα για τον τρόπο με τον οποίο τα κύτταρα στη φάση του MGUS είναι σε θέση να επηρεάσουν το περιβάλλον περιβάλλον της οστεογένεσης και να δημιουργήσουν τις πρώτες εστίες μυελώματος.

Μια ακόμη πρωτεΐνη με αρκετά σημαντική συμβολή στο πολλαπλό μυέλωμα είναι και η DKK1. Κύριος ρόλος της είναι η δράση της ως ανταγωνιστής του σηματοδοτικού μορίου wnt το οποίο διαδραματίζει σημαντική λειτουργία στην ανάπτυξη αρκετών ειδών κυττάρων και είναι άρρηκτα συνδεδεμένο με την επιβίωση. Η πρωτεΐνη αυτή όταν συνδεθεί με μια άλλη που ονομάζεται LRP5/6 αλλά και με μια ακόμη που καλείται Kremen 1/2, σχηματίζεται ένα σύμπλοκο μεταξύ των LRP και Kremen με αποτέλεσμα την είσοδό τους στο εσωτερικό του κυττάρου. Στη συνέχεια, το δραστικό τμήμα στην περίπτωση του πολλαπλού μυελώματος αποτελεί μόνο η LRP, δηλαδή η συμβολή της Kremen στην ασθένεια είναι μόνο η διευκόλυνση της εισόδου της LRP. Όταν εισαχθεί λοιπόν, η LRP παρεμποδίζει ένα από τα επιμέρους τμήματα του μονοπατιού που σχετίζεται με τον αυξητικό παράγοντα wnt[11]. Πιο συγκεκριμένα, όταν ο wnt συνδεθεί σε υποδοχέα της μεμβράνης ορισμένων κυττάρων, πυροδοτεί μια αλυσίδα γεγονότων με πρώτο στάδιο τη δέσμευση της ελεύθερης β κατενίνης του κυττάρου για να συνδυαστεί με μια άλλη πρωτεΐνη την TCF. Αυτό το σύμπλοκο στη συνέχεια ενεργοποιεί τους τελικούς στόχους του μονοπατιού του wnt που σχετίζονται με την ανάπτυξη και την επιβίωση του κυττάρου. Το συγκεκριμένο σηματοδοτικό μόριο σχετίζεται και με την εμφάνιση ενός άλλου καρκίνου, του APC.(οικογενής αδενωματώδης πολυποδίαση εντέρου) Όσον αφορά στο πολλαπλό μυέλωμα, η δράση της DKK1 οδηγεί τελικά σε αναστολή ενός μονοπατιού που ρυθμίζεται από το wnt και ελέγχει τη διαφοροποίηση των πρώιμων οστεοβλαστών σε ώριμους και έτσι συμβάλλει στην περαιτέρω μείωση του αριθμού τους και στην ενίσχυση του καρκινικού αποτελέσματος. Μια επιπλέον πρωτεΐνη που δρα με παρόμοιο τρόπο είναι και η σκληροστίνη. Και η συγκεκριμένη συνδέεται με την πρωτεΐνη LRP5/6 και το σύμπλοκο αυτό εμποδίζει τη φυσιολογική πρόσδεση της τελευταίας με τους φυσιολογικούς της συνδέτες. Πέραν τούτου, η σκληροστίνη μαζί με τη νογγίνη, παρεμποδίζουν τη σύνδεση των οστεομορφογενετικών παραγόντων τύπου 1 και 2 στους αντίστοιχους υποδοχείς τους με συνέπεια την εξάλειψη τροφοδοσίας των οστεοβλαστών με απαραίτητα στοιχεία για την ανάπτυξη τους.

## 1.6 Γενετική του πολλαπλού μυελώματος

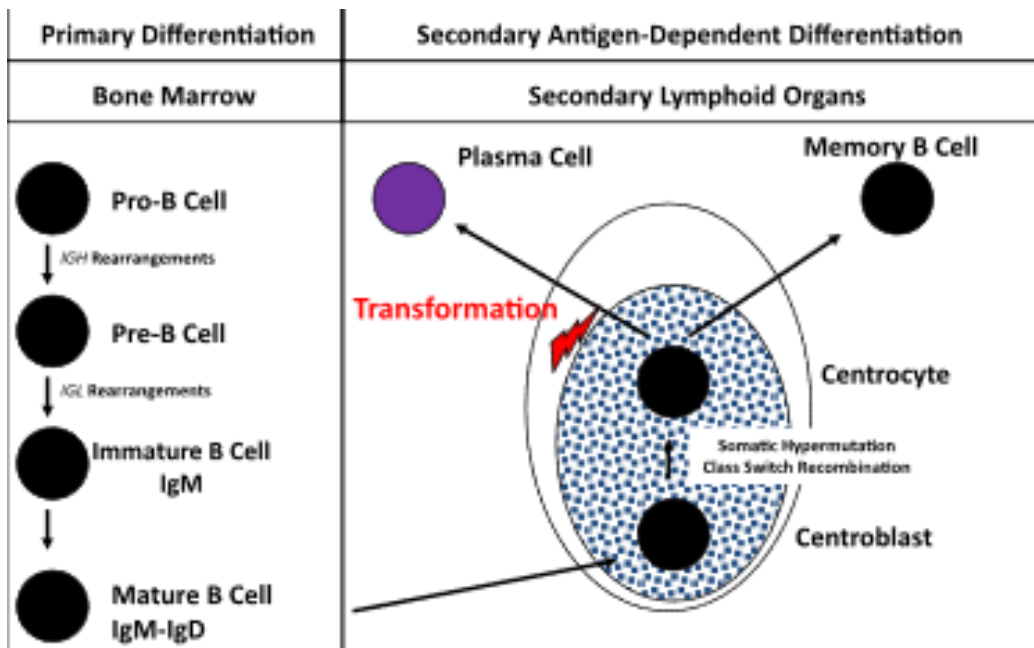
Η διερεύνηση της παθοφυσιολογίας της νόσου από μόνη της δεν είναι ικανή να εξηγήσει τις αιτίες που αποτελούν την εναρκτήρια σπίθα εκδήλωσης της. Τέτοιας φύσεως απαντήσεις μας δίνει η γενετική μελέτη του πολλαπλού μυελώματος, η οποία μάλιστα εμφανίζει μια μεγάλη ποικιλία. Ειδικότερα, οι βασικές κατηγορίες γενετικών διαταραχών που σχετίζονται με το νόσημα εκτείνονται σε μεγάλη σφαίρα και κυρίως περιλαμβάνουν τις ανωμαλίες κατά τη γενετική αναδιάταξη των γονιδίων των

ανοσοσφαιρινών, τη χρωμοσωματική ανευπλοειδία, τις χρωμοσωματικές μεταθέσεις, την απορρύθμιση ορισμένων σηματοδοτικών μονοπατιών με αποτέλεσμα την πρόωμη και έκτοπη έκκριση διαμεσολαβητικών μορίων σχετιζόμενων με καρκινογένεση και τις επιγενετικές τροποποιήσεις. Το πολλαπλό μύελωμα αν και όχι τόσο συχνό σε εμφάνιση, παρουσιάζει περίπλοκη βιολογία και το γενετικό του υπόβαθρο αποτελεί το στάδιο με τα περισσότερα σκοτεινά σημεία μέχρι στιγμής. Μεγαλύτερη ακόμη σημασία της περαιτέρω διερεύνησης των γενετικών αιτιών της ασθένειας προσδίδει και το γεγονός πως παρουσιάζει σε πολύ μεγάλο ποσοστό οιγογενειακή επανεμφάνιση που συνεπάγεται ισχυρή βάση κληρονομισιμότητας. Κυρίως πλήττονται άτομα μεγαλύτερης ηλικίας, όπως και στους περισσότερους καρκίνους άλλωστε, καθώς απαιτείται να παρέλθει μεγάλο χρονικό διάστημα για να λάβουν χώρα μια σειρά μεγάλου αριθμού μεταλλάξεων.[15][17]

Οι ανοσοσφαιρίνες υπόκεινται σε ένα δαιδαλώδες σύστημα ανασυρμαλόγησης και ανακατατάξης το οποίο αρχίζει να διαδραματίζεται πολύ πριν ωριμάσουν στην πλήρως λειτουργική μορφή τους. Τα συγκεκριμένα μόρια εμφανίζουν μια εξαιρετική ιδιομορφία κατά τη σύνθεσή τους, δε διαθέτουν ένα ενιαίο γονίδιο αλλά περισσότερα από ένα με τμήματα αυτών να ανακατεύονται συνεχώς ωστόσο σχηματιστούν τα σωστά. Οι ανοσοσφαιρίνες των αντισωμάτων αποτελούνται από τέσσερες αλυσίδες ανά δύο όμοιες με τις δύο από αυτές να ονομάζονται ελαφριές και τις δύο άλλες βαριές. Οι αλυσίδες μεταξύ τους συνδέονται με δισουλφιδικούς δεσμούς σε ένα σχήμα που θυμίζει το κεφαλαίο ύψιλον. Η κάθε ελαφριά αλυσίδα έχει δύο τμήματα και η κάθε βαριά τέσσερα, συνολικά δηλαδή το αντίσωμα διαθέτει τέσσερα ελαφριά και οκτώ βαριά τμήματα. Από αυτά, τα δύο ελαφριά και τα έξι βαριά είναι πανομοιότυπα σε όλα τα είδη ενώ μόνο τα υπόλοιπα δύο ελαφριά και δύο βαριά είναι αυτά που συνιστούν το μεταβλητό κομμάτι. Ωστόσο το συγκεκριμένο σημείο είναι και το κύριο δραστικό συστατικό του μορίου καθώς του επιτρέπει να προσδένει επιλεκτικά και με εξαιρετική ακρίβεια ειδικά αντιγόνα και να τα οδηγεί προς καταστροφή, προφυλάσσοντας τον οργανισμό από δυνητικές βλάβες. Αξίζει να αναφερθεί ότι και η σταθερή περιοχή διαδραματίζει σημαντικό ρόλο διευκολύνοντας τη χημειοταξία του αντισώματος όπως επίσης αναγνώρισης του από μακροφάγα για την μετέπειτα εξουδετέρωση του αντιγόνου με το οποίο είναι συνδεδεμένο.[16]

Το γεγονός πως το πολλαπλό μύελωμα αφορά στην απρόσφορη και ανεξέλεγκτη παραγωγή πλασματοκυττάρων φαίνεται να έχει άμεση συσχέτιση με διαδικασίες που διαδραματίζονται τόσο πριν την ωρίμανση των β-λεμφοκυττάρων όσο και μετά από αυτήν. Όπως είναι γνωστό, όλα τα αντισώματα προκύπτουν από τη δράση β-λεμφοκυττάρων ύστερα από διαφοροποίηση τους σε ένα συγκεκριμένο τύπο. Τα πλασματοκύτταρα αποτελούν το τελευταίο βαθμό ωρίμανσης των β-λεμφοκυττάρων. Το πρώτο στάδιο ξεκινά με την ανασύσταση του γονιδίου των βαριών αλυσίδων.[17] Πιο συγκεκριμένα, το γονίδιο αυτό έχει ένα τεράστιο μέγεθος (2 mbs) και μέσα στο οποίο βρίσκονται διασκορπισμένα διάφορα μικρότερα που αφορούν στη σύνθεση των βαριών αλυσίδων. Υπάρχουν τρία τέτοια υπογονιδιακά σύνολα τα οποία ονομάζονται V, D, J αντίστοιχα αλλά και ένα τέταρτο που είναι πανομοιότυπο σε κάθε αλυσίδα και δεν υφίσταται τροποποιήσεις. Τα τρία αυτά τμήματα γονιδίου επιδέχονται τεράστιες ανασυστάσεις με μερικά κοψίματα, αφαιρέσεις τμημάτων και επανακολλήσεις ούτως ώστε να σχηματιστεί το ώριμο DNA των βαριών αλυσίδων και στη συνέχεια να μεταγραφεί από κοινό υποκινητή. Το ένζυμο που αναλαμβάνει να διατελέσει την επανασύνδεση καλείται RAG. Μέχρι αυτήν τη στιγμή, το β λεμφοκύτταρο βρίσκεται στην πρώτη φάση. Σε επόμενο στάδιο και εάν το κύτ-

ταρο κριθεί ότι παρήγαγε γενετικά χρήσιμο προϊόν στο προηγούμενο, περνά στη φάση παραγωγής των ελαφριών αλυσίδων. Εδώ υπάρχουν δύο διαφορετικά γονίδια, τα IGLk και IGLλ, τα οποία επίσης υπόκεινται σε ανασύσταση όχι όμως μεταξύ τους αλλά το καθένα με τον εαυτό του. Εάν και αυτό το στάδιο (φάση δύο) θεωρηθεί επιτυχές, τα λεμφοκύτταρα Β είναι ελεύθερα να μεταναστεύσουν στο θύμο αδένι για να διαφοροποιηθούν σε πλασματοκύτταρα και πλέον να εξειδικεύσουν τη δράση τους παράγοντας τα γνωστά IGM, IGG, IGG και IGE (φάση 3). [16][17][20] Αρκετά στοιχεία μέχρι στιγμής φαίνεται να υποδεικνύουν πως οι μεταλλάξεις που απαιτούνται για την πρόκληση του αρχικού σταδίου του πολλαπλού μυελώματος πηγάζουν από τις βαριές μορφές ανακατατάξεις που λαμβάνουν χώρα κατά τη διαδικασία ωρίμανσης των β-λεμφοκυττάρων. Αυτές οι συχνές μεταφορές γονιδιακών τμημάτων, οι ανασυνδυασμοί και η εξάλειψη του διασυνδεδεμένου DNA αφήνουν αρκετές πιθανότητες ανοικτές για εκτεταμένη συσσώρευση μεταλλάξεων ενώ σε ασθενείς με διαγνωσμένο πολλαπλό μυέλωμα έχουν ανιχνευθεί διαφορές στον ανασυνδυασμό των τμημάτων VDJ. Παρουσιάζεται κατόπιν μια εικόνα με ενδεικτική απεικόνιση των σταδίων ωρίμανσης των β-λεμφοκυττάρων.



**Εικόνα 1.7** B cell differentiation. *Genetics of multiple myeloma: another heterogeneity level? Blood, 125(12), 1870–1876.* Διακρίνουμε τα επιμέρους στάδια διαφοροποίησης των β-λεμφοκυττάρων καθώς επίσης και την τελική ωρίμανση στο θύμο αδένι που παράγει δύο διακριτά είδη κυττάρων: α) β-κύτταρα μνήμης β) πλασματοκύτταρα.

Υπάρχει ένας τεράστιος αριθμός μεταθέσεων γονιδιακών ή χρωμοσωματικών τμημάτων κατά την ανασύσταση του γενετικού υλικού των β-λεμφοκυττάρων που έχουν ενοχοποιηθεί για τη συσχέτιση τους με την πάθηση. Όσον αφορά στην πρώτη κατηγορία, σημαντικές είναι οι μεταθέσεις γονιδιακών τμημάτων του γονιδίου των βαριών αλυσίδων που βρίσκεται στο χρωμόσωμα 14 και περιοχή βραχίονα q32 καθώς επίσης και οι μεταθέσεις των δύο γονιδίων κ και λ των ελαφριών αλυσίδων στις περιοχές

p22 του χρωμοσώματος 2 και 11q του χρωμοσώματος 22 αντίστοιχα. Αυτά τα τρία σημεία εμφανίζουν αυξημένη ροπή για λανθασμένη μετάθεση και συνένωση κομματιών DNA. Και τα τρία είδη μεταλλάξεων έχουν ως αποτέλεσμα την ενεργοποίηση ενός ογκογονιδίου το οποίο φέρεται να είναι φυσιολογικά ανενεργό και να βρίσκεται στο χρωμόσωμα 14, σε κοντινή απόσταση από τους ενισχυτές του γονιδίου που κωδικοποιεί τις βαριές αλυσίδες. Αυτή η ενεργοποίηση φαίνεται να ξεκινά από μια αφαίρεση ενός κομματιού ιντρονίου που βρίσκεται κοντά στο συγκεκριμένο ενισχυτή και οδηγεί σε αύξηση της δραστηριότητάς του. Εκτός από ασθενείς με πολλαπλό μυέλωμα, τα συγκεκριμένα στοιχεία έχουν ταυτοποιηθεί και σε μη ασθενείς με MGUS, ένα ακόμα δεδομένο που ενισχύει την υπόθεση σύνδεσης με το νόσημα.[19] Μια άλλη περίπτωση μετάθεσης είναι εκείνη που αφορά στη μεταφορά ενός τμήματος από το χρωμόσωμα 4 στο χρωμόσωμα 14, της t(4; 14). Και σε αυτήν την περίπτωση έχουμε αλλαγή στη δράση δύο, των FGFR3 και MMSET, τοποθετώντας τα κοντά σε έναν ενισχυτή του γονιδίου των ελαφριών αλυσίδων. Το πρώτο που διαθέτει φυσιολογικά λειτουργία υποδοχέα κινάσης τυροσίνης, τις περισσότερες φορές (70%) απενεργοποιείται πιθανώς λόγω του ότι η μετάθεση το απομάκρυνε από τα τοπικά στοιχεία που είναι απαραίτητα για την σωστή λειτουργία του, δίχως όμως να γνωρίζουμε ποια ακριβώς είναι αυτή και κατ'επέκταση χωρίς να είμαστε σε θέση να καθορίσουμε τη σημασία του γονιδίου στην ογκογένεση. Από την άλλη, το γονίδιο MMSET υπερδραστηριοποιείται και είναι ενεργό σε όλες τις περιπτώσεις αυτού του είδους της μετάθεσης. Ο φυσιολογικός του ρόλος είναι η κωδικοποίηση ενός παράγοντα με δραστηριότητα μεθυλοτρανσφεράσης ιστονών, κατατάσσοντάς τον στα ένζυμα που σχετίζονται με την αναδιαμόρφωση της χρωματίνης, μιας πλήρως δυναμικής διαδικασίας. Αν και από τη λειτουργία και μόνο φαίνεται να υπάρχει κάποια σύνδεση με το πολλαπλό μυέλωμα, ακόμη δεν έχει πλήρως αποσαφηνιστεί ο ρόλος του.[19] Δύο άλλες λιγότερο συχνές μεταθέσεις είναι οι t(6; 14) και t(11; 14). Η πρώτη σχετίζεται και πάλι με μετάθεση κοντά σε έναν από τους ενισχυτές των γονιδίων της ελαφριάς αλυσίδας. Η δεύτερη περίπτωση όμως στο 17% των συνολικών του πολλαπλού μυελώματος, οδήγησε σε αυξημένη ενεργοποίηση δύο γονιδίων CCND1 και CCND3 που σχετίζονται με την κυκλίνη D, της οποίας η αναστολή έχει αναγνωριστεί ως σχετική με το πολλαπλό μυέλωμα σε πολλές έρευνες. Τέλος δύο ακόμη μεταθέσεις οι t(14; 16), t(14; 20) έχουν αξιολογηθεί ως ύποπτες. Αμφότερες έχουν συνδεθεί με την υπερέκφραση ενός ογκογονιδίου τύπου MAF. Η πρώτη φαίνεται να οδηγεί σε εναλλακτικό του φυσιολογικού μάτισμα το γονίδιο CCND2 που σχετίζεται με την κυκλίνη D, συνδεόμενο απευθείας με τον υποκινητή του γονιδίου. Η δεύτερη μετάθεση οδηγεί σε υπερπαραγωγή του MAFB το οποίο με τη σειρά του αναστέλλει τη δράση του CCND2.[18] Να σημειωθεί ότι όλα τα γονίδια της οικογένειας του συμπλόκου της κυκλίνης D είναι απαραίτητα για τη φωσφορυλίωση διάφορων συστατικών του κυττάρου, με στόχο την απενεργοποίηση ή ενεργοποίηση τους όπως για παράδειγμα στην περίπτωση του ρετινοβλαστώματος όπου φωσφορυλιώνει διάφορους στόχους ελέγχοντας τη μετάβαση από τη φάση G1 στη φάση S.

Οι ανευπλοειδίες αποτελούν έναν επιπλέον σημαντικό παράγοντα κινδύνου για το νόσημα. Η παρέκκλιση από το φυσιολογικό αριθμό χρωμοσωμάτων είναι μια πολύ συχνά εμφανιζόμενη κατάσταση στον καρκίνο. Συνήθως προκύπτουν σε κάποιο από τα επιμέρους στάδια της μίτωσης, το διαχωρισμό δηλαδή των αυτοσωμικών χρωμοσωμάτων κατά την κυτταρική διαίρεση, οδηγώντας σε κυτταρικούς απογόνους με περίσσεια ή έλλειμμα χρωμοσωμάτων. Στην περίπτωση του πολλαπλού μυελώματος, σε περισσότερους από τους μισούς ασθενείς παρουσιάζεται ανευπλοειδία είτε με τη μορφή αύξησης είτε μείωσης των χρωμοσωμάτων. Αξίζει να σημειωθεί πως τα περισσότερα από τα

λάθη που συμβαίνουν σε αυτές τις περιπτώσεις, λαμβάνουν χώρα σε ένα και μόνο από τα στάδια της μίτωσης και δεν απαιτούν τη συσσώρευση μεταλλάξεων σε προηγηθέντα χρόνο. Ένα από τα κύρια χρωμοσώματα που έχουν ταυτοποιηθεί πολλαπλώς ως τροποποιημένα είναι το χρωμόσωμα 1. Πιο συγκεκριμένα, έχουν εντοπιστεί σε αρκετές περιπτώσεις ασθενών με πολλαπλό μυέλωμα, δύο διαφορετικές χρωμοσωματικές ανωμαλίες. Στην πρώτη περίπτωση έχουμε ένα επιπλέον του φυσιολογικού τμήματος από τον μακρύ βραχίονα του χρωμοσώματος 1 και στη δεύτερη έλλειμα του κοντού βραχίονα. Το πρώτο φαινόμενο έχει παρατηρηθεί σε ποσοστό 35-40% των ασθενών με πολλαπλό μυέλωμα ενώ το δεύτερο σε ποσοστό 30%. Υπάρχουν αρκετά επίσης περιστατικά τα οποία εμφανίζουν και τις δύο παρεκκλίσεις ταυτοχρόνως.[18] Όσον αφορά στον πλεονάζοντα βραχίονα, συνήθως συνοδεύεται από μεταθέσεις τύπου t(4;14) ή t(14;16) οι οποίες μπορεί να παρουσιάζονται και μαζί. Το τμήμα αυτό που ονομάζεται 1q21, φέρει αρκετά γονίδια των οποίων ο ρόλος δεν έχει διαλευκανθεί. Ωστόσο, περιέχει μια σειρά γονιδίων μεταξύ αυτών και τα CKS1B, ANP32E, BCL-9 και PDZK1 τα οποία έχουν ταυτοποιηθεί ως ογκογονίδια σε άλλες γνωστές περιπτώσεις καρκίνου. Το ANP32E συγκεκριμένα φέρει δραστηριότητα φωσφατάσης, επηρεάζοντας τη διαμόρφωση της χρωματίνης και συνεπώς ρυθμίζει τη μεταγραφική λειτουργία σειράς γονιδίων. Αν και δεν γνωρίζουμε τα ακριβή μονοπάτια στα οποία συμμετέχει, έχει βρεθεί πως ο χρόνος ζωής των κυττάρων καθορίζεται σε σημαντικό βαθμό από την ενέργειά της.[18] Η απώλεια τμήματος του μικρού βραχίονα του χρωμοσώματος 1 φαίνεται πως επίσης συνδέεται άμεσα με το μυέλωμα. Υπάρχουν δύο περιοχές ενδιαφέροντος πάνω στο βραχίονα, οι 1p12 και 1p32.3. Η πρώτη περιέχει το γονίδιο FAM46C όπου πρόκειται για ογκοσταλτικό παράγοντα κατά μεγάλη πιθανότητα μιας και η δράσης του δεν είναι πλήρως γνωστή. Ωστόσο σε αρκετές περιπτώσεις πολλαπλού μυελώματος στις οποίες ολόκληρο το χρωμόσωμα 1 είναι ακέραιο, το συγκεκριμένο γονίδιο είναι μεταλλαγμένο. Επομένως υπάρχει μεγάλη πιθανότητα είτε να απουσιάζει είτε να είναι αλλοιωμένο, όπου όμως σε κάθε περίπτωση έχει συσχετιστεί με πολλαπλό μυέλωμα. Έχει επίσης ταυτοποιηθεί η δράση του στην πρωτεϊνική παραγωγή από τα ριβοσώματα και συγκεκριμένα στο στάδιο της επιμήκυνσης της αμινοξικής αλυσίδας. Το κομμάτι 1p32.3 περιλαμβάνει τα γονίδια FAF1 και CDKN2C. Το πρώτο φαίνεται να σχετίζεται με τον προγραμματισμένο θάνατο του κυττάρου ενώ το δεύτερο ανήκει στη μεγάλη οικογένεια τους συμπλόκου της κυκλίνης D όπου επανειλημμένως μέλη της έχουν βρεθεί να εμπλέκονται σε φωσφορυλιώσεις οι οποίες ελέγχουν τη μετάβαση από το ένα στάδιο του κυτταρικού κύκλου στο επόμενο. Τα δύο γονίδια μπορεί να έχουν χάσει είτε το ένα είτε και τα δύο αλληλόμορφα τους ενώ συνήθως και τα δύο γονίδια απουσιάζουν σε ασθενείς με πολλαπλό μυέλωμα.[18][19] Δύο ακόμα σημαντικά χρωμοσωματικά ελλείματα αποτελούν εκείνα των χρωμοσωμάτων 17 και 13. Η απώλεια του χρωμοσωματικού μικρού βραχίονα 17 ή ολόκληρου του χρωμοσώματος είναι λιγότερο συχνή. Και στις δύο περιπτώσεις, τις περισσότερες φορές συνυπάρχει μετάθεση t(4; 14) ενώ επιπλέον έχει βρεθεί μείωση της έκφρασης του γονιδίου TP53, ενός εκ των σημαντικότερων ογκοκατασταλτικών γονιδίων.[19] Ειδικότερα, το TP53 κωδικοποιεί φυσιολογικά μία ανασταλτική πρωτεΐνη η οποία συμβάλλει σε πολλές διαφορετικές λειτουργίες όπως στην ενεργοποίηση της επιδιόρθωσης του DNA όταν αυτό υπόκειται σε βλάβη, στη σωστή και μη ατελέσφορη δράση των τελομερών καθώς επίσης και στον έλεγχο του κυτταρικού περιβάλλοντος κατά τη μετάβαση από τη φάση G1 στη φάση S του κυτταρικού κύκλου ενώ θεωρείται ορόσημο στην έρευνα του καρκίνου. Πολύ πιθανόν λοιπόν η πλήρης απουσία ή η μερική παρουσία(ετεροζυγωτία) να συνδέεται με το πολλαπλό μυέλωμα. Το χρωμόσωμα

13 επίσης έχει ταυτοποιηθεί να παρουσιάζει έλλειμμα σε περίπου από τις μισές των περιπτώσεων του πολλαπλού μυελώματος. Μπορεί να απουσιάζει είτε πλήρως είτε μερικώς (13q). Ένα από τα γονίδια που περιλαμβάνει η περιοχή 13q είναι εκείνο του RB1, το οποίο υπολειτουργεί. Το RB1 έχει ως φυσιολογικό ρόλο την παραγωγή ανασταλτικής πρωτεΐνης η οποία εμποδίζει τη μεταγραφή ορισμένων ογκογονιδίων, κρατώντας την κυτταρική διαίρεση σε ισορροπία. Ορισμένα γεγονότα όμως που οδηγούν στην απώλεια του, επιτρέπουν στα κυτταρικά μιτογόνα να δρουν ανεξέλεγκτα, οδηγώντας σε καρκίνο. Η απώλεια του βραχίονα 13q αλλά και ολόκληρου του χρωμοσώματος 13 φαίνεται να εμφανίζεται συχνά σε συνδυασμό με χρωμοσωματικές μεταθέσεις αλλά και την απώλεια του κομματιού 17p13.[18][19]

Όπως αναφέρθηκε, αρκετά σηματοδοτικά μονοπάτια παραποιούνται κατά την εκδήλωση του πολλαπλού μυελώματος. Ανάμεσα τους το MAPK μονοπάτι, το μονοπάτι των JAK-STAT υποδοχέων, της κινάσης φωσφατιδυλοϊνοσιτόλης (PI3K) αλλά και αλλαγές στα πρότυπα επιδιόρθωσης του DNA και διαχείρισης του RNA. Το μονοπάτι της MAPK πρόκειται για αλληλουχία αλυσιδωτών φωσφορυλιώσεων ως απάντηση από ερεθίσματα του εξωτερικού περιβάλλοντος του κυττάρου με τελικούς αποδέκτες μια σειρά φωσφορικών κινασών οι οποίες ενεργοποιούνται και παρουσιάζουν ποικίλες δράσεις. Δύο γονίδια, τα NRAS και KRAS, κατέχουν εξέχοντα ρόλο σε διαφορετικά είδη καρκίνου. Τα γονίδια αυτά διαμεσολαβούν στο μονοπάτι των MAPK κινασών και σε πολλές από τις περιπτώσεις πολλαπλού μυελώματος έχουν απορρυθμισμένη δράση. (πάνω από 10%) Ασθενείς με μετάλλαξη στο γονίδιο KRAS εμφανίζουν επιθετικότερη μορφή πολλαπλού μυελώματος ενώ έχουν γίνει προσπάθειες αναστολής της δράσης τους φαρμακευτικά. Ένα λιγότερο συχνά αναμεμιγμένο γονίδιο των MAPK είναι και το BRAF (περίπου 5%) με δράση κωδικοποίησης κινάσης σερίνης/θρεονίνης. Το συγκεκριμένο έχει βρεθεί να εμπλέκεται σε αρκετούς καρκίνους μεταξύ άλλων και τη λευχαιμία αλλά δεν έχει αποσαφηνιστεί η αξία του στην πρόκληση πολλαπλού μυελώματος.[21] Όσον αφορά στο μονοπάτι των JAK-STAT κινασών, πρόκειται για υποδοχείς οι οποίοι έχουν τη δυνατότητα μετά από τη σύνδεσή τους με έναν επαγωγικό εξωκυττάριο παράγοντα, να αυτοφωσφορυλιώνονται και στη συνέχεια μια ομάδα πρωτεϊνών ονόματι STAT να συνδέονται στις φωσφορυλιωμένες περιοχές, να δημιουργούν ένα συσσωμάτωμα φωσφορικών ομάδων και STAT πρωτεϊνών και να μεταναστεύουν στον πυρήνα ρυθμίζοντας την έκφραση γονιδίων. Ο κύριος επαγωγικός παράγοντας των JAK-STAT υποδοχέων κινασών αποτελεί η ιντερλευκίνη 6. Σε πάνω από τις μισές περιπτώσεις πολλαπλού μυελώματος έχει διαπιστωθεί υπέρμετρη λειτουργία των υποδοχέων αυτών και συγκεκριμένα εκείνου του τύπου που συνδέεται με την STAT 3. Ως εκ τούτου, το κύτταρο οδηγείται σε υπερενεργοποίηση του γονιδίου που κωδικοποιεί την πρωτεΐνη BCL-XL, η οποία σχετίζεται με την απόπτωση και ειδικότερα ότι την καταστέλλει. Στην περίπτωση του πολλαπλού μυελώματος, η συγκεκριμένη πρωτεΐνη έχει βρεθεί ότι ασκεί επιπλέον αντίσταση σε σχήματα χημειοθεραπείας χωρίς όμως να είναι γνωστός ο τρόπος που αυτό επιτελείται. Έχει δοκιμαστεί η χρήση φαρμάκων με στόχο την καταστολή της λειτουργίας της STAT3 και είχε ως αποτέλεσμα τη φυσιολογική απόπτωση του κυττάρου έναντι της αντίστασης που προβάλλει όταν η BCL-XL είναι υπερλειτουργική σε ασθενείς με πολλαπλό μύελωμα.[18] Ένα άλλο μονοπάτι που ενεργοποιείται από την ιντερλευκίνη 6 και τον αυξητικό παράγοντα ινσουλίνης τύπου 1, είναι εκείνο της κινάσης της φωσφατιδυλοϊνοσιτόλης. Η κινάση αυτή φωσφορυλιώνει και έτσι ενεργοποιεί μια σειρά διαφορετικών στόχων, μεταξύ άλλων και την AKT, μια κινάση με ενεργότητα σερίνης θρεονίνης. Εκείνη στη συνέχεια φωσφορυλιώνει ένζυμα όπως τα mTOR, GSK-3B και FKHR



όπου σχετίζονται και αυτά με την αυξημένη αντοχή σε απόπτωση. Το μονοπάτι της κινάσης της φωσφατιδυλοϊνοσιτόλης είναι μεταλλαγμένο σε περισσότερους από τους μισούς ασθενείς με πολλαπλό μυέλωμα, προσδίδοντας αυξημένη αντίσταση στο φυσιολογικό προγραμματισμένο κυτταρικό θάνατο.

Ο καρκίνος είναι ένα νόσημα που χαρακτηρίζεται από αρκετές μεταλλάξεις σε επίπεδο DNA οι οποίες προϋποθέτουν αρκετό χρόνο για να συσσωρευθούν και να εκφραστούν φαινοτυπικά. Ένα κάρδιο γονίδιο που υφίσταται αλλοίωση είναι και το PARG1. Το γονίδιο αυτό κωδικοποιεί το ομώνυμο ένζυμο το οποίο εμπλέκεται στο κόψιμο του DNA σε διαδικασίες επιδιόρθωσης. Σε ασθενείς με πολλαπλό μυέλωμα φαίνεται να σχετίζεται με αυξημένη δραστηριότητα και κατ'επέκταση σε διατάραξη της φυσιολογικής δομής του DNA. Έχει επιχειρηθεί η χρήση ειδικών αναστολέων του συγκεκριμένου ενζύμου με θετικές εκβάσεις όσον αφορά στα επίπεδα έντασης της ασθένειας. Ένα ακόμη γονίδιο, το DIS3, παρουσιάζει μεταλλαγές σε ασθενείς πολλαπλού μυελώματος. Σε έρευνα που πραγματοποιήθηκε σε 38 ασθενείς, βρέθηκαν αρκετές διαφορετικές μεταλλάξεις σε αυτό το γονίδιο του οποίου ο φυσιολογικός ρόλος έγκειται στην κωδικοποίηση RNAάσης όπου δρα κατά το σχηματισμό πρωτεϊνών των εξωσωμάτων, κατά το σχηματισμό δηλαδή κυστιδίων που περιβάλλονται από πρωτεϊνικά πλέγματα προκειμένου να μεταναστεύσουν από το κύτταρο. Στους περισσότερους ασθενείς βρέθηκε μετάλλαξη του ενός εκ των δύο αλληλομόρφων του γονιδίου με ταυτόχρονη πλήρη απουσία του δεύτερου. Επιπλέον, η τοποθεσία του είναι στο χρωμόσωμα 13 και μάλιστα σχετικά κοντά με τη θέση που εντοπίζεται το RB1 που όπως αναφέρθηκε έχει άμεση επίδραση στο πολλαπλό μυέλωμα. Η θεωρία πίσω από τη συνεισφορά του στο πολλαπλό μυέλωμα είναι πιθανώς η ελλατωματική κωδικοποίηση των πρωτεϊνών που είναι απαραίτητη για το πρωτεϊνικό πλέγμα του εξωσώματος μέσω της πρότερης παραγωγής ελλατωματικού RNA και κατά συνέπεια συσσώρευση των στοιχείων που επρόκειτο να μεταφερθούν στο αρχικό κύτταρο σε μεγάλες ποσότητες. Επιπλέον, το συγκεκριμένο γονίδιο φαίνεται πως το mRNA που παράγει, μεταφέρεται μέσω του εξωσώματος και σε άλλα κύτταρα. Έτσι μεταλλάξεις που το τροποποιούν πιθανώς να οδηγούν σε προβληματική παραγωγή πρωτεϊνών και σε διαφορετικά κύτταρα πέραν της αρχικής εστίας και δημιουργία μυελώματος. Και οι επιγενετικές αλλαγές διαδραματίζουν σημαντικό ρόλο στην πρόκληση πολλαπλού μυελώματος. Η μεθυλίωση του DNA πρόκειται για διαδικασία κατά την οποία δινοκυλεοτίδια CpG(κυτοσίνη-γουανίνη)δέχονται μεθυλομάδες (-CH3) μεταβάλλοντας τη δραστηριότητα συγκεκριμένων περιοχών του DNA. Όπως και στα περισσότερα είδη καρκίνων έτσι και στο πολλαπλό μυέλωμα έχουν εντοπιστεί νησίδες DNA με εκτεταμένη μεθυλίωση αλλά και άλλες με υπομεθυλίωση. Ως εκ τούτου, όλες αυτές οι περιοχές παρουσιάζουν αυξομειώσεις στην έκφραση των αντίστοιχων γονιδίων που επικαλύπτουν, είτε αυτές βρίσκονται πάνω στην αλληλουχία του ίδιου του DNA του γονιδίου είτε σε κάποιες από τις πολλές ρυθμιστικές περιοχές του. Στο πολλαπλό μυέλωμα έχει βρεθεί κυρίως υπερμεθυλίωση σε περιοχές γονιδίων εμπλεκόμενων με την κυτταρική σηματοδότηση και μάλιστα εμφανίζουν διακριτά πρότυπα τα οποία διερευνώνται τη δεδομένη στιγμή και σίγουρα σχετίζονται με τη μετάβαση από το στάδιο του MGUS στο πλήρως δραστήριο πολλαπλό μυέλωμα μιας και εμπλέκονται στους μηχανισμούς επιβίωσης, επιδιόρθωσης του DNA και ανταλλαγής γενετικού υλικού. Οι ιστόνες, πρωτεΐνες οι οποίες συνδέονται με DNA για να σχηματίσουν το πρώτο επίπεδο οργάνωσης του DNA, τη χρωματίνη, συνδέονται και αυτές με ορισμένες μεταλλαγμένες τροποποιήσεις που οδηγούν στο νόσημα. Ένα ομοιωτικό γονίδιο, το HOXA9, του οποίου ο φυσιολογικός ρόλος έγκειται στη μορφογένεση και την κυτταρική δια-

φοροποίηση, παρουσιάζει συχνά μεταλλάξεις σε ασθενείς. Φυσιολογικά κωδικοποιεί ένα μεταγραφικό παράγοντα ο οποίος όταν συνδεθεί με το DNA οδηγεί σε ρύθμιση σειράς λειτουργιών σχετικές με τα παραπάνω και η ενεργότητα του καθορίζεται από επιγενετικές μετατροπές σε ιστονικές ουρές. Στο πολλαπλό μυέλωμα φαίνεται να υπερεκφράζεται και πιθανώς αυτό να έχει αντίκτυπο στον ατελέσφορο κυτταρικό πολλαπλασιασμό του νοσήματος ενώ οι παρεκκλίσεις αυτές είναι το αποτέλεσμα ιστονικών τροποποιήσεων. Επίσης δείχτηκε πως απώλεια του γονιδίου σε έρευνες τύπου knockout, οδηγεί σε απορρύθμιση της κυτταρικής διαφοροποίησης, γεγονός που ενέτεινε τις εικασίες ενοχοποίησής του στο πολλαπλό μυέλωμα. [18][19][21]

## 1.7 Συμπτώματα, κριτήρια διαφοροδιάγνωσης και θεραπεία

Το πολλαπλό μυέλωμα είναι ένα νόσημα που πολλές φορές λειτουργεί αθόρυβα ενώ γίνεται αντιληπτό φαινοτυπικά όταν είναι πλέον πολύ αργά. Σε αρκετές περιπτώσεις ο λόγος που περνάει απαρατήρητο είναι διότι αφενός δεν εμφανίζει έντονα συμπτώματα παρά μόνο όταν πλέον βρίσκεται σε προχωρημένο και συνήθως μη αναστρέψιμο στάδιο και αφετέρου γιατί η διαφοροδιάγνωση του από άλλα παρόμοια σε βιοχημικές τιμές σύνδρομα δυσκολεύουν την έγκαιρη διάγνωση του. Με βάση το μνημονικό κανόνα CRAB υπάρχουν τέσσερα κύρια συμπτώματα που όταν εμφανιστούν αξίζουν διερεύνησης προκειμένου να διαπιστωθεί η ύπαρξη πολλαπλού μυελώματος. Αυτά είναι η υπερασβεστιαμία, η νεφρική ανεπάρκεια τελικού σταδίου, η αναιμία και οι οστικές κακώσεις. Από μόνα του βέβαια αυτά τα στοιχεία δεν αποδεικνύουν πως σίγουρα η ασθένεια υφίσταται. Έχουν αναπτυχθεί και κάποιοι επιπλέον βιοχημικοί έλεγχοι με στόχο να εντοπιστούν συγκεκριμένοι βιοδείκτες οι τιμές των οποίων μπορούν να βοηθήσουν στη διαγνωστική διαδικασία. Σε αυτούς ανήκουν η μέτρηση του αριθμού των πλασματοκυττάρων στο μυελό των οστών το ποσοστό των οποίων φυσιολογικά δε θα πρέπει να ξεπερνά το 60% του συνόλου των κυττάρων, ο αριθμός των ελεύθερων ελαφριών αλυσίδων στον ορό του αίματος καθώς επίσης και ανησυχητικά ευρήματα σε CT, MRI ή PET-CT. [22]

Ακολουθεί ένας πίνακας όπου παρουσιάζονται οι βασικές τιμές του μνημονικού κανόνα CRAB που χρησιμοποιείται προς επικύρωση ύπαρξης πολλαπλού μυελώματος:

**Πίνακας 1.4** *Updated Diagnostic Criteria and Staging System for Multiple Myeloma. American Society of Clinical Oncology Educational Book, (36), e418–e423.* Οι τιμές των βιοδεικτών που συνοδεύουν τη διάγνωση πολλαπλού μυελώματος κατά το σύστημα CRAB και τα ενδεικτικά εύρη αυτών. [24]

### Κριτήρια CRAB

• Υπερασβεστιαμία με τιμές ασβεστίου ορού >0.25mmol/L πάνω από το φυσιολογικό όριο των 2.75mmol/L
• Τελική νεφρική ανεπάρκεια με δείκτη κάθαρσης κρεατινίνης GFR <40ml/min ή κρεατινίνη ορού >117μmol/L
• Αναιμία με επίπεδα >2g/dl κάτω από το φυσιολογικό κατώτατο όριο ή τιμές <10g/dl
• Τουλάχιστον μία οστική κάκωση διαγνωσμένη από εξέταση MRI, CT ή PET-CT

Όπως αναφέρθηκε, η διαφοροδιάγνωση είναι ιδιαίτερα δύσκολη καθώς υπάρχουν σύνδρομα με πολύ πανομοιότυπη μορφή ως προς τα βιοχημικά ευρήματα. Ο παρακάτω πίνακας θέτει τις λεπτές γραμμές που διαφοροποιούν το πολλαπλό μυέλωμα από τους κοντινούς συγγενείς του:

**Πίνακας 1.5** *Multiple Myeloma: Diagnosis and Treatment. Mayo Clinic Proceedings, 91(1), 101–119.* Ο συγκεκριμένος πίνακας παρουσιάζει αναλυτικά τα στοιχεία που χρειάζονται προκειμένου να διενεργηθεί σύγκριση μεταξύ πολλαπλού μυελώματος και παρόμοιων νοσημάτων. Να σημειωθεί ότι οι περισσότερες από τις συγγενείς αυτές ασθένειες δεν παρουσιάζουν συμπτώματα, ενώ εμφανίζουν βιοχημικές τιμές και κλινικά ευρήματα σε πανομοιότυπα επίπεδα, περιπλέκοντας περαιτέρω τη διαφοροδιάγνωση. [23]

<p><u>Διαταραχή</u></p> <p>Μη IgM μονοκλωνική γαμμοπάθεια ανεξιχνίαστης σημασίας (MGUS)</p>	<p><u>Περιγραφή ασθένειας</u></p> <p><u>Πρέπει να πληρούνται τα παρακάτω:</u></p> <ul style="list-style-type: none"> <li>▪ Ορός μονοκλωνικής πρωτεΐνης (όχι τύπου IgM) &lt; 3g/dL</li> <li>▪ Κλώνος πλασματοκυττάρων οστού &lt; 10%</li> <li>▪ Απουσία καταστροφής οργάνων τελικού σταδίου με συνοδά συμπτώματα υπερκαλιαιμίας, νεφρικής ανεπάρκειας, οστικών κακώσεων ή αναιμίας (CRAB)</li> </ul>
<p>Υποδόσκον πολλαπλό μυέλωμα</p>	<p><u>Πρέπει να πληρούνται τα παρακάτω:</u></p> <ul style="list-style-type: none"> <li>▪ Ορός μονοκλωνικής πρωτεΐνης (IgG ή IgA) &gt; 3g/dL ή μονοκλωνική πρωτεΐνη ούρων &gt; 500mg/24h με παράλληλο αριθμό πλασματοκυττάρων μυελού των οστών 10%-60%</li> <li>▪ Απουσία καθοριστικών για το μυέλωμα λοιπών στοιχείων ή στοιχείων που υποδεικνύουν αμυλοείδωση</li> </ul>
<p>Πολλαπλό μυέλωμα</p>	<p><u>Πρέπει να πληρούνται τα παρακάτω:</u></p> <ul style="list-style-type: none"> <li>• Πλασματοκύτταρα μυελού των οστών &gt; 10% ή βιοψία που να</li> </ul>

αποδεικνύει την ασθένεια με βάση βιοχημικές τιμές.

- Ενδείξεις για τελικού επιπέδου βλάβη σε όργανα που μπορεί να αποδοθεί σε ασθένεια εξάπλωσης πλασματοκυττάρων με συνοδά στοιχεία:
  - 1)Υπερασβεστιαμία με επίπεδα στον ορό του αίματος $>0.25\text{mmol/L}$
  - 2)Νεφρική ανεπάρκεια με κάθαρση κρεατινίνης $<40\text{mL/min}$  ή κρεατινή στον ορό του αίματος με τιμές $>2\text{mg/dl}$
  - 3)Αναιμία με τιμές αιμοσφαιρίνης $>2\text{g/dL}$  πάνω ή κάτω του φυσιολογικού (περίπου  $12\text{g/dl}$ ) ή τιμή αιμοσφαιρίνης $<10\text{g/dL}$
  - 4)Οστικές λύσεις με μία ή παραπάνω αποτυπωμένες σε ακτινογραφία οστών ή αξονική τομογραφία ή τομογραφία εκπομπής ποζιτρονίων
  - 5)Ποσοστό πλασματοκυττάρων στο μυελό των οστών $>60\%$ Αριθμός ελαφριών ελεύθερων αλυσίδων στον ορό του αίματος $>100\text{mg/L}$

MGUS συνδεόμενη με παρουσία ελαφριών αλυσίδων

στο αίμα

Πρέπει να πληρούνται τα παρακάτω:

- Αποκλίσεις στις τιμές των ελαφριών ελεύθερων αλυσίδων ( $<0.256$  ή  $>1.65$ )
- Αυξημένα επίπεδα κ ελαφριών αλυσίδων $>1.65$  με ταυτόχρονη αυξημένη παρουσία λ ελαφριών αλυσίδων $>0.26$  του φυσιολογικού
- Μη έκφραση βαριών αλυσίδων ανοσφαιρινών
- Απουσία βλάβης οργάνων τελικού επιπέδου που μπορεί να αποδοθεί σε ασθένεια εξάπλωσης πλασματοκυττάρων
- Ποσοστό πλασματοκυττάρων μυελού των

	οστών<10%Μονοκλωνική πρωτεΐνη ούρων<500mg/24h
IgM μονοκλωνική γαμμοπάθεια ανεξιχνίαστης σημασίας(MGUS)	<p><u>Πρέπει να πληρούνται τα παρακάτω</u></p> <ul style="list-style-type: none"> <li>• IgM μονοκλωνική πρωτεΐνη στον ορό του αίματος&lt;3g/dL</li> <li>• Ποσοστό διείσδυσης μυελού των οστών από πλασματοκύτταρα&lt;10%</li> <li>• Μη ύπαρξη ενδείξεων για αναιμία,αλλοίωσης επιμέρους οργάνων,αύξηση του ιξώδους,λεμφαδενοπάθεια,ηπατοσπληνομεγαλία που μπορεί να αποδοθεί σε κάποιο υποβόσκον νόσημα που σχετίζεται με ανώμαλη επέκταση λεμφοκυττάρων</li> </ul>

Η αντιμετώπιση του πολλαπλού μυελώματος ποικίλει και σχετίζεται τόσο με την ένταση των συμπτωμάτων όσο και με το στάδιο στο οποίο βρίσκεται. Το βασικότερο όλων, όπως και σε όλες τις περιπτώσεις καρκίνου, είναι η έγκαιρη διάγνωση όσο κατά το δυνατόν αυτό είναι εφικτό πριν τα όργανα στόχοι πληχθούν ανεπανόρθωτα. Υπάρχουν αρκετές διαφορετικές μέθοδοι προσέγγισης ανάλογα πάντα με την περίπτωση. Ορισμένα από τα φάρμακα εκλογής που χρησιμοποιούνται πολλές φορές συνδυαστικά είναι τα παρακάτω: Η λεναλιδομίδη με δράση αναστολέα αγγειογένεσης και κυτοκινών παραγόμενων από τον όγκο ενώ επίσης λειτουργεί ως επαγωγέας της υγιούς απόπτωσης με στόχο την προστασία από την περαιτέρω ογκογένεση, η δεξαμεθαζόνη, φάρμακο με ανοσοκατασταλτική δράση, η μπορτεζομίμη όπου αναστέλλει τη δράση του πρωτεοσώματος στο κυτταρόπλασμα και η κυκλοφωσφαμίδη με επίσης ανοσοκατασταλτική δράση δημιουργώντας επιθυμητές βλάβες στο DNA των καρκινικών κυττάρων προκειμένου σε δεύτερο χρόνο αυτά να καταστραφούν μέσω της απόπτωσης. Ένα λιγότερο συχνά χρησιμοποιούμενο φάρμακο είναι η θαλιδομίδη όπου ναι μεν έχει ογκοκατασταλτική δραστηριότητα μέσω μείωσης του παράγοντα TNF-α (tumor necrosis factor α) αλλά ωστόσο παρουσιάζει αρκετές παρενέργειες όπως τερατογένεση σε κυήματα και σχηματισμό θρόμβων αίματος οι οποίοι κατά την λύση τους μπορεί να αποβούν ακόμα και θανατηφόροι. Το τελευταίο για να συμβεί απαιτείται μεγάλη ποσότητα του φαρμάκου ενώ οι δοσολογίες είναι αυστηρά ορισμένες στα διάφορα κλινικά θεραπευτικά σχήματα οπότε είναι δύσκολο να λάβει χώρα. Συνήθως οι

διαφορετικές θεραπευτικές οδοί χρησιμοποιούν ένα ανοσοκατασταλτικό φάρμακο προκειμένου να κατασταλούν οι υπέρμετρες αντιδράσεις από τη μεριά του οργανισμού μέσω της αναστολής κυτοκινών, τ λεμφοκυττάρων και άλλων παραγόντων και ένα αντινεοπλασματικό φάρμακο για να περιορίσουν την αγγειογένεση του νεοσυντιθέμενου όγκου ή να καταστρέψουν κάποια από τα επιμέρους στοιχεία του κυττάρου(μεμβράνη,οργανίδια). Σε πολλαπλό μυέλωμα υψηλού κινδύνου χρησιμοποιείται τα τελευταία χρόνια το σχήμα KRd που απαρτίζεται από καρφιλζομίμη, δεξαμεθαζόνη και λεναλιδομίδη. Ωστόσο το σχήμα βρίσκεται ακόμα υπό κλινικές δοκιμές ενώ έχει χορηγηθεί σε ασθενείς τελικού σταδίου όπου όμως εμφάνισε άμβλυνση της έντασης του νοσήματος αλλά όχι ίαση. Επίσης το φάρμακο καρφιλζομίμη εμφανίζει σημαντική καρδιοτοξικότητα.[23] Ασθενείς που παρουσιάζουν τελική βλάβη σε όργανα συνήθως δεν επανακάμπτουν ανεξαρτήτως της χορήγησης θεραπευτικής αγωγής. Το προσδόκιμο ζωής σε αυτές τις περιπτώσεις, με τα σύγχρονα σχήματα, μπορεί να προσεγγίσει την πενταετία. Συνεπώς με τα συμβατικά φάρμακα μιλάμε για παράταση της ζωής και όχι εξάλειψη της ασθένειας. Αξίζει να αναφερθεί ότι ένα από τα φάρμακα που χρησιμοποιείται ευρέως για την αντιμετώπιση του πολλαπλού μυελώματος είναι η μπορτεζομίμη. Συνδυαστικά με κάποιο ανοσοκατασταλτικό, η μπορτεζομίμη είναι το φάρμακο εκλογής σε περιπτώσεις επανεμφάνισης της ασθένειας εάν αυτή είχε υποχωρήσει. Κάποια συχνά χρησιμοποιούμενα σχήματα είναι το RD όπου περιλαμβάνει λεναλιδομίδη μαζί με δεξαμεθαζόνη, το VD με μπορτεζομίμη και δεξαμεθαζόνη και το VTD όπου πρόκειται για το VD με προσθήκη είτε θαλιδομίδης είτε κυκλοφωσφαμίδης.[23][24] Μια λιγότερο συμβατική θεραπευτική οδός είναι η αυτόλογη μετάγγιση εμβρυϊκών κυττάρων αίματος. Στην περίπτωση αυτή, χρησιμοποιούνται υγιή πολυδύναμα αιμοποιητικά κύτταρα που προέρχονται από τον ίδιο τον ασθενή προκειμένου να ενεθούν και να αναπληρώσουν τις κατεστραμμένες αποθήκες του μυελού των οστών. Κατά τη χρήση χημειοθεραπευτικών φαρμάκων, πολλά από τα φυσιολογικά κύτταρα του οργανισμού και σε αυτήν την περίπτωση κυρίως του μυελού των οστών, έχουν αποδεκατιστεί ως προς τον αριθμό τους με αποτέλεσμα τα όργανα να υπολειπώνονται. Επιπλέον ο μυελός των οστών αποτελεί μαζί με το σπλήνα το κέντρο της αιμοποίησης. Συνεπώς η απέκδυση του μυελού από κύτταρα μειώνει κατά πολύ τον όγκο του διαθέσιμων ερυθροκυττάρων αλλά και των λοιπών κυττάρων του αίματος. Σε πολλές περιπτώσεις πολλαπλού μυελώματος όπου έχει λάβει χώρα έγκαιρα η διάγνωση, προσφέρεται η δυνατότητα φύλαξης των υγιών κυττάρων και χρησιμοποίησής τους σε απώτερο χρόνο. Ένα πολύ σημαντικό πλεονέκτημα αυτού του είδους θεραπείας είναι ότι δεν υπόκειται σε περιορισμούς λόγω θεμάτων ιστοσυμβατότητας και επομένως δεν υπάρχει περίπτωση να απορριφθούν από τον οργανισμό ή να υπάρξει ανάγκη για τεχνητή ανοσοκαταστολή.[23][24] Τέλος, υπάρχει και η παρηγορητική θεραπεία για ασθενείς πολλαπλού μυελώματος σε τελευταίο στάδιο. Αυτή περιλαμβάνει τη ψυχολογική υποστήριξη, η οποία βέβαια οφείλει να εφαρμόζεται ακόμη και στους καλύτερα ανταποκρινόμενους ασθενείς και μια σειρά από φάρμακα για την κατά συνθήκη αντιμετώπιση των εκδηλώσεων της νόσου όπως τα μη στεροειδή αντιφλεγμονώδη και διάφορα οπιούχα παυσίπονα.

## 1.8 Βιοδείκτες στην Πρωτεομική του καρκίνου

Στην προσπάθεια τους οι επιστήμονες να αποκτήσουν ολοένα και περισσότερο εξειδικευμένα και ακριβή μέσα για την ανίχνευση ασθενειών σε όσο το δυνατότερο πρωιμότερο στάδιο, τις τελευταίες δεκαετίες έχουν αναπτύξει βιοχημικές τεχνικές που τους επιτρέπουν να διερευνούν για βιολογικά στοιχεία τα οποία χαρακτηρίζουν ένα νόσημα. Τα στοιχεία αυτά ονομάζονται βιοδείκτες, είναι ενδεικτικά του νοσήματος και στόχος τους είναι να χρησιμοποιούνται ως σημεία αναφοράς για την επιβεβέβαιωση της εμφάνισης σειράς προβλημάτων υγείας στον οργανισμό. Υπάρχουν πολλά είδη βιοδεικτών τόσο σε επίπεδο γενετικής(πολυμορφισμοί, μεταλλαγμένα γονίδια)όσο και σε επίπεδο πρωτεομικής(πεπτίδια)τα οποία έχουν πλέον άμεση κλινική σημασία μιας και η έγκαιρη διάγνωση αποτελεί το μισό της επιτυχίας κατά την αντιμετώπιση των διαφόρων ασθενειών. Τα στοιχεία με βάση τα οποία κρίνονται οι βιοδείκτες είναι κυρίως δύο:η ευαισθησία όπου μετρά το μέγεθος της πιθανότητας όταν η ασθένεια υπάρχει ο δείκτης επίσης να εντοπίζεται και η ειδικότητα όπου μετρά το μέγεθος της πιθανότητας όταν η ασθένεια δεν υπάρχει ο βιοδείκτης επίσης να μην υπάρχει.Είναι προφανές πως το επιθυμητό αποτέλεσμα πάντα είναι αυτό που προσεγγίζει τη μονάδα και στις δύο περιπτώσεις και όσο πιο κοντά είναι ο λόγος σε αυτήν τόσο μεγαλύτερη και η κλινική σημασία του βιοδείκτη. Υπάρχουν ωστόσο αρκετά εμπόδια κατά τις μελέτες βιοδεικτών όπως για παράδειγμα περιορισμοί που πηγάζουν από τον εξοπλιστικό μηχανισμό, τα συστήματα βιοπληροφορικής καθώς επίσης και από λάθη κατά τη διενέργεια των εκάστοτε μελετών λόγω ανεπαρκούς δείγματος, προσωπικού σφάλματος κ.α. Επιπλέον, για να χαρακτηριστεί ένας βιοδείκτης ως βέλτιστος και αξιόπιστος, θα πρέπει να εμφανίζεται και να εντοπίζεται με την ίδια συχνότητα σε ασθενείς μιας συγκεκριμένης νόσου, ανεξάρτητα της ηλικίας, του φύλου και της φυλετικής καταγωγής.Έτσι λοιπόν η διαδικασία ανάπτυξης νέων βιοδεικτών παρουσιάζει μεγάλες προκλήσεις. Ωστόσο η ανάγκη για ολοένα και ταχύτερη αλλά και αποτελεσματικότερη θεραπεία έναντι των διαφόρων καρκινικών μορφών καθιστά επιτακτική τη διάνοιξη νέων διαγνωστικών οδών ανάμεσα στις οποίες οι βιοδείκτες κατέχουν εξέχοντα ρόλο.[25][27]

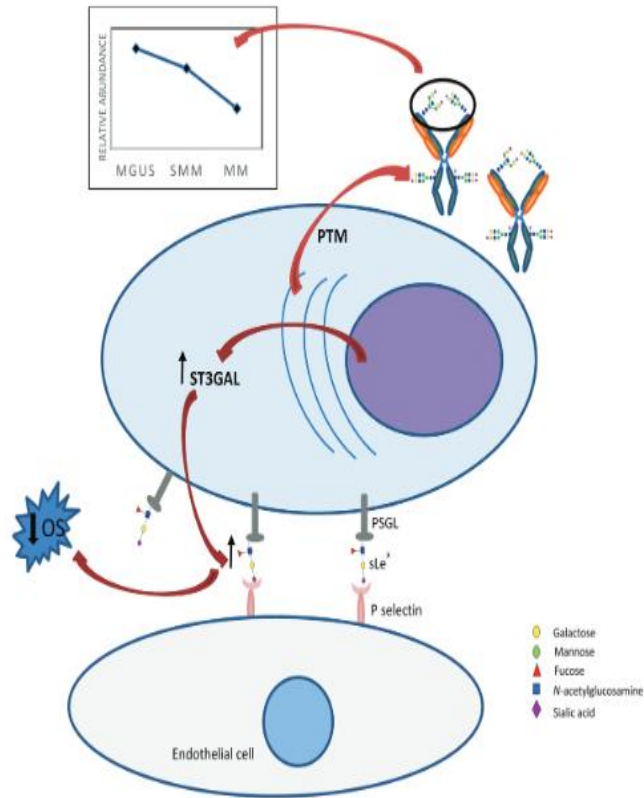
Στον καρκίνο, πολλά από τα επιμέρους στοιχεία που απαρτίζουν ένα κύτταρο, παρουσιάζουν τροποποιήσεις οι οποίες μπορούν να αξιοποιηθούν για τη στοιχειοθέτηση ενός συγκεκριμένου νοσήματος. Ειδικότερα στην πρωτεομική μελέτη, ένα από τα σημαντικότερα σημεία είναι το γεγονός πως οι ποσοτικές μεταβολές που συμβαίνουν στο φυσιολογικό πεπτίδιο ή ολόκληρη πρωτεΐνη είναι πολύ διαφορετικές από τις ποσοτικές μεταβολές σε τροποποιημένα πεπτίδια ή ολόκληρες πρωτεΐνες. Γνωρίζοντας ποια είναι αυτή η βιολογική τροποποίηση, μπορούμε να υπολογίσουμε το λόγο της φυσιολογικής μορφής προς τη μεταλλαγμένη και στη συνέχεια, δεδομένου ότι η υπόθεση μας έχει αρκετά στοιχεία και επαρκή επαναληψιμότητα ως προς το επιθυμητό αποτέλεσμα, να αποφασίσουμε για το αν ένα πεπτίδιο θα μπορούσε να χρησιμοποιηθεί ως δυναμικός βιοδείκτης για τη συγκεκριμένη ασθένεια. Μερικές από αυτές τις τροποποιήσεις είναι οι ακετυλιώσεις, οι μεθυλιώσεις, οι φωσφορυλιώσεις και οι γλυκοζυλιώσεις.[26]

Η πρωτεομική είναι ένα από τα μεγαλύτερα κομμάτια της επιστήμης που ασχολείται με την ανακάλυψη νέων βιοδεικτών στον καρκίνο. Τα καρκινικά κύτταρα έχουν την τάση να αλλάζουν το μεταβολικό προφίλ τους σε σχέση με τα υγιή, με αποτέλεσμα να παρατηρούνται αλλαγές στην παραγωγή και μεταφορά διαφορετικών πρωτεϊνών και πεπτιδίων. Το αίμα, μαζί με το πλάσμα που συνιστά το 55% του συνολικού του όγκου, αποτελεί μια από τις καλύτερες δεξαμενές για την αναζήτηση νέων βιοδεικτών καθώς είναι το κύριο υγρό του σώματος μέσω του οποίου μεταφέρονται σχεδόν όλα τα διαφορετικά πεπτιδία και πρωτεϊνικά παράγωγα. Επίσης ένας καλός υποψήφιος για ανεύρεση βιοδεικτών αποτελεί και κάποιο τμήμα ενός όγκου που προέρχεται από μια βιοψία. Σε κάθε περίπτωση, η διαδικασία ξεκινάει με προετοιμασία του δείγματος για φασματοσκοπία μάζας, τρέξιμο της φασματοσκοπίας και στη συνέχεια βιοπληροφορική ανάλυση. Η ανάπτυξη της τεχνολογίας φασματομετρίας (maldi, quadrupole, ion orbit trap) αλλά και βιοπληροφορικής επιτρέπει την ποσοτική ανάλυση ενός συγκεκριμένου πεπτιδίου και μας δίνει τη δυνατότητα να αξιολογήσουμε τη σημασία του ως σημείο αναφοράς για ένα νόσημα. [25][26]

Η γλυκοζυλίωση είναι μια μεταμεταφραστική τροποποίηση πρωτεϊνών η οποία είναι αρκετά συνήθης στον καρκίνο. Μάλιστα στην περίπτωση που το άτομο νοσεί, εμφανίζονται στικτά πρότυπα γλυκοζυλίωσης, αρκετά διαφορετικά από εκείνα των φυσιολογικών κυττάρων και κατ'επέκταση ο εντοπισμός τους τους προσδίδει μεγάλη στατιστική αξία ως πιθανούς βιοδείκτες. Όσον αφορά στο πολλαπλό μυέλωμα, έχουν διεξαχθεί αρκετές έρευνες ως προς τη γλυκοζυλίωση ανοσοσφαιρινών και πρωτεϊνών της επιφάνειας του κυττάρου. Οι ανοσοσφαιρίνες συγκεκριμένα περιέχουν αρκετές γλυκοζυλιωμένες περιοχές τόσο στο μεταβλητό όσο και στο σταθερό τμήμα τους. Ωστόσο στην περίπτωση του μυελώματος, αλλάζει σε μεγάλο βαθμό το προφίλ των γλυκοπρωτεϊνών σε περιεκτικότητα ενός σακχάρου που ονομάζεται σιαλικό οξύ. Το οξύ αυτό, που επίσης καλείται και νευραμινικό, διαθέτει μια ελεύθερη καρβοξυλομάδα με αποτέλεσμα να έχει αρνητικό φορτίο. Αυτό είναι ιδιαίτερα σημαντικό καθώς το καθιστά αρκετά ενεργό μέσω ιοντικών αλληλεπιδράσεων με παρακείμενα μόρια και ιδίως με μόρια λεκτινών, πρωτεϊνικών υποδοχέων δηλαδή στην κυτταρική επιφάνεια οι οποίοι αναγνωρίζουν συγκεκριμένους υδατάνθρακες σε γλυκοπρωτεΐνες και συνδέονται μαζί τους. Τα IgG αντισώματα, τα οποία εμφανίζουν αρκετές μεταλλάξεις στο πολλαπλό μυέλωμα και είναι ένα από τα κύρια σημεία της ασθένειας, παρουσιάζουν διαφορές στα πρότυπα γλυκοζυλίωσης τους σε σχέση με τα φυσιολογικά. Έτσι έχουμε αυξημένη προσθήκη καταλοίπων σιαλικού οξέος σε βαριές αλυσίδες των ανοσοσφαιρινών καθώς επίσης και στις ελαφριές. Ένα στοιχείο το οποίο είναι αρκετά αξιοσημείωτο είναι το ότι οι διαφορές στα επίπεδα γλυκοζυλίωσης δεν σχετίζονται μόνο με την ύπαρξη ή όχι της ασθένειας αλλά και με τη σοβαρότητα της και τη μετάβαση από το στάδιο του απλού όγκου σε πολλαπλό μυέλωμα. Για παράδειγμα υπάρχουν τρεις διαφορετικές περιοχές στο αντίσωμα που ονομάζονται A2G2S1, A2BG2S1 και A2BG2S2, των οποίων το ποσοστό γλυκοζυλίωσης είναι συνιφασμένο με την ταξινόμηση της ασθένειας σε ένα εύρος που ξεκινά με την απλή MGUS και καταλήγει στο πολλαπλό μυέλωμα. Συνεπώς, όλα αυτά τα διαφορετικά στάδια γλυκοζυλίωσης θα μπορούσαν να αξιολογηθούν ως πιθανοί βιοδείκτες που να προσδιορίζουν το κατά πόσο έχει εξελιχθεί το νόσημα και να επιτρέπουν την καλύτερη αντιμετώπιση του. Η δράση του σιαλικού οξέος έχει ενοχοποιηθεί σε αρκετούς διαφορετικούς καρκίνους και ειδικότερα για τη συμβολή του στη μετανάστευση καρκινικών κυττάρων. Ένα γονίδιο που ονομάζεται ST3GAL 6 φαίνεται να είναι υπερλειτουργικό στο πολλαπλό μυέλωμα και κωδικοποιεί έναν τύπο σιαλικού οξέος που ονομάζεται



LEWIS. Το συγκεκριμένο έχει ταυτοποιηθεί ως επιβαρυντικό για την επιβίωση ασθενών. Η αυξημένη μεταγραφή του γονιδίου αυτού σε συνδυασμό με την υπερέκφραση μιας σειράς σελεκτινών, κυρίως E και P, οδηγεί σε διευκόλυνση της μετανάστευσης των κυττάρων παρότι δεν είναι σαφής ο τρόπος με τον οποίο συμβάλλει ολόκληρο το κύκλωμα των LEWIS σιαλικού και των σελεκτινών με ακρίβεια[26]. Ακολουθεί μια εικόνα που περιγράφει συνοπτικά το συγκεκριμένο σύστημα:



**Εικόνα 1.8** *Current and future biomarkers for risk-stratification and treatment personalisation in multiple myeloma. Molecular Omics.* Παρουσιάζεται η αυξημένη παραγωγή Lewis σιαλικού οξέος από το γονίδιο *ST3GAL* και η αλληλεπίδραση αυτού με τη σελεκτίνη P ενώ στο μικρό πινακίδιο υπερθεματίζεται η αντιστοιχία των επιπέδων μεταμεταφραστικών τροποποιήσεων με το στάδιο της ασθένειας. Το χαρακτηριστικό προφίλ τροποποιήσεων θα μπορούσε να αποτελέσει έναν καλό αρωγό στην προσπάθεια αναζήτησης νέων βιοδεικτών. [26]

Είναι βέβαιο πως οι βιοδείκτες και το ενδιαφέρον για την ανάπτυξη ολόένα και ειδικότερων, κατέχουν σημαντική θέση στο κομμάτι της διαγνωστικής ιατρικής. Οι πολλές δυσκολίες τόσο σε επίπεδο εντοπισμού όσο και σε επίπεδο καθορισμού της ευαισθησίας που αυτοί προσφέρουν, αποτελούν καίρια εμπόδια κατά τη διεξαγωγή των εκάστοτε ερευνών. Όσο όμως ασθένειες όπως το πολλαπλό μυέλωμα αλλά και εν γένει οι διάφοροι τύποι καρκίνου φαίνεται να δυσχεραίνουν πάρα πολύ το κλινικό έργο μιας και τις περισσότερες φορές δεν υπάρχει τρόπος αντιμετώπισης μέσω των συμβατικών μέσων όπως η χειρουργική και τα καθιερωμένα θεραπευτικά σχήματα, η ανάγκη για

καινούς και περισσότερο ειδικούς βιοδείκτες καθίσταται πιο επιτακτική από ποτέ, καθώς η έγκαιρη διάγνωση μπορεί να συντελέσει καταλυτικά στην έκβαση της κατάστασης της υγείας ενός ασθενούς.[28]

## ΚΕΦΑΛΑΙΟ 2 ΠΡΟΓΡΑΜΜΑΤΑ ΤΗΣ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ ΑΝΑΛΥΣΗΣ

### 2.1 Βιοπληροφορική και Πρωτεομική

Τις τελευταίες δεκαετίες ο ολοένα και αυξανόμενος αριθμός ερευνών και ερευνητικών προγραμμάτων έχει δημιουργήσει μια αδιανόητα υψηλή ποσότητα βιολογικών δεδομένων. Η τεράστια ανάγκη για την ανάλυση και την ερμηνεία των δεδομένων αυτών καλύπτεται από την εξελισσόμενη επιστήμη της βιοπληροφορικής. Η βιοπληροφορική ορίζεται ως το σύνολο των εργαλείων υπολογισμού και ανάλυσης στην καταγραφή και την ερμηνεία των βιολογικών δεδομένων. Είναι ένα διατμηματικό επιστημονικό πεδίο, το οποίο χρησιμοποιεί την επιστήμη υπολογιστών, τα μαθηματικά, τη φυσική καθώς και τη βιολογία. Η βιοπληροφορική θεωρείται πλέον απαραίτητη για τη διαχείριση των δεδομένων στους χώρους της σύγχρονης βιολογίας και ιατρικής.

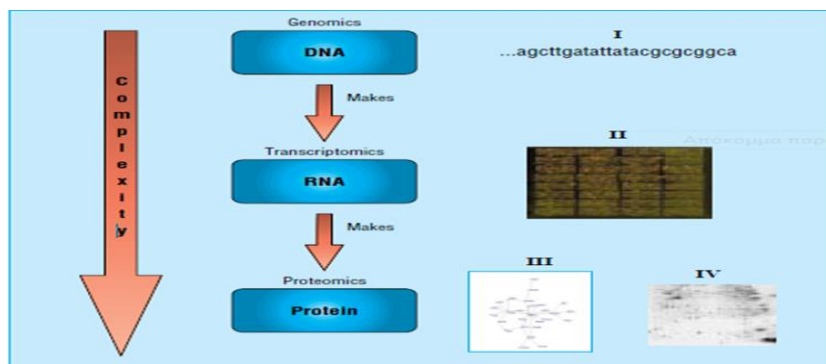
Η συνεχώς αυξανόμενη ποσότητα δεδομένων από την κλινική έρευνα έχει καταστήσει απαραίτητη τη δημιουργία υπολογιστικών βάσεων δεδομένων οι οποίες θα περιλαμβάνουν μια εύχρηστη πύλη πληροφοριών καθώς και προγράμματα λογισμικών αλγορίθμων, τα οποία θα καθιστούν πιο αποδοτική τη διαχείριση τους. Εξαιτίας της ποικιλόμορφης φύσης των αναδύομενων δεδομένων, επί του παρόντος, δεν υπάρχει μια και μοναδική πλήρης βάση δεδομένων η οποία να μπορεί να ικανοποιεί όλα τα ερευνητικά πεδία. Παρόλα αυτά, λόγω της υπέρμετρης ερευνητικής διακλάδωσης, έχουν δημιουργηθεί πλήρως εξειδικευμένες βάσεις. Αυτές περιλαμβάνουν τόσο δημόσιες όσο και ιδιωτικές δεξαμενές βιολογικών πληροφοριών οι οποίες μπορούν να φανούν εξαιρετικά χρήσιμες σε κλινικούς και εργαστηριακούς ερευνητές.

Βασικά εργαλεία όσων ασχολούνται με τη βιοπληροφορική, είναι τα προγράμματα λογισμικού και το ίντερνετ. Ορισμένες από τις πλέον βασικές δραστηριότητες, όπως η ανάλυση αλολουθιών του DNA ή των πρωτεϊνικών προφίλ, γίνεται με τη χρήση διάφορων προγραμμάτων και βάσεων δεδομένων που βρίσκονται διαθέσιμα στο διαδίκτυο. Οποιοσδήποτε έχει πρόσβαση στο διαδίκτυο και στις κατάλληλες ιστοσελίδες, από κλινικούς ερευνητές, μοριακούς βιολόγους έως και απλούς πολίτες, μπορεί πλέον να ανακαλύψει χωρίς χρέωση αποτελέσματα ερευνών που έχουν προκύψει και δημοσιευτεί με τη χρήση βιοπληροφορικών εργαλείων.[29][30]

Η βιοπληροφορική χρησιμοποιείται σήμερα σε ένα εξαιρετικά ευρύ φάσμα σημαντικών διαδικασιών στις οποίες συμπεριλαμβάνεται η ανάλυση της έκφρασης και της μεταβολής των γονιδίων, η ανάλυση και η πρόβλεψη της γονιδιακής και πρωτεϊνικής δομής και λειτουργίας, η πρόβλεψη και η ανίχνευση των γονιδιακών ρυθμιστικών δικτύων, η προσομοίωση περιβαλλόντων για τη δημιουργία ολόκληρων

κυτταρικών μοντέλων, η δημιουργία σύνθετων μοντέλων των ρυθμιστικών δυναμικών και δικτύων των γονιδίων και η παρουσίαση και η ανάλυση των μοριακών οδών, με στόχο την κατανόηση των σύνθετων αλληλεπιδράσεων μεταξύ γονιδίων και νοσημάτων.[29][30] Σε πιο μικρή κλίμακα, υπάρχουν απλούστερες βιοπληροφορικές διαδικασίες οι οποίες μπορούν να φανούν χρήσιμες στον κλινικό ερευνητή. Οι διαδικασίες αυτές ποικίλουν, από το σχεδιασμό εκκινητών (βραχείες ολιγονουκλεοτιδικές αλληλουχίες που χρησιμοποιούνται για την ενίσχυση του DNA σε πειράματα αλυσιδωτής αντίδρασης πολυμεράσης), έως σύγχρονα προγράμματα βιοχημικών αναλύσεων.(εξετάσεις DNA,βιοχημικές τιμές κ.α)

Από την ολοκλήρωση της χαρτογράφησης του ανθρώπινου γονιδιώματος κι έπειτα, η έμφαση της μελέτης των ερευνητών έχει μετακινηθεί από τα γονίδια αυτά καθαυτά προς την κατεύθυνση των προϊόντων τους. Η γενομική(ή γονιδιωματική) αφορά καθαρά και μόνο στη γονιδιακή υπόσταση ενώ τα επόμενα βήματα μελετώνται από τους αλληλοσυνδεόμενους κλάδους της μεταγραφομικής, της πρωτεομικής και της μεταβολομικής. Συνεπώς, άρχισε σιγά σιγά η εστίαση στην κατανόηση της δομής και της λειτουργίας αυτών των γονιδιακών προϊόντων, με τις πρωτεΐνες να κατέχουν εξέχοντα ρόλο στο ερευνητικό γίγνεσθαι. Αυτή η μετάβαση άρχισε να λαμβάνει χώρα ύστερα από την κατανόηση του ότι η ανάλυση και η ερμηνεία των βιολογικών δεδομένων δε στηρίζεται αποκλειστικά στην πληροφορία που αποδίδει η γονιδιωματική, αλλά και στο επίπεδο του πρωτεόματος και του μεταγραφώματος, ανοίγοντας τους ορίζοντες για μια περισσότερο ολιστική προσέγγιση των πραγμάτων. Η πρωτεομική είναι η ανάλυση του συνολικού αριθμού των πρωτεϊνών (πρωτέωμα) που εκφράζει ένα κύτταρο και η μεταγραφομική αναφέρεται στην ανάλυση του συνόλου των μορίων του μεταγραφικού RNA που παράγει ένα κύτταρο (μεταγράψωμα). Η τεχνολογία μικροδιάταξης του DNA καθορίζει το επίπεδο έκφρασης των γονιδίων και περιλαμβάνει την έκφραση του γενότυπου και την αλληλούχιση του DNA. Η διάταξη της γονιδιακής έκφρασης προσφέρει ταυτόχρονη ανάλυση των επιπέδων έκφρασης του αγγελιαφόρου RNA σε χιλιάδες γονίδια που αφορούν σε καλοήθεις και κακοήθεις όγκους όπως το χηλοειδές και το μελάνωμα. Τα προφίλ έκφρασης κατηγοριοποιούν τους όγκους και παρέχουν πιθανούς θεραπευτικούς στόχους.[29]



**Εικόνα 2.1** Science, Medicine and the future: Bioinformatics, *BMJ* 2002;324:1020 Σχηματικό διάγραμμα που δείχνει τα επιμέρους στάδια της επεξεργασίας των γονιδιακών δεδομένων. Η ανάλυση και η ερμηνεία των βιολογικών δεδομένων μελετά την πληροφορία σε κάθε επίπεδο, ξεκινώντας από το γονιδίωμα προς το πρωτέωμα και το μεταγράψωμα. [29]

Η πρωτεϊνική έρευνα στο χώρο της βιοπληροφορικής στηρίζεται σε βάσεις δεδομένων επισημειωμένων πρωτεϊνών και δεδομένων που έχουν προκύψει κυρίως από τις τεχνικές της ηλεκτροφόρησης δυο διαστάσεων και φασματομετρίας μάζας. Μετά το διαχωρισμό, την ταυτοποίηση και την κατηγοριοποίηση μιας πρωτεΐνης, ο επόμενος στόχος της βιοπληροφορικής είναι η πρόβλεψη της δομής της. Τέλος, οι βιολόγοι που μελετούν τη δομή χρησιμοποιούν επίσης τη βιοπληροφορική για να χειριστούν τον τεράστιο αριθμό περίπλοκων δεδομένων που προκύπτουν από την κρυσταλλογραφία ακτίνων Χ, τον πυρηνικό μαγνητικό συντονισμό αλλά και την παρατήρηση μέσω ηλεκτρονικού μικροσκοπίου με απώτερο στόχο την κατανόηση της χωροταξίας τους.[29][30]

Η κλινική πρωτεομική είναι ένα συνεχώς αναπτυσσόμενο πεδίο αφιερωμένο στην ανακάλυψη, επιβεβαίωση και κατηγοριοποίηση πρωτεϊνών με βάση τη σημασία τους στο πεδίο της υγείας. Γενικά, οι περισσότερες κλινικές πρωτεομικές μελέτες περιλαμβάνουν την τμηματοποίηση ενός συνόλου κλινικών δειγμάτων από μια ομάδα περιπτώσεων καθώς και ένα σύνολο κλινικών δειγμάτων από μια ομάδα ελέγχου και έπειτα τη σύγκριση υγιών και ασθενών κλασμάτων. Οι ροές δεδομένων που προκύπτουν από αυτές τις διαδικασίες στη συνέχεια υφίστανται μια σύνθετη ανάλυση αναγνώρισης μοτίβων, παρέχοντας πληθώρα χρήσιμων πληροφοριών για τις ιδιότητες που διέπουν τις υπό εξέταση πρωτεΐνες. Κατ'αυτόν τον τρόπο, κινούμαστε για να εντοπίσουμε και τους πιθανούς βιοδείκτες, πεπτίδια ή πρωτεΐνες με διακριτή παρουσία σε ορισμένα νοσήματα που συνεισφέρουν στη διαγνωστική ακρίβεια των αντίστοιχων ασθενειών.[31][32]

Οι πρωτεΐνες συμμετέχουν σχεδόν σε κάθε διαδικασία της κυτταρικής ζωής, από την κατάλυση βιοχημικών αντιδράσεων στα πλαίσια του μεταβολισμού μέχρι και την ολοκλήρωση κυτταρικών κυκλωμάτων ως διαμεσολαβητές διαφορετικών μονοπατιών. Η απορρύθμιση της πρωτεϊνικής έκφρασης έχει ως αποτέλεσμα παθολογικές καταστάσεις όπως ο καρκίνος, οι νευροεκφυλιστικές νόσοι και ποικίλες μεταβολικές διαταραχές. Οι πρωτεΐνες συνθέτονται με τη μετάφραση των πληροφοριών που κωδικοποιούνται από ένα αρχικό mRNA μετάγραφο. Συχνά βρίσκονται σε συνεχή δραστηριότητα και κινητικότητα, κάτι το οποίο καθιστά τη μελέτη των μηχανισμών ελέγχου της πρωτεϊνικής ομοιόστασης ιδιαίτερως θελκτική.[31] Πολλές πρωτεΐνες αποτελούν μέρος μεγάλων πολυμερών συμπλεγμάτων τα οποία λειτουργούν κάτω από συγκεκριμένες συνθήκες και συγκεντρώσεις. Οι πρόσφατες εξελίξεις που προκύπτουν από την επιστοίβαξη μεγάλης κλίμακας γενομικών, πρωτεομικών και μεταγραφικών δεδομένων, θέτουν σημαντικές προκλήσεις για τη βιοπληροφορική, προκλήσεις που δεν έχουν ακόμα ξεπεραστεί. Στην περίπτωση της πρωτεομικής έρευνας, αντιξοότητες προκύπτουν σε όλα τα επίπεδα, από την προετοιμασία των δειγμάτων και τη συγκέντρωση των δεδομένων, έως την ενσωμάτωση των ακατέργαστων στοιχείων και τη λειτουργική ερμηνεία ενός μεγάλου όγκου πληροφορίας. Βασιζόμενοι σε μια σειρά σύγχρονων εργαλείων βιοπληροφορικής, μπορούμε να ρίξουμε φως σε αδιέξοδα που προέκυψαν από τις μέχρι πρότινος συμβατικές μεθόδους της κλασικής έρευνας και να συνδυάσουμε αποτελέσματα ερευνών κατά απρόβλεπτο τρόπο, ανοίγοντας νέα μονοπάτια στα ερευνητικά δρώμενα.[32][33]

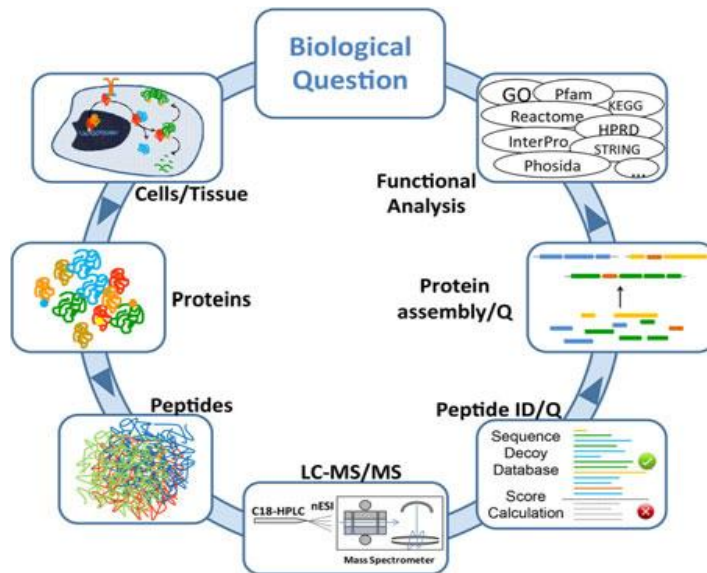
Η πρωτεομική δε θα ήταν εφικτό να γνωρίσει ανάπτυξη χωρίς τα προαναφερθέντα επιτεύγματα και τις ανακαλύψεις της γενομικής η οποία παρείχε την αρχική βάση. Η διαθεσιμότητα των γενομικών και πρωτεομικών δεδομένων και παράλληλα η βελτιστοποίηση των βιοπληροφορικών και βιοχημικών

εργαλείων έχει αυξήσει τις προσδοκίες. Επιπλέον, έχουν εμφανιστεί τεράστια πλεονεκτήματα όπως η βελτιωμένη διάγνωση μέσω της χρήσης πρωτεϊνικών βιοδεικτών, η προστασία απέναντι σε μολυσματικές ασθένειες χρησιμοποιώντας χαμηλού κόστους εμβόλια και πιο ορθολογικό σχεδιασμό φαρμάκων, η βελτίωση στην ποσότητα και την ποιότητα της αγροτικής παραγωγής και η ανάπτυξη τεχνικών που συμβάλλουν στην οπτικοποίηση και την κατανόηση των περίπλοκων μηχανισμών της ασθένειας στο μοριακό επίπεδο. Η πρωτεομική ανάλυση έχει επίσης αποδειχθεί χρήσιμη για την ταυτοποίηση βιοδεικτών των ιστών και του ορού σχετιζόμενους με ανθρώπινες κακοήθειες.[32]

Μια μεγάλη ποικιλία αναλυτικών μεθόδων έχει χρησιμοποιηθεί για τη διαλογή, την ταυτοποίηση και τον χαρακτηρισμό των πρωτεϊνών. Κάθε μια από αυτές τις τεχνολογίες έχει τα προτερήματά της, αλλά όλες έχουν ταυτόχρονα τους περιορισμούς τους. Η ηλεκτροφόρηση δυο διαστάσεων για παράδειγμα είναι μια πολύ καλή μέθοδος για τη ξεδιάλυνση πρωτεϊνών αλλά απαιτεί ένα ιδιαίτερα απαιτητικό πλαίσιο χειρισμού του δείγματος, έχει σχετικά περιορισμένη επαναληψιμότητα και δεν δίνει από μόνη της ξεκάθαρα αποτελέσματα όταν διενεργείται σε μεγάλο όγκο πεπτιδίων. Αν και η τεχνολογία τζελ δυο διαστάσεων υπάρχει εδώ και δεκαετίες και χρησιμοποιείται ευρέως, έχει αρχίσει να αντικαθίσταται από νέες τεχνολογίες που έχουν τελειοποιηθεί ειδικά για την πρωτεομική ανάλυση. Οι διατάξεις πρωτεϊνών-αντισωμάτων έχουν εξαιρετικά ενδιαφέρουσες προοπτικές ως προς την κατεύθυνση της βελτιστοποίησης της πρωτεομικής ανάλυσης, αλλά η λειτουργικότητά τους περιορίζεται από διάφορα τεχνικά προβλήματα. Οι περιορισμοί περιλαμβάνουν την έλλειψη αλγορίθμων υψηλής διεκπαιρωτικότητας για την έκφραση και τον καθαρισμό πρωτεϊνών καθώς και την αδυναμία τους να δημιουργήσουν μεγάλες σειρές καλώς καθορισμένων αντισωμάτων. Χρειάζεται επιπρόσθετη έρευνα και ανάπτυξη προτού οι προοπτικές αυτής της τεχνολογίας να μπορέσουν να είναι πλήρως αξιοποιήσιμες.[31] Το αποτέλεσμα της πρωτεομικής ανάλυσης, ανεξάρτητα της μεθόδου που θα χρησιμοποιηθεί, είναι μια μακρά λίστα δεδομένων. Για να επιτύχουμε την κατανόηση και την ερμηνεία αυτών των δεδομένων και για να δημιουργήσουμε επαληθεύσιμα αποτελέσματα σχετικά τα αντίστοιχα βιολογικά ερωτήματα, χρειάζεται η περαιτέρω ταξινόμηση και φιλτράρισμα τους. Το πρώτο στάδιο για μια λειτουργική ανάλυση μιας μεγάλης λίστας πρωτεϊνών είναι η σύνδεση του ονόματος της πρωτεΐνης με ένα διακριτό χαρακτηριστικό.[33] Ενώ τα ονόματα των γονιδίων έχουν τυποποιηθεί, τα ονόματα των πρωτεϊνών διαφέρουν ανάμεσα σε διαφορετικές βάσεις δεδομένων, ή ακόμη και σε διαφορετικές εκδόσεις της ίδιας βάσης δεδομένων. Παρότι πολλές από τις μεγαλύτερες βάσεις δεδομένων έχουν καθολικοποιηθεί σε μεγάλο βαθμό, αυτό αποτελεί συνήθως μια δύσκολη διαδικασία, που συχνά μπορεί να οδηγήσει σε σημαντική απώλεια πληροφοριών. Ταυτόχρονα, υπάρχουν πολλοί διαδικτυακοί αλγόριθμοι οι οποίοι συνδέουν τα ονόματα των πρωτεϊνών με τα ονόματα των αντίστοιχων γονιδίων τους.[33][34]

Συχνή είναι και η ανάλυση δεδομένων από μεταβολικές οδούς. Μια μεταβολική οδός περιγράφει τη σειρά των χημικών αντιδράσεων σε ένα κύτταρο που οδηγεί σε ένα παρατηρήσιμο βιολογικό αποτέλεσμα. Οι πρωτεΐνες που συμμετέχουν στη χημική αντίδραση καθώς και αυτές που λειτουργούν ως ρυθμιστικοί παράγοντες, συνδυάζονται στις λεγόμενες βάσεις δεδομένων μεταβολικών οδών. Οι πρωτεΐνες ή λίστες γονιδίων μπορούν να μελετηθούν ώστε να βρεθεί σε ποιες μεταβολικές οδούς συμμετέχουν. Αυτή είναι μια πιο ουσιαστική προσέγγιση γιατί μεταφέρει την ερμηνεία των δεδομένων μακριά από τη γονιδιοκεντρική έμφαση προς την ταυτοποίηση των λειτουργικών βιολογικών

διαδικασιών. Υπάρχει μεγάλος αριθμός εργαλείων και βάσεων δεδομένων διαθέσιμα στους ερευνητές για την εξαγωγή στοιχείων σχετιζόμενων με τις μεταβολικές οδούς από απλές πηγές βιολογικών αποτελεσμάτων. Επιπλέον έχουν δημιουργηθεί εξαιρετικά εξειδικευμένες βάσεις δεδομένων για τη μελέτη διαδικασιών των σημάτων μεταγωγής. Αυτού του είδους οι βάσεις, μπορούν να χρησιμοποιηθούν για την ανίχνευση πρωτεϊνών και γονιδίων σχετιζόμενων με καρκίνο, μέσα από ένα δαιδαλώδες δίκτυο πληροφοριών.[33][34]



**Εικόνα 2.2** Bioinformatic analysis of proteomics data, BMC Systems Biology Η ροή των εργασιών στην πρωτεομική. Τα δείγματα που μας ενδιαφέρουν λαμβάνονται και οι πρωτεΐνες διαχωρίζονται μέσω φυγοκέντησης από τα λοιπά κυτταρικά στοιχεία. Τα πεπτίδια που προκύπτουν ύστερα από τη χρήση ειδικών διαλυμάτων(buffers), χωρίζονται μέσω χρωματογραφίας και εισάγονται στο μηχάνημα φασματομετρίας μάζας όπου και καταγράφονται η αναλογία μάζας φόρτισης και τα φάσματα διάσπασης. Τα προκύπτοντα αποτελέσματα αναλύονται σε επόμενο χρόνο από βιοπληροφορικά συστήματα. [33]

Πολλά δεδομένα επίσης προκύπτουν από την ανάλυση των αλληλεπιδράσεων των πρωτεϊνών μεταξύ τους. Η πλειοψηφία αυτών δεν δρουν ως ανεξάρτητες οντότητες. Σχηματίζουν είτε παροδικά, είτε μόνιμα συμπλέγματα με άλλες πρωτεΐνες, τα οποία δρουν ως εφελθήρια σχηματισμού περίπλοκων κυτταρικών δομών. Οι πληροφορίες που αναφέρονται στις πρωτεϊνικές αλληλεπιδράσεις σε συμπλέγματα τοποθετούνται σε βάσεις αλληλεπιδράσεων που σχετίζονται με τις βιολογικές διαδικασίες στις οποίες εμπλέκονται. Οι πρωτεϊνικές αλληλεπιδράσεις συχνά αναπαρίστανται ως μεγάλα δίκτυα τα οποία παρουσιάζουν το βαθμό συνδεσιμότητας μέσω συστημάτων τύπου γράφου με κόμβους και γραμμές.[33]

## 2.2 Cytoscape

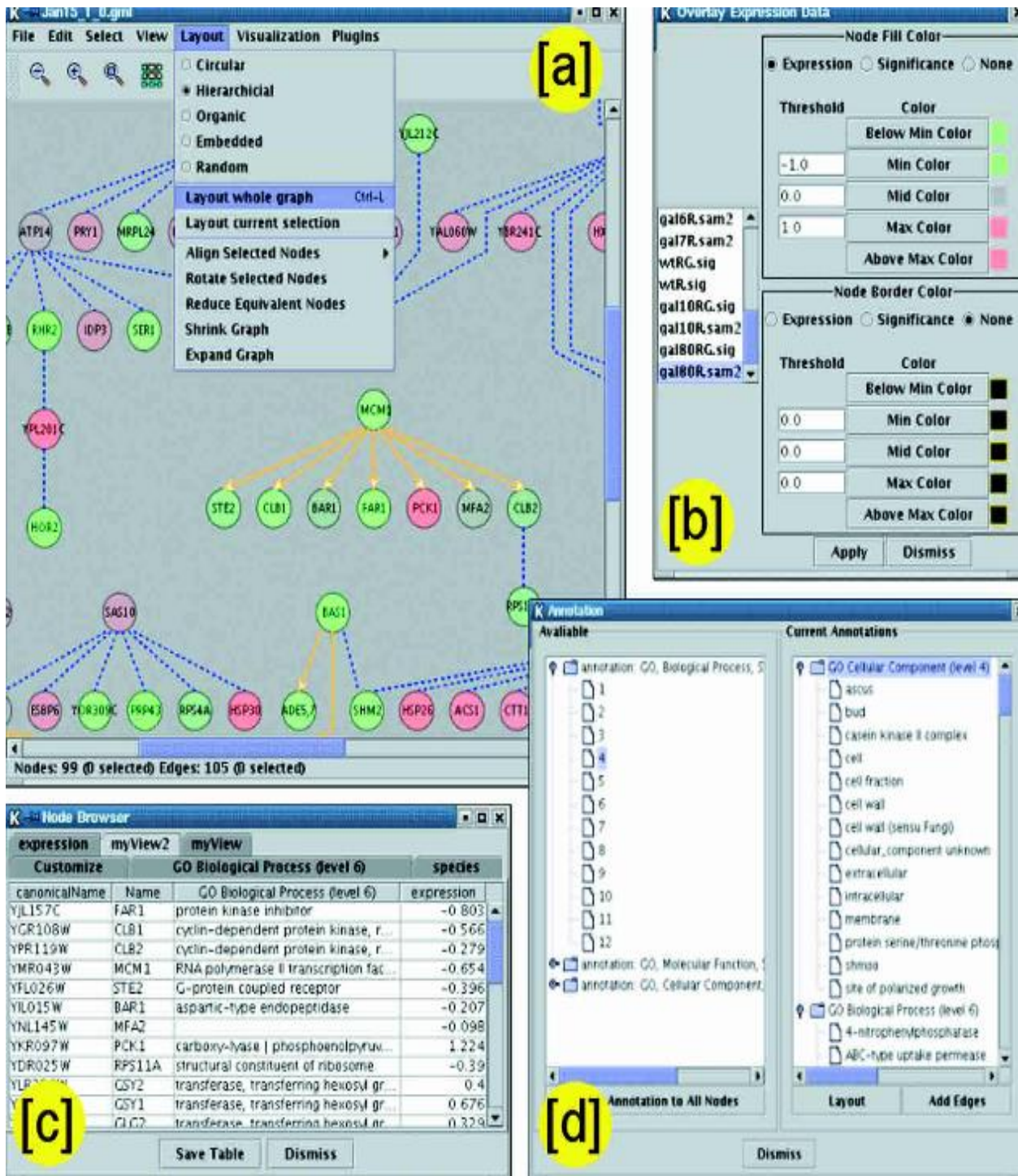
Το Cytoscape είναι μια εφαρμογή βιοπληροφορικής η οποία τα τελευταία χρόνια κυριαρχεί ως το πρώτο εργαλείο οπτικοποίησης βιολογικών δεδομένων μεγάλου όγκου. Ξεκινώντας το 2001 σαν μια εφαρμογή με αρχικό στόχο την ποσοτικοποίηση πειραματικών δεδομένων, έγινε γρήγορα αποδεκτό ως ένα πρόγραμμα με μεγάλη γκάμα εφαρμογών. Τις τελευταίες δεκαετίες, ένα από τα μεγαλύτερα προβλήματα στο χώρο της έρευνας, ήταν η ύπαρξη τεράστιου όγκου δεδομένων βιολογικής σημασίας τα οποία δεν μπορούσαν να σταχυολογηθούν κατά ρυθμό ανάλογο με τον οποίο παράγονται. Επιπρόσθετα, πολλές από αυτές τις πληροφορίες δεν μπορούσαν εύκολα να αξιολογηθούν καθώς εμφανίζονταν ως μεμονομένα στοιχεία χωρίς να είναι δυνατή η εμβάθυνση σε αυτά αφού δεν υπήρχε μέχρι τις αρχές του 2000 ένας τρόπος που να επιτρέπει τη σύγκριση και την οπτικοποίηση τους. Με την ανάπτυξη του διαδικτύου, άνθιση γνώρισαν οι λεγόμενες βάσεις δεδομένων, δηλαδή η δημιουργία δεξαμενών βιολογικής πληροφορίας οι οποίες όμως είχαν τη δυνατότητα όχι μόνο να φιλοξενούν αλλά και να συνδέουν αυτά τα στοιχεία ανάλογα με τα κυτταρικά διαμερίσματα στα οποία εντοπίζονταν, τη βιολογική τους δραστηριότητα αλλά και τη συμμετοχή τους στα διάφορα μεταβολικά μονοπάτια. Γρήγορα λοιπόν σχηματίστηκε ένας κυκλώνας πληροφορίας ο οποίος ήταν αδύνατο να αποτυπωθεί ως προς τη σημασία του χωρίς να οπτικοποιηθεί. Το Cytoscape δημιουργήθηκε με σκοπό να καλύψει αυτό το κενό των ερευνητών στη βιολογία, δίνοντας τους τη δυνατότητα να αποκτήσουν οπτική πληροφορία από τα αποτελέσματα των ερευνών τους αλλά και να τα συγκρίνουν ή να τα συνδυάσουν με στόχο τη διάνοιξη νέων οδών και ιδεών στο χώρο της βιολογικής επιστήμης.[37]

Η λειτουργία του cytoscape στηρίζεται σε ένα σύστημα αρχιτεκτονικής το οποίο δομείται από ορισμένα επιμέρους στοιχεία τα οποία επιτρέπουν την ειδική ερμηνεία ενός σύνθετου ή απλούστερου φάσματος πληροφοριών. Ειδικότερα, το βασικότερο στοιχείο του προγράμματος είναι μια σειρά λογισμικών αλγορίθμων που του δίνουν τη δυνατότητα να χρησιμοποιείται ως γράφος. Συνεπώς, στον πυρήνα λειτουργίας του, πρόκειται για ένα σύστημα όπου συνδέει πληροφορίες μεταξύ τους, με τα αρχικά δεδομένα να παρουσιάζονται ως κόμβοι στο αντίστοιχο γράφημα και τις συνδέσεις μεταξύ τους ως γραμμές. Τόσο οι κόμβοι όσο και οι γραμμές έχουν μια σειρά από δικά τους χαρακτηριστικά. Οι κόμβοι διαφοροποιούνται ως προς το μέγεθος αλλά και ως προς το χρώμα, υποδεικνύοντας το ποσοστό συμμετοχής αλλά και τη βιολογική οικογένεια στην οποία ανήκει το κάθε βιολογικό στοιχείο ενώ οι γραμμές διαφέρουν ως προς το πάχος και ως προς το μέγεθος ανάλογα με την ένταση της σχέσης που παρουσιάζουν. Η δυνατότητα χρησιμοποίησης χρωμάτων και η πληθώρα επιλογών κατά την εισαγωγή των δεδομένων αλλά και οι διάφορες επιλογές στη μέθοδο επεξεργασίας τους, ανοίγει το δρόμο σε δυναμικές μελέτες και συνάμα πρωτότυπες μιας και υπάρχει μικρός περιορισμός ως προς τους διαφορετικούς συνδυασμούς που μπορούμε να δοκιμάσουμε κάθε φορά. Το δεύτερο χαρακτηριστικό του cytoscape είναι η άντληση πληροφορίας από δημοσιευμένες και επιστημονικά τεκμηριωμένες πηγές καθώς και η έννοια των επισημειώσεων που συνοδεύουν αυτές. Ύψιστης σημασίας αποτελεί για κάθε πρόγραμμα βιοπληροφορικής, οι βάσεις δεδομένων του να χαρακτηρίζονται από μεγάλη αξιοπιστία και αμεροληψία. Το cytoscape έχει εδώ και σχεδόν δύο δεκαετίες πλέον καθιερωθεί στη συνείδηση του επιστημονικού τοπίου ως ένα από τα προγράμματα με τη μεγαλύτερη ασφάλεια

αληθείας ως προς τα δεδομένα που χρησιμοποιεί. Επιπλέον, οι διάφορες βιολογικές πληροφορίες φέρουν και σειρά επισημειώσεων, διαφορετικών δηλαδή στοιχείων που αφορούν σε αυτά ανάλογα με το βιολογικό τους ρόλο, την κυτταρική τοποθεσία, τη συμμετοχή σε σηματοδοτικά μονοπάτια, επιτρέποντας στο πρόγραμμα όταν χρησιμοποιεί τα διαφορετικά φίλτρα να αξιοποιεί κάθε φορά μόνο τα επιθυμητά για την εκάστοτε έρευνα χαρακτηριστικά.[35][37]Όπως αναφέρθηκε, η κύρια καινοτομία που εισήγαγε το cytoscape στους κόλπους της βιοπληροφορικής ήταν η δυνατότητα εξαγωγής σύνθετων γραφημάτων. Ωστόσο το τρίτο και καιριότερο στοιχείο του προγράμματος, αποτελεί η χρησιμοποίηση για πρώτη φορά περίπλοκων αλγορίθμων που επιτρέπουν όχι μόνο την απλή παράθεση πληροφορίας μέσω συνδέσεων και χρωματισμών αλλά και την ιεράρχηση των επιμέρους δεδομένων. Αυτό επιτρέπει ακόμα και σε κάποιον που πιθανώς δε γνωρίζει λεπτομέρειες για την έρευνά μας, να μπορεί γρήγορα να πλοηγηθεί στην εικόνα που του παρουσιάζουμε και να εστιάσει στον πηρύνα των ερευνητικών μας ενδιαφερόντων. Αξίζει να σημειωθεί πως αν και υπάρχουν διαφορετικοί αλγόριθμοι ανάλογα με τα φίλτρα που χρησιμοποιούμε κάθε φορά, το cytoscape χαρακτηρίζεται από υψηλή επάρκεια, με την έννοια πως οι περισσότεροι αλγόριθμοι που διαθέτει πλην συγκεκριμένων ειδικών περιπτώσεων, καλύπτουν σχεδόν το πλήρες εύρος των ειδών βιολογικής οπτικοποίησης και αυτό το καθιστά το πρόγραμμα εκλογής για την αναπαράσταση δικτύων βιολογικής πληροφορίας.[35]Επιπλέον, μας προσφέρει τη δυνατότητα όχι μόνο να δούμε αλλά και να επεμβούμε στο παρόν γράφημα. Μπορούμε εν παραδείγματι να αφαιρέσουμε επιμέρους στοιχεία από αυτό, να τα μεταφέρουμε αλλά και να αλλάξουμε το σχήμα τους, ανάλογα με το τι μας ενδιαφέρει κάθε φορά. Η αλλαγή θέσης δε φέρει καμία επίπτωση στην αξία του γραφήματος καθώς οι επιμέρους σχέσεις διατηρούνται ακέραιες, απλώς οι συνδέσεις(γραμμές) αλλάζουν μήκος ανάλογα με το πόση διαφορά παρατηρείται από την αρχική θέση. Πέραν του γραφήματος, το cytoscape παράγει και μια σειρά πινάκων με πληροφορίες για το τι βλέπουμε αλλά και για επιπλέον χαρακτηριστικά όπως της οικογένειας που ανήκουν τα διάφορα γονίδια,πρωτεΐνες,βακτήρια,φάρμακα και οτιδήποτε άλλο χρησιμοποιήσαμε ως πρώτη ύλη. Οι βάσεις δεδομένων από τις οποίες επιτρέπει την άντληση πληροφορίας έχουν ιλιγγιώδες μέγεθος και εκτείνονται από GO(gene ontology), path reactome, biological processes έως drug interactions και protein-protein interaction. Επιπρόσθετα, καθιστά εφικτή ακόμα και την εισαγωγή δικών μας δεδομένων που δε βρίσκονται σε κάποια γνωστή βάση δεδομένων και τη σύγκρισή τους είτε αναμεταξύ τους είτε με ήδη υπάρχουσες. Γίνεται εύκολα λοιπόν αντιληπτό πως ακόμη και στη βασική του μορφή, το cytoscape προσφέρει απύθμενες δυνατότητες επεξεργασίας βιολογικών δεδομένων και πολλές φορές εξαγωγή συμπερασμάτων που ειδάλλως δεν θα ήταν δυνατό να λάβουν χώρα.

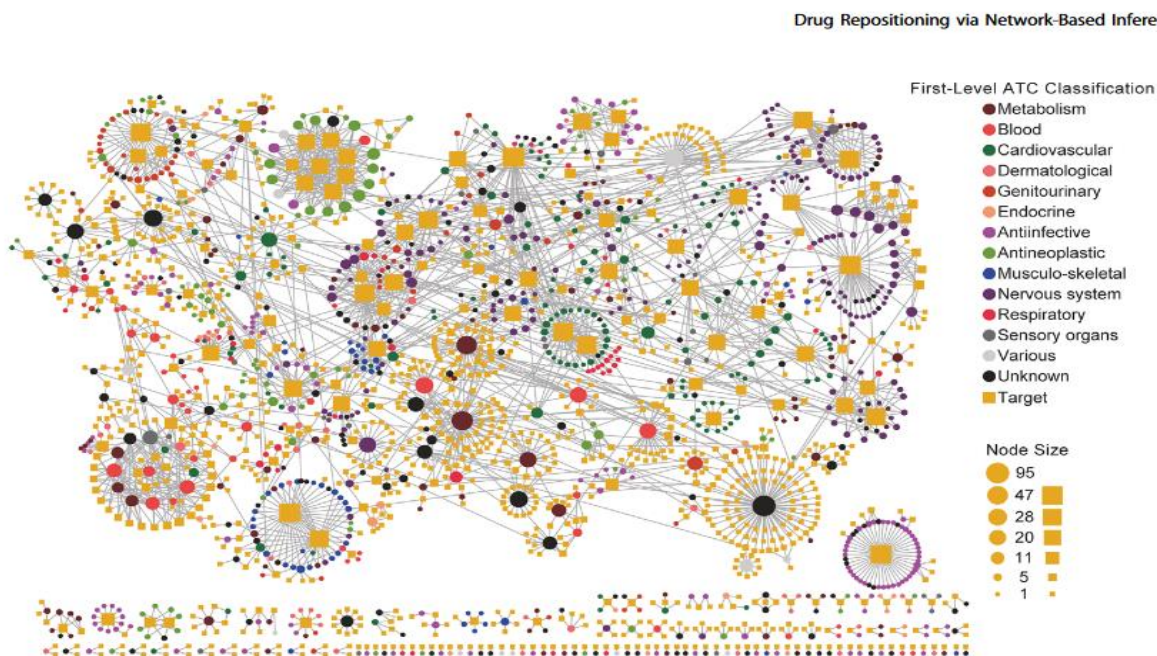
Ακολουθεί μια εικόνα στην οποία περιλαμβάνονται τα κύρια κομμάτια που απαρτίζουν το cytoscape ως λογισμική πλατφόρμα:





**Εικόνα 2.3** Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13(11), 2498–2504. Η εικόνα περιγράφει τέσσερα διαφορετικά στάδια που προηγούνται της ανάλυσης που πρόκειται να διενεργήσουμε: α)παρουσιάζει τις διάφορες επιλογές ιεράρχησης β)τις ρυθμίσεις με τις οποίες η βιολογική πληροφορία επρόκειτο να συνδυαστεί με την οπτικοποίηση γ)Πίνακας με στοιχεία που χαρακτηρίζουν το παραγόμενο γράφημα δ)Ο τρόπος με τον οποίο οι διάφορες επισημειώσεις επισυνάπτονται στα βιολογικά δεδομένα. [35]

Μέχρι τώρα μιλήσαμε για το γενικό πλαίσιο λειτουργίας του cytoscape και ενώ ήδη φαίνεται ως ένα ιδιαίτερα μεγαλεπίβολο εργαλείο βιοπληροφορικής με αναρίθμητες εφαρμογές, δε σταματά εκεί. Η πρωτοποριακή του δυναμική κυρίως έγκειται στο γεγονός πως προσφέρει τη δυνατότητα εγκατάστασης εξωτερικών του cytoscape, plugins. Πρόκειται για σειρά αλγορίθμων οι οποίοι είναι σε θέση να επιτελούν ένα πολύ ειδικό έργο, όταν οι βασική εργαλειοθήκη του cytoscape δεν είναι αρκετή για να το εκτελέσει. Για να γίνει δεκτό ένα plugin θα πρέπει να πληροί μια σειρά χαρακτηριστικών. Αφενός θα πρέπει να παρουσιάζονται τα πλήρη στοιχεία του συγγραφέος αλλά να διασφαλείζεται η εγκυρότητα τους. Εάν αξιολογηθεί το εν λόγω plugin ως να φιλοξενηθεί στο cytoscape, θα πρέπει να επισημειωθούν κάποιες βασικές λειτουργίες του εν συντομία όπως για παράδειγμα γύρω από ποιο βιολογικό αντικείμενο περιστρέφεται και το τι θα μπορέσει να αποφέρει στους χρήστες του. Ο χρήστης του cytoscape είναι σε θέση να ελέγξει τους κατολόγους των plugins που έχουν ενσωματωθεί στο πρόγραμμα και δίχως καμία οικονομική επιβάρυνση, να τα κατεβάσει και να τα χρησιμοποιήσει. Σε κάποιες περιπτώσεις, ορισμένοι συγγραφείς δεν επιτρέπουν την ελεύθερη χρήση των λογισμικών τους παρά μόνο ύστερα από προσωπική άδεια μέσω κάποιου email. Με αυτόν τον τρόπο έχει δημιουργηθεί μια τεράστια βιβλιοθήκη υποπρογραμμάτων που πρακτικά λύνει τα χέρια των εκάστοτε ερευνητών. [38] Ακολουθεί μια εικόνα που περιγράφει το πόσο πολύ μπορεί να περιπλεχθεί η οπτική πληροφορία από τη μελέτη αλληλεπιδράσεων ενός και μόνο φαρμάκου:



**Εικόνα 2.4** Prediction of Drug-Target Interactions and Drug Repositioning via Network-Based Inference. *PLoS Computational Biology*, 8(5), e1002503. Παρουσιάζεται ένα δαιδαλώδες πλέγμα σχέσεων και αλληλοεπικοινωνιών σχετικό με τις δράσεις που έχει ένα συγκεκριμένο φάρμακο ανά βιολογικό σύστημα καθώς και τα δίκτυα των μεταξύ τους συνδέσεων μέσω του προγράμματος cytoscape. Από την εικόνα, είναι πασιφανές πως δύσκολα μπορεί κάποιος να οδηγηθεί σε ασφαλή συμπεράσματα χωρίς ενδελεχέστερη μελέτη. [36]

## 2.3 ClueGO

Το ClueGO πρόκειται για plugin του cytoscape με υψηλή ειδικότητα όσον αφορά σε σύγκριση γονιδίων αλλά και πρωτεϊνών. Ο τεράστιος όγκος πληροφορίας ως προϊόν της βιολογικής έρευνας σε συνδυασμό με το γεγονός πως κάθε δεδομένο αντλείται από μια βάση δεδομένων με συγκεκριμένο ύφος(είτε γενετική είτε πρωτεομική), γέννησε την ανάγκη για plugins με μεγάλη ευαισθησία και παροχή δυνατοτήτων αλληλοσύνδεσης αυτών. Το ClueGO χρησιμοποιεί κυρίως δύο βάσεις δεδομένων, το GO(gene ontology) στην οποία τα γονίδια που περιλαμβάνονται είναι ταξινομημένα με επισημειώσεις σχετικές με την τοπολογία του μεταγραφικού τους προϊόντος, τη βιολογική τους λειτουργία καθώς και τη μοριακή τους υπόσταση και το KEGG όπου περιέχει κυρίως δεδομένα σχετικά με μεταβολικά μονοπάτια. Μία ακόμη βάση δεδομένων είναι και η Biocarta, η οποία επίσης αφορά σε μεταβολικά μονοπάτια. Μέσω του plugin, όλα τα παραπάνω μπορούν να διασταυρωθούν και να παράξουμε σύνθετα αποτελέσματα, κάτι που δε θα ήταν δυνατό αν εξετάζαμε τα επιμέρους στοιχεία μεμονωμένα. Ένα επιπλέον προσόν του ClueGO αποτελεί το γεγονός πως διευκολύνει σε μεγάλο βαθμό την αξιολόγηση της οπτικής πληροφορίας καθώς διαθέτει αλγορίθμους φιλτραρίσματος αρκετά διαφορετικούς από εκείνους του κοινού cytoscape. Η χρήση του προγράμματος είναι αρκετά απλή. Αρχικά τοποθετούμε τη δική μας λίστα η οποία περιέχει τα ονόματα ή τον κωδικό των επιθυμητών γονιδίων και στη συνέχεια επιλέγουμε τι είδους ανάλυση επιθυμούμε να τρέξουμε. Το ClueGO δύναται να παρουσιάσει όλα τα παραπάνω ταυτόχρονα, δηλαδή να κατατάξει τα δεδομένα μας σε ομάδες ανάλογα με την κυτταρική λειτουργία, τοποθεσία και μοριακή δομή αλλά και το καθένα από αυτά μεμονωμένα. Όπως και στο απλό cytoscape και εδώ υπάρχουν κόμβοι(δεδομένα) και ευθείες γραμμές(συνδέσεις)με διαφορετικά μεγέθη και χρώματα ούτως ώστε να υπερθεματίζεται ο εμπλουτισμός της κάθε ομάδας. Επίσης μπορούμε να εισάγουμε και δεδομένα δικής μας προσωπικής έρευνας ούτως ώστε το ClueGO να επιτελέσει τις επιθυμητές συγκρίσεις. Μπορούμε να μετακινήσουμε τους κόμβους προς οπτική διευκόλυνσή μας κατά το δοκούν.

Καθώς τα δίκτυα που θέλουμε να οπτικοποιήσουμε γίνονται ολοένα και πιο δαιδαλώδη, δημιουργείται η ανάγκη για εφαρμογή στατιστικών ελέγχων.[39] Η στατιστική μέθοδος, η οποία μπορεί να επιλεγεί από το χρήστη, ασκεί έλεγχο στις εισαχθείσες πληροφορίες και παρουσιάζει στο τέλος μόνο τις επιτρεπόμενες. Για να γίνει αυτό καλύτερα αντιληπτό, χρησιμοποιώντας τα διάφορα στατιστικά φίλτρα μπορούμε για παράδειγμα να κρατήσουμε μόνο γονίδια τα οποία έχουν συγκεκριμένο ποσοστό πιθανότητας(μεγαλύτερο ή μικρότερο από τη δοθείσα αρχική τιμή που θέσαμε)να συμμετέχουν σε μια συγκεκριμένη βιολογική διαδικασία.Το σύστημα τότε θα παρουσιάσει μόνο συγκεκριμένες τιμές και θα παραβλέψει άλλες.Αυτό προσδίδει μεγάλη σημασία στο πρόγραμμα καθώς μας βοηθά να εστιάζουμε σε περιοχές ενδιαφέροντος και αφαιρεί όλο το «θόρυβο» που θα προσέθετε στην τελική εικόνα μας. Οι συχνότερα χρησιμοποιούμενοι στατιστικοί έλεγχοι είναι οι bonferroni και benjamini hochberg.

Η λειτουργία του ClueGo στηρίζεται στη μαθηματική έννοια της υπεργεωμετρικής κατανομής. Όσον αφορά σε αυτόν το στατιστικό όρο, πρόκειται για κατανομή της στατιστικής η οποία σχετίζεται με διανομή με βάση ένα συγκεκριμένο χαρακτηριστικό, ανάμεσα σε δύο ομάδες, χωρίς αντικατάσταση των μελών των ομάδων. Η υπεργεωμετρική κατανομή διαφέρει από την διωνυμική κατανομή μόνο ως προς την έλλειψη αντικαταστάσεων. Έτσι, συχνά χρησιμοποιείται σε τυχαία δειγματοληψία για στατι-

στικό ποιοτικό έλεγχο. Ένα απλό καθημερινό παράδειγμα θα ήταν η τυχαία επιλογή μελών για μια ομάδα από έναν πληθυσμό κοριτσιών και αγοριών. Ο γενικός της τύπος είναι ο εξής:

$$P(X = x) = \frac{\binom{k}{x} \binom{N - k}{n - x}}{\binom{N}{n}},$$

**Εικόνα 2.5** Παρουσιάζεται ο γενικός τύπος της υπεργεωμετρικής κατανομής όπου  $n$  ο αριθμός των δειγμάτων που επιλέχθηκαν,  $N$  ο αριθμός του συνόλου των δειγμάτων,  $x$  ο αριθμός των δειγμάτων με ένα επιθυμητό στοιχείο  $A$ ,  $k$  το σύνολο των δειγμάτων που περιέχουν το χαρακτηριστικό στοιχείο  $A$ ,  $n-x$  ο αριθμός των δειγμάτων με ένα επιθυμητό στοιχείο  $B$  και  $n-k$  το σύνολο των δειγμάτων με το χαρακτηριστικό στοιχείο  $B$ . Η παραπάνω σχέση δηλαδή εκφράζει την πιθανότητα να επιλέξουμε αναμέσα σε ένα δείγμα με δύο πληθυσμούς που περιέχουν δύο δακριτά χαρακτηριστικά ο καθένας, ένα συγκεκριμένο αριθμό δειγμάτων που εκφράζουν μόνο το ένα από τα δύο χαρακτηριστικά, που προέρχονται δηλαδή μόνο από την ίδια ομάδα.**[41]**

Για να ξεκινήσει το ClueGO χρειάζονται τέσσερα βήματα. Αρχικά θέτουμε το είδος των οργανισμών που μας ενδιαφέρει να μελετήσουμε. Μπορούμε σε αυτό το βήμα να επιλέξουμε παραπάνω από ένα στην περίπτωση που η έρευνα μας αφορά σε περισσότερα του ενός είδους. Εδώ γίνεται επίσης η επιλογή του μεγέθους των κόμβων καθώς επίσης επιλέγεται και ο χρωματισμός. Όλα τα παραπάνω μπορούν να τροποποιηθούν και μετά το πέρας του σχεδιασμού του γραφήματος. Ακολούθως, μπορούμε να επιλέξουμε το είδος της ομαδοποίησης που επιθυμούμε. Οι κύριες επιλογές είναι η ομαδοποίηση κατά ομάδα, η ομαδοποίηση κατά σμήνος και η ομαδοποίηση κατά σημασία. Αυτό το στάδιο αφορά κυρίως στη μέθοδο οπτικοποίησης αν και η πιο συνηθισμένη επιλογή είναι η πρώτη (default). Σε επόμενο χρόνο, επιλέγουμε το είδος της αναζήτησης που επιθυμούμε. Εδώ περιλαμβάνονται όλες οι διαθέσιμες βάσεις δεδομένων καθώς επίσης και υποτιμήματα αυτών. Για παράδειγμα μπορεί να επιθυμούμε να χρησιμοποιήσουμε μόνο τη βάση δεδομένων του GO (gene ontology) αλλά μόνο όσον αφορά στο κομμάτι που αναφέρεται στην κυτταρική τοποθεσία και όχι στα υπόλοιπα ή σε δύο κατηγορίες και όχι σε άλλες. Μπορούμε φυσικά να επιλέξουμε και όλες τις βάσεις δεδομένων ταυτόχρονα και το σύστημα τότε θα φιλτράρει τα επικαλυπτόμενα δεδομένα. Ακολούθως, επιλέγουμε το είδος του στατιστικού ελέγχου που θέλουμε να εφαρμοστεί και επίσης ρυθμίζουμε το όριο πάνω από το οποίο τα αποτελέσματα θα γίνουν δεκτά και θα παρουσιαστούν στο τελικό γράφημα. Τέλος τρέχουμε την ανάλυση. Εκτός του γραφήματος, το ClueGO δίνει τη δυνατότητα εξαγωγής μιας σειράς αρχείων excel όπου περιέχουν πληροφορίες για τα γονίδια τα οποία εισαγάγαμε όπως για παράδειγμα τι ποσοστό αυτών ανήκει σε μια συγκεκριμένη ομάδα ή σε περισσότερες από μια ταυτόχρονα καθώς επίσης και από ποια βάση αντλήθηκαν τα δεδομένα.**[40]**

The screenshot displays the ClueGO software interface. On the left, the Control Panel includes options for ClueGO Function (ClueGO+CluePedia), ClueGO Settings (Ontologies/Pathways), ClueGO Functional Analysis, and ClueGO Settings (ClueGO Settings). The main window shows a network of terms/pathways with labels such as 'Antigen processing and presentation', 'Rheumatoid arthritis', 'positive regulation of response to external stimulus', and 'cell chemotaxis'. Below the network is a Table Panel with a 'Results for Analysis4' table. The table has columns for GO term, Ontology Source, GO Levels, % Associ., Nr. Genes, Term PValue, Term PValue Correc., and Associated Genes Found. The table lists various GO terms and their associated genes, with some terms highlighted in green.

**Network of terms/pathways**

**Result table: overrepresentation of terms/pathways**

GO ID	GO Term	Ontology Source	GO Levels	% Associ.	Nr. Genes	Term PValue	Term PValue Correc.	Associated Genes Found
2695	negative regulation of leukocyte activation	GO_BiologicalPr...	[3, 4, 5]	8.197	10	8.6 E-6	4.4 E-4	BTN2A2, CD274, CNR1, GPER1, I...
50670	regulation of lymphocyte proliferation	GO_BiologicalPr...	[5, 6, 7]	7.558	13	9.6 E-7	5.5 E-5	BTN2A2, CD1D, CD274, IDO1, IFN...
50864	regulation of B cell activation	GO_BiologicalPr...	[5, 6, 7]	7.216	7	4.3 E-4	1.3 E-2	BTX, IFNG, IL7, MEF2C, MINDA, TB...
31295	T cell costimulation	GO_BiologicalPr...	[3, 6, 7]	7.5	6	9.1 E-4	2.5 E-2	CD274, ICOS, PDCD1, PDCD1LG2,
50672	negative regulation of lymphocyte proliferation	GO_BiologicalPr...	[5, 6, 7]	10.714	6	1.3 E-4	5.0 E-3	BTN2A2, CD274, IDO1, LILRB2, M...
42129	regulation of T cell proliferation	GO_BiologicalPr...	[6, 7, 8]	8.264	10	8.0 E-6	4.2 E-4	BTN2A2, CD1D, CD274, IDO1, IFN...
42102	positive regulation of T cell proliferation	GO_BiologicalPr...	[6, 7, 8]	9.091	7	1.0 E-4	4.4 E-3	CD1D, CD274, IFNG, LILRB2, PDC...
560	TGF Beta Signaling Pathway	WikiPathways_2...	[1]	7.407	4	6.9 E-3	9.7 E-2	BTX, BTN2A2, CCL3, CD180, CD27...
32103	positive regulation of response to external stimulus	GO_BiologicalPr...	[3, 4, 5]	7.653	15	1.1 E-7	6.8 E-6	CCL3, CD180, CNR1, CXCL1, CXC...
50921	positive regulation of chemotaxis	GO_BiologicalPr...	[3, 4, 5]	6.122	6	2.5 E-3	5.4 E-2	CCL3, CXCL10, CXCR2, IL8, PDGFR...
2687	positive regulation of leukocyte migration	GO_BiologicalPr...	[3, 4, 5]	9.756	8	1.9 E-5	9.8 E-4	CCL3, CXCL10, CXCR2, KAM1, IL8,
50726	positive regulation of inflammatory response	GO_BiologicalPr...	[4, 5, 6]	10.976	9	2.2 E-6	1.2 E-4	CCL3, CNR1, CXCL1, IDO1, PLA2G...
2712	regulation of B cell mediated immunity	GO_BiologicalPr...	[5, 7]	13.514	5	1.5 E-4	5.8 E-3	IFNG, TBX21, TNF, TNFSF41
50707	regulation of cytokine secretion	GO_BiologicalPr...	[4, 5, 6]	8.247	8	6.6 E-5	2.9 E-3	BTN2A2, CCL3, CD274, F2R, IFNG,
50715	positive regulation of cytokine secretion	GO_BiologicalPr...	[4, 5, 6]	7.812	5	2.0 E-3	4.8 E-2	CCL3, CD274, F2R, IFNG, TNF
4612	Antigen processing and presentation	KEGG_26.02.20...	[1]	9.756	8	1.9 E-5	9.8 E-4	BTX, BTN2A2, CCL3, CD180, CD1...
4650	Natural killer cell mediated cytotoxicity	KEGG_26.02.20...	[1]	6.667	9	1.2 E-4	5.0 E-3	ICAM1, IFNG, KLRC1, KLRC2, KLR...
5321	Inflammatory bowel disease (IBD)	KEGG_26.02.20...	[1]	7.246	5	2.8 E-3	5.6 E-2	HLA-DPMR, IFNG, STAT1, TBX21, TN...

**Εικόνα 2.6** <http://proteomicsnews.blogspot.com/2018/08/cluego-another-promising-downstream.html>  
 Παρουσιάζονται τα επιμέρους βήματα των επιλογών του ClueGO. Στο κάτω μέρος παρατίθενται οι λεπτομέρειες του γραφήματος οι οποίες μπορούν να εκδοθούν σε excel. [42]

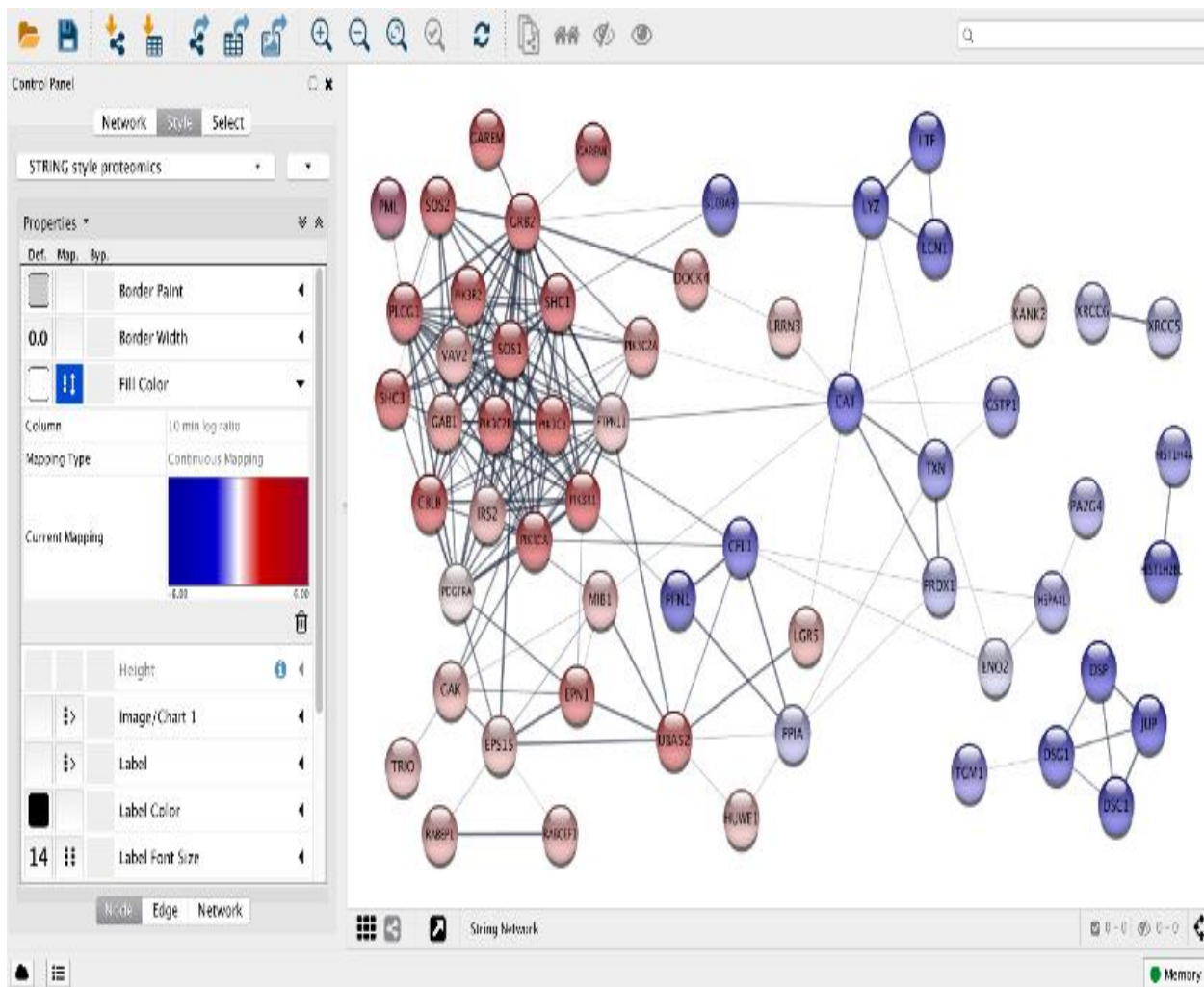
## 2.4 STRING

Τόσο τα γονίδια όσο και οι πρωτεΐνες παρουσιάζουν μια τεράστια ποικιλία ενώ ακόμη μεγαλύτερη είναι η συσσώρευση ολόένα και περισσότερης πληροφορίας από τα διάφορα βιολογικά εργαστήρια. Είναι εξαιρετικά σημαντικό λοιπόν να υπάρχουν προγράμματα βιοπληροφορικής που να μπορούν να διαχειριστούν και να διυλίσουν αυτά τα δεδομένα, προσφέροντας μια πιο χειροπιαστή εικόνα σε σχέση με το ποιες διαδικασίες εμπλέκονται αλλά και ποιος ο ρόλος τους σε αυτές. Μια τέτοια εφαρμογή είναι το STRING, ένα λογισμικό βιοπληροφορικής το οποίο προσπαθεί να συνδέσει γονίδια και

πρωτεΐνες σε δίκτυα που παρουσιάζουν αξιοποιήσιμη βιολογική πληροφορία. Υπάρχουν ωστόσο πάμπολλα άλλα προγράμματα που επιτελούν παρόμοιο έργο. Το STRING όμως διαφοροποιείται από τα υπόλοιπα λόγω του ότι εστιάζει περισσότερο σε θέματα που αφορούν στην πρωτεϊνική ή γονιδιακή αλληλεπίδραση και στην παροχή δικτύων πιθανοτήτων όταν τα βιολογικά δεδομένα δεν επαρκούν για την εξαγωγή απόλυτα σίγουρων συμπερασμάτων. Πρόκειται δηλαδή για πρόγραμμα το οποίο πραγματοποιεί ποιοτικό έλεγχο των δεδομένων που αντλεί από τις επιμέρους βάσεις δεδομένων και αποδίδει ένα συγκεκριμένο βαθμό(score) που αντανακλά το πόσο πιθανές είναι οι αλληλεπιδράσεις που προτείνει στα γραφήματα που σχεδιάζει.[43] Η εισαγωγή των δεδομένων γίνεται από ήδη υπάρχουσες βάσεις όπως και στα περισσότερα προγράμματα οπτικής ποσοτικοποίησης και στηρίζεται κυρίως στην αναζήτηση συγκεκριμένων σημάτων στο γονιδίωμα τα οποία μπορεί να ταυτοποιηθεί ότι ανήκουν σε ένα συγκεκριμένο μεταβολομικό μονοπάτι, σε εντοπισμό αλληλεπιδράσεων τύπου πρωτεΐνης-πρωτεΐνης και σε αναζήτηση λέξεων κλειδιά. Όσον αφορά στον πρώτο όρο, το πρόγραμμα διενεργεί έλεγχο εύρεσης δημοσιευμένων στοιχείων που να περιλαμβάνουν το γονίδιο που μας ενδιαφέρει ως υπεύθυνο για την παραγωγή ενός συγκεκριμένου μορίου. Κατά αυτόν τον τρόπο, δημιουργεί συσχετίσεις και πιθανά δένδρα τα οποία βαθμολογεί ανάλογα με την ένταση της επικάλυψης των δεδομένων από τις βάσεις που προήλθαν. Στις πρωτεϊνικές αλληλεπιδράσεις η λογική είναι παρόμοια. Και στις δύο περιπτώσεις, όταν τα δεδομένα δεν επαρκούν για να είναι σίγουρο το πρόγραμμα ότι ισχύουν αδιαμφισβήτητα, χρησιμοποιεί αλγορίθμους πιθανοτήτων συγκρίνοντας τα αποτελέσματα με άλλα παρεμφερή και καταλήγει σε συμπεράσματα για το αν αυτές οι διασταυρώσεις φέρουν αρκετά υψηλούς βαθμούς αποτελεσμάτων ούτως ώστε να θεωρηθούν πολύ πιθανές. Αυτό είναι ιδιαίτερα σημαντικό καθώς σε αντίθετη περίπτωση θα είχαμε απώλεια τεράστιου όγκου πληροφορίας λόγω του ότι βιολογικά μια συγκεκριμένη σύνδεση δεν έχει ταυτοποιηθεί ή μελετηθεί, ενώ ταυτόχρονα επιτρέπει την παράκαμψη εμποδίων που προκύπτουν από τις βιολογικές μελέτες. Το τρίτο στοιχείο του STRING, η αξιοποίηση των λέξεων κλειδιά, είναι μια τακτική που ακολουθείται από όλα τα προγράμματα βιοπληροφορικής. Τα διάφορα δεδομένα είναι κατηγοριοποιημένα με τέτοιο τρόπο ούτως ώστε να φέρουν επισημειώσεις που αφορούν στα χαρακτηριστικά τους όπως για παράδειγμα το ρόλο, την κυτταρική τους τοποθεσία, τη συσχέτιση τους ως σηματοδοτικά μόρια κ.α. Μια πρωτεΐνη για παράδειγμα μπορεί να φέρει επισημειώσεις σχετικές με το RNA από το οποίο προήλθε, τις πρωτεΐνες με τις οποίες συσσωματώνεται καθώς επίσης και με τις μεταμεταφραστικές τροποποιήσεις της. Αυτές οι επισημειώσεις επιταχύνουν σε μεγάλο βαθμό την εύρεση και διασταύρωση πληροφοριών ενώ μπορούν επίσης να συμβάλουν στην αύξηση του score πιθανότητας που υπολογίζει το STRING. Σε κάθε περίπτωση, τα αποτελέσματα ύστερα από την εκτέλεση του προγράμματος παρουσιάζονται τόσο με τα ανάλογα δίκτυα όσο και με το συνοδευόν κείμενο που τα δικαιολογεί. Από τα scores και από τους περιορισμούς που επιτάσσει η προσωπική μας έρευνα, μπορούμε να αποφασίσουμε για το αν θα δεχτούμε ορισμένες από τις συνδέσεις ή θα τις απορρίψουμε. Μπορούμε εξ'αρχής βέβαια να ορίσουμε τα όρια των scores που θέλουμε να παρουσιάζονται.[43]

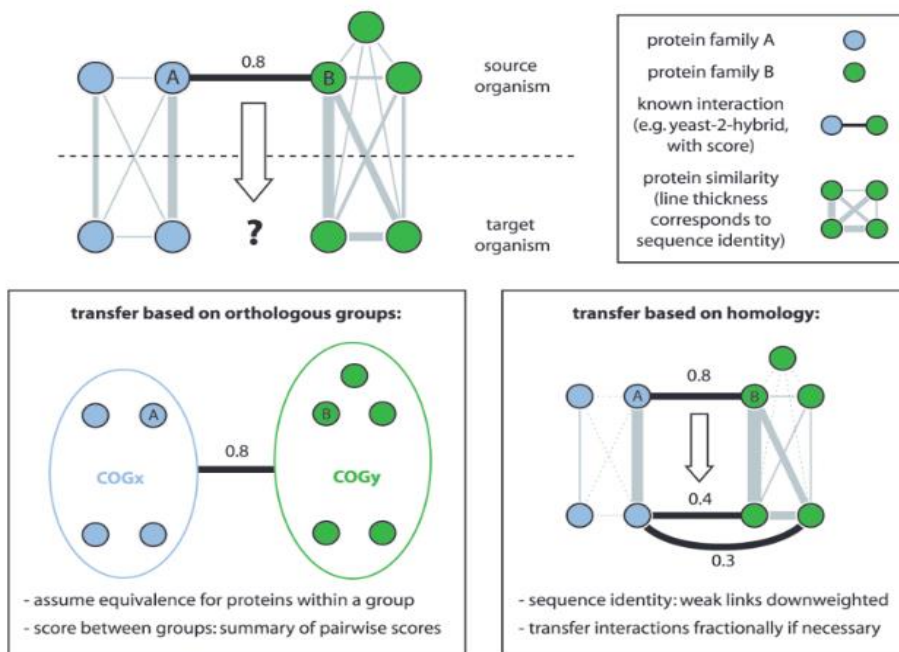
Ένα επιπλέον προσόν του STRING είναι και η δυνατότητα να υπολογίζει στις πιθανές συνδέσεις σε πρωτεΐνες ή γονίδια τα οποία απέχουν μεταξύ τους και δε βρίσκονται σε άμεση αλυσιδωτή σχέση. Είναι αλήθεια ότι ιδίως οι πρωτεϊνικές αλληλεπιδράσεις πολλές φορές δε λαμβάνουν χώρα άμεσα αλλά με έμμεσο τρόπο μέσω της ρύθμισης στοιχείων της μιας πρωτεΐνης από την άλλη. Ένα τέτοιο παρά-

δειγμα αποτελεί ένα κλειστό σύστημα θετικής ανατροφοδότησης όπου η μία πρωτεΐνη ρυθμίζει την παραγωγή μιας άλλης μέσω άμεσης αλληλεπίδρασης με στοιχεία που καθορίζουν τη μεταγραφική δραστηριότητα της δεύτερης πρωτεΐνης. Το STRING δύναται να συγκρίνει τα βιολογικά δεδομένα κατά τέτοιο τρόπο που ακόμα και αν δύο πρωτεΐνες δεν εμφανίζουν άμεση συσχέτιση, να αναλύονται ως προς τη βιολογική τους δραστηριότητα και να προτείνονται πιθανές οδοί μέχρι να φτάσουμε από τη μια πρωτεΐνη στην άλλη με βάση το αν ανήκουν σε κοινά μεταβολικά μονοπάτια. Και σε αυτήν την περίπτωση αποδίδονται τα αντίστοιχα scores πιθανοτήτων για να έχουμε εικόνα του μεγέθους απόκλισης από την πραγματικότητα.[43][44]



**Εικόνα 2.7** *The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research, 45(D1), D362–D368. Μια τυπική εικόνα ενός διαγράμματος τύπου STRING. Διακρίνονται οι κλασικές γραμμές και οι κλασικοί κόμβοι που χαρακτηρίζουν όλα τα δίκτυα με ιδιότητες γράφου. Έχουμε επίσης τη δυνατότητα να μεταποιήσουμε την κλασική(default) μορφή του διαγράμματος, αλλάζοντας το χρώμα των κόμβων και των γραμμών, το μέγεθος, το σχήμα καθώς και το συνοδεύον τους κείμενο. [43]*

Ένα ακόμα σημαντικό στοιχείο του STRING, είναι η εφαρμογή του συστήματος τύπου GOG. Η GOG είναι μια αναζήτηση σε βάσεις δεδομένων κατά τη σύγκριση διαφορετικών ειδών οργανισμών, με στόχο να διαπιστωθεί είτε το κατά πόσο διατηρημένα μεταξύ των ειδών είναι ορισμένα λειτουργικά μονοπάτια, είτε διότι γνωρίζουμε τη δράση της πρωτεΐνης στο ένα μόνο από τα δύο είδη και θέλουμε να την ταυτοποιήσουμε στο άλλο. Το σύστημα GOG, αφού ανατρέξει στα αντίστοιχα ερευνητικά δεδομένα, υπολογίζει το score μεταξύ της σχέσης που συναντάται σε ένα είδος και στη συνέχεια προσπαθεί να το αποδώσει προσεγγιστικά στο δεύτερο. Για να το πετύχει αυτό, λαμβάνει υπόψιν και την περιρρέουσα ατμόσφαιρα των δύο πρωτεϊνών, δηλαδή τι προηγείται και τι ακολουθεί μετά από τη σύνδεση μεταξύ τους. Επιπλέον, υπολογίζει και τις πιθανές συνδέσεις των δύο πρωτεϊνών με άλλες πρωτεΐνες του μεταβολικού μονοπατιού και καταληκτικά αποφασίζει για το αν τα αποτελέσματα ευσταθούν και αξίζει να συμπεριληφθούν στο τελικό διάγραμμα. Τόσο το πρώτο όσο και το δεύτερο score συμψηφίζονται πριν το πρόγραμμα καταλήξει στο τελικό αποτέλεσμα.[44]



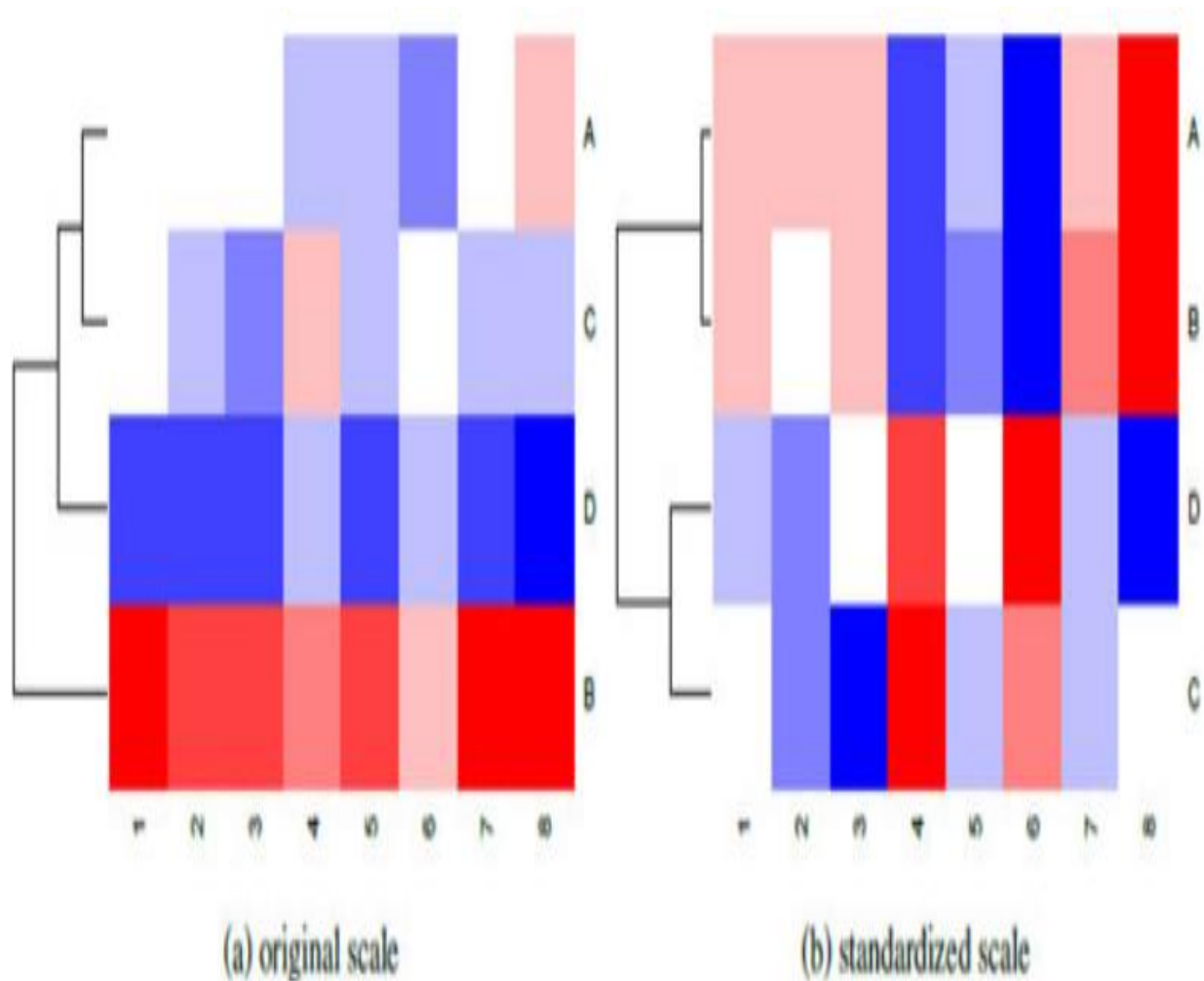
**Εικόνα 2.8** STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Research*, 33(Database issue), D433–D437. Παρουσιάζεται το σύστημα GOG, πως υπολογίζει τα scores ανάμεσα στις πρωτεΐνες ενδιαφέροντος και πως τα μεταφέρει από το ένα είδος στο άλλο. Για να επιτελέσει τους εκάστοτε ελέγχους, αντλεί πληροφορίες από τις βάσεις δεδομένων, τις συνδυάζει με κατάλληλους αλγορίθμους και οπουδήποτε υπάρχει αμφιβολία, θέτει ένα score για να αξιολογήσει την πιθανότητα η απεικονιζόμενη σχέση να ευσταθεί.[45]

## 2.4 Heatmap

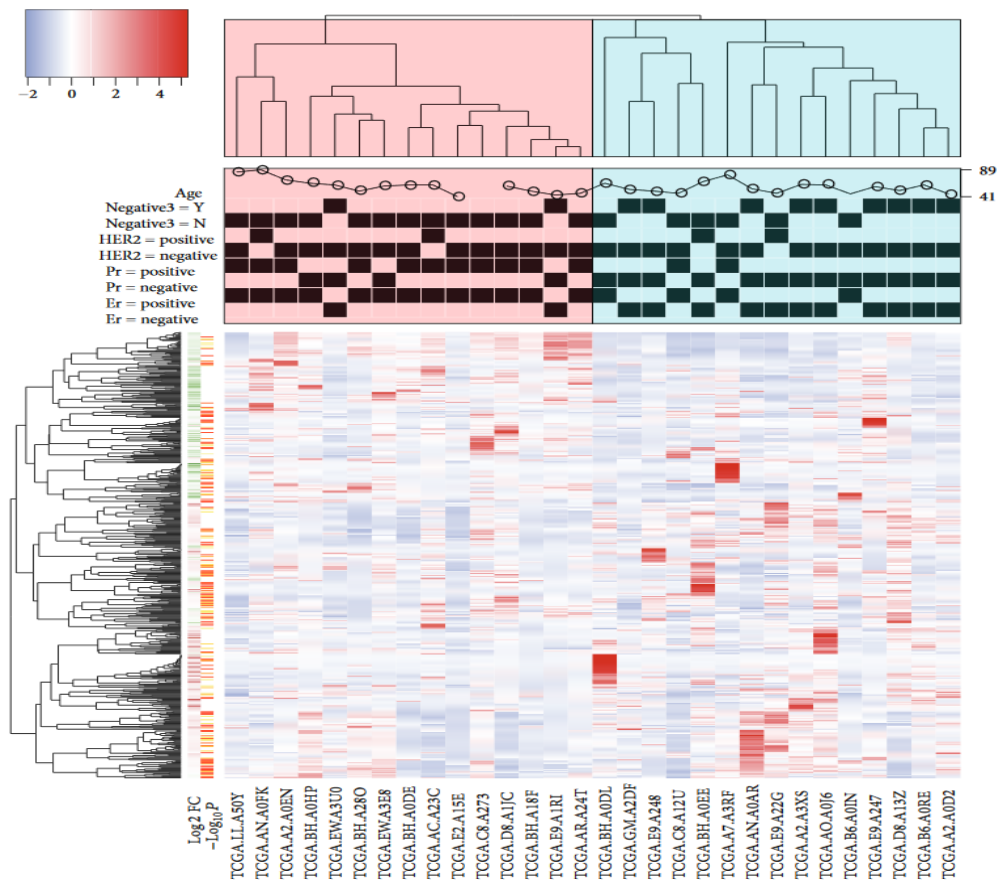
Ανάμεσα στα πλέον ευρέως χρησιμοποιούμενα εργαλεία βιοπληροφορικής, ανήκει και το Heatmap diagram. Πρόκειται για πολύ σημαντικό πολυεργαλείο το οποίο βρίσκει εφαρμογή σε πολλά επιστημονικά πεδία και όχι μόνο στο βιολογικό χώρο. Ως στόχο έχει την παρουσίαση ποσοτικής πληροφορίας



και τη διευκόλυνση της σύγκρισης δεδομένων προερχόμενων από δεξαμενές μεγάλου όγκου, οι οποίες στην αρχική τους μορφή είναι τελείως ανεπεξέργαστες και δεν μπορούν να οδηγήσουν σε ασφαλή συμπεράσματα για το τι πραγματικά συμβαίνει στο δείγμα μας. Το Heatmap χρησιμοποιεί διάφορους αλγορίθμους ομαδοποίησης και συμπαρουσίασης των μεταβλητών που εμείς ορίζουμε και στη συνέχεια αποδίδει χρώματα ούτως ώστε να υπερθεματίσει με ευδιάκριτο τρόπο τις επιμέρους ομάδες που σχηματίζονται. Τα είδη διαγραμμάτων τύπου heatmap διακρίνονται κυρίως, χωρίς όμως να περιορίζονται μόνο σε αυτά, σε δύο κύριες κατηγορίες, στα heatmaps με μορφή διαγράμματος απλής διάταξης μεταβλητών και στα heatmaps με μορφή δενδρογράμματος. Η πρώτη κατηγορία πρόκειται για διαγράμματα στα οποία στον έναν άξονα τοποθετείται το σύνολο των δειγμάτων στα οποία θέλουμε να διενεργήσουμε τη συγκριτική μελέτη και στο δεύτερο το επιθυμητό χαρακτηριστικό με βάση το οποίο θέλουμε να τα διακρίνουμε. Η δεύτερη περίπτωση ομοιάζει με την πρώτη ως προς τις θεμελιώδεις αυτές αρχές, με τη διαφορά όμως του ότι προσφέρει τη δυνατότητα σχεδιασμού δενδρογράμματος ανάλογα με τη συγγένεια που εμφανίζουν τα εικονιζόμενα στοιχεία του δείγματος ως προς το εκάστοτε χαρακτηριστικό.[46][48] Στη συνέχεια πραγματοποιείται η διαδικασία της ομαδοποίησης. Ο κύριος αλγόριθμος που χρησιμοποιείται είναι εκείνος με το όνομα agglomerative hierarchical clustering algorithm, δηλαδή ένας αλγόριθμος ομαδοποίησης δεδομένων με γνώμονα τη συμπαρουσίαση δεδομένων χωρισμένων κατά ομάδες με βάση το βαθμό συσχέτισής τους με ένα συγκεκριμένο χαρακτηριστικό. Ο αλγόριθμος μπορεί να εφαρμοστεί σε οποιονδήποτε από τους δύο άξονες μας ενδιαφέρει. Γίνεται πασιφανές πως εάν εν παραδείγματι ο ένας άξονας αντιπροσωπεύει μια λίστα πρωτεϊνών και ο άλλος τα επιθυμητά στοιχεία με βάση τα οποία θέλουμε να τις διαφοροποιήσουμε, ο ρόλος των heatmaps είναι να φέρουν κοντά τις πρωτεΐνες με παρόμοια scores για το εκάστοτε χαρακτηριστικό και να δημιουργήσει μικρές ή μεγάλες γεωγραφικές αποικίες στο διάγραμμα με παράλληλη υπέρθεση του αντίστοιχου χρώματος, δηλωτικού αυτού του χαρακτηριστικού. Αυτή η ιδιότητα (scaling) κατέχει εξέχουσα σημασία, ιδίως όταν ο αριθμός των δειγμάτων είναι τεράστιος και τα χαρακτηριστικά που θέλουμε να ξεχωρίσουμε είναι περισσότερα του ενός. Από την άλλη, μέσω αυτής της τεχνικής μπορούν να δημιουργηθούν πολύπλοκα διαγράμματα και πολλές ομάδες να μην είναι πάντα διακρίτες ως κάτι το ξεχωριστό. Σε αυτήν την περίπτωση, έχουμε τη δυνατότητα να ορίσουμε εκ προοιμίου τα όρια των scores με βάση τα οποία θέλουμε να τρέξουμε την εφαρμογή. Κατ'αυτόν τον τρόπο, μπορούμε να αυξομειώσουμε την ευαισθησία της έρευνάς μας και να κατασκευάσουμε είτε περισσότερο είτε λιγότερο λεπτομερή παρουσίαση.[45][46] Όσον αφορά στο δενδρόγραμμα, η δημιουργία του στηρίζεται στο βαθμό της ομαδοποίησης που επιτελέσαμε και ο σχεδιασμός των γραμμών είναι εξαρτώμενος από τον αριθμό των διαφορετικών ομάδων που προέκυψαν μετά το scaling. Όσο τα scores αρχίζουν να αποκλίνουν από τις αρχικές τιμές τους, (ως αρχικές ορίζουμε τα scores με τιμή κοντά ή ακριβώς ίση με εκείνη που έχει αποδοθεί στο εκάστοτε χαρακτηριστικό) αρχίζουν να σχηματίζονται ολοένα και περισσότερες διακλαδώσεις. Καθώς αυτές συνεχίζουν να αυξάνουν σε ύψος, (εάν κοιτάμε τα αρχικά επίπεδα ομαδοποίησης ως τη βάση) αρχίζουμε να απομακρυνόμαστε από αυτές και να τείνουμε να προσεγγίσουμε τη γραμμή της μηδενικής πιθανότητας, η οποία έχει το μεγαλύτερο ύψος όλων. Ακολουθούν δύο εικόνες χαρακτηριστικές των βασικών ειδών τύπου heatmap:



**Εικόνα 2.9** A tutorial in displaying mass spectrometry-based proteomic data using heat maps. *BMC Bioinformatics*, 13(Suppl 16), S10. Στο διάγραμμα α παρουσιάζονται τα δεδομένα χωρίς να έχει γίνει ομαδαποίηση ενώ στο β έχει προηγηθεί το *scaling*. Μπορούμε να διακρίνουμε πως στη δεύτερη περίπτωση είναι πολύ πιο ευκρινείς οι διαφορές στην απόκλιση σε σχέση με το πρώτη. Η ένταση των χρωμάτων αντανακλά την ένταση των χαρακτηριστικών. Στο αριστερό σκέλος και των δύο εικόνων υπάρχει ένα δενδρόγραμμα που αποτυπώνει τις ομάδες συνάφειας στις οποίες ανήκουν στοιχεία από το δείγμα μας. [47]



**Εικόνα 2.10** *Advanced Heat Map and Clustering Analysis Using Heatmap3. BioMed Research International, 2014, 1–6.* Χαρακτηριστική εικόνα ενός δενδρογράμματος από γονιδιακή ανάλυση τύπου Heatmap3. Παρατηρούμε τα πυκνοσχεδιασμένα clusterings που αντικατοπτρίζουν την ανακατατάξη των δεδομένων ύστερα από το scaling. Στο Heatmap 3 παρουσιάζονται παράλληλα και οι φαινότυποι των αντίστοιχων χαρακτηριστικών. Heatmap 3 διαγράμματα παρουσιάζουν εξαιρετικό ενδιαφέρον ως προς τις δυνατότες που προσφέρουν σε περιπτώσεις που το μέγεθος του δείγματός μας είναι αρκετά μεγάλο και επιθυμούμε μια λεπτομερή κατηγοριοποίηση του. [48]

Προκειμένου να δημιουργηθεί μια εικόνα που ο αναγνώστης να μπορεί να πλοηγηθεί εύκολα μελετώντας την, το heatmap αποδίδει μια σειρά χρωμάτων ανάλογα με τις ομαδοποιήσεις που έχει επιτελέσει. Η επιλογή των χρωμάτων γίνεται από εμάς και ουσιαστικά αντανακλά την πιθανότητα της μηδενικής υπόθεσης. Για να γίνει αυτό περισσότερο αντιληπτό,ας υποθέσουμε ότι έχουμε ένα δείγμα ενός συγκεκριμένου αριθμού πρωτεϊνών και θέλουμε να τις συγκρίνουμε ως προς τρία χαρακτηριστικά. Θεωρούμε πως ο κάθετος άξονας παρουσιάζει τις πρωτεΐνες μας ενώ ο οριζόντιος τα τρία χαρακτηριστικά που διερευνούμε. Σε αυτήν την περίπτωση, θα πρέπει αρχικά να επιλέξουμε ένα χρώμα

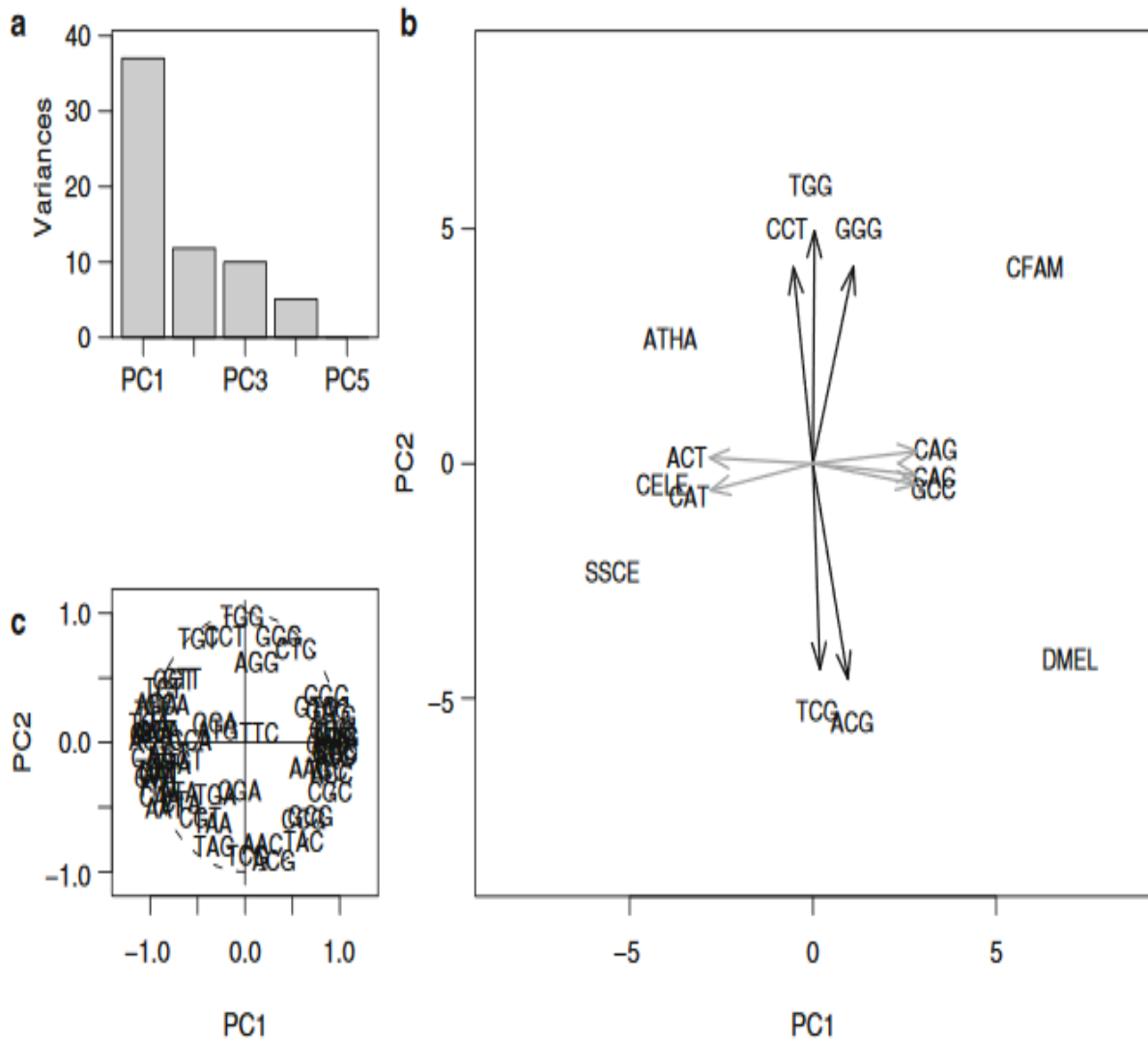
που θα αποδίδεται κάθε φορά που θα συναντάται στο διάγραμμα ένα σημείο το οποίο θα περιλαμβάνει ως  $\gamma$  συντεταγμένη ένα από τα τρία χαρακτηριστικά. Ωστόσο για ένα χαρακτηριστικό δεν αρκεί να οριστεί μόνο ένα χρώμα. Αυτό πηγάζει από το γεγονός πως εάν δύο διαφορετικές πρωτεΐνες παρουσιάζουν το ίδιο χαρακτηριστικό αλλά σε διαφορετική ένταση, δηλαδή δεν ισχύει η δυαδική αρχή για τα μελετούμενα χαρακτηριστικά τύπου 0 ή 1, τότε θα πρέπει η αντίστοιχη ένταση να είναι σε συμφωνία και με μια διαφορετική απόχρωση του προκαθορισμένου χρώματος. Ουσιαστικά για κάθε χαρακτηριστικό δεν επιλέγουμε ένα αλλά τρία χρώματα, με τα δύο να αντιπροσωπεύουν τα άκρα του συνόλου του εύρους των εντάσεων που μπορούν να εμφανιστούν και το τρίτο την μηδενική πιθανότητα. Το πρόγραμμα υπολογίζει τις αντίστοιχες πιθανότητες, αποδίδει τα ανάλογα χρώματα χρώματα χρησιμοποιώντας και τις αντίστοιχες επικαλύψεις αυτών για τιμές που δεν βρίσκονται στα άκρα ή στο κέντρο του συνολικού εύρους και σχεδιάζει το αντίστοιχο δενδρόγραμμα. Στο προηγούμενο παράδειγμα των τριών χαρακτηριστικών επομένως θα ορίζαμε εννέα διαφορετικά χρώματα. Ύστερα από το scaling, πρωτεΐνες με παρόμοια scores τοποθετούνται μαζί ενώ πρωτεΐνες με μακρινά scores τοποθετούνται ξέχωρα. Κατ'αυτόν τον τρόπο όταν τρέξουμε το πρόγραμμα, συνήθως εμφανίζονται στικτές περιοχές με συγκεκριμένο χρωματισμό, διευκολύνοντας την εξαγωγή της οπτικής αλλά και ποσοτικής πληροφορίας. Τα διαγράμματα τύπου heatmap μπορούν να σχεδιαστούν με ποικίλους τρόπους. Ο πιο συνηθισμένος σε υψηλής λεπτομέρειας έρευνες είναι η χρήση προγραμμαμένων ή ο de novo σχεδιασμών πακέτων εντολών στη γλώσσα προγραμματισμού R, ανάλογα με το προφίλ του διαγράμματος που επιθυμούμε να κατασκευάσουμε. Σε αυτό το στάδιο μπορούμε να επιλέξουμε τα χρώματα, το είδος του clustering, το είδος των μεταβλητών και τα στοιχεία των  $x$  και  $\gamma$  αξόνων. Σε προηγμένες εκδόσεις του Heatmap όπως το Heatmap 3, μπορούμε να συμπαρουσιάσουμε και άλλα στοιχεία όπως ο φαινότυπος των γονιδιακών μας δεδομένων. Μια άλλη λύση είναι η χρήση ιντερνετικών προγραμμάτων από sites που έχουν έτοιμες τις εντολές και τοποθετώντας απλά τις επιθυμητές, σχεδιάζουν εύκολα και γρήγορα το προσωπικό μας διάγραμμα.[48]

## 2.6 PCA

Το PCA ή αλλιώς Principal Component Analysis, πρόκειται όπως άλλωστε υποδεικνύει και το όνομά του για ανάλυση βασισμένη σε πρωταρχικά στοιχεία. Ενδελεχέστερα, είναι μια μέθοδος που στόχο έχει την εμβάθυνση σε ορισμένα μόνο από τα δεδομένα μας, εστιάζοντας σε αυτά και αφήνοντας στο περιθώριο άλλα που δε θεωρούνται εξίσου στατιστικώς σημαντικά. Κατ'αυτόν τον τρόπο, επιτυγχάνεται μείωση του αρχικού όγκου, ιεράρχηση των δεδομένων και παράλληλη σχετικά μικρή απώλεια πληροφορίας. Το PCA χρησιμοποιεί μια σειρά αλγορίθμων με σκοπό τη συσχέτιση των εισαχθέντων στοιχείων και κατ'επέκταση τη λήψη αποφάσεων για το ποια από αυτά θεωρούνται υψηλής στατιστικής αξίας. Τα δεδομένα εισάγονται με μορφή αρχείου excel κατά τον κλασικό τρόπο, με οριζόντιες και κατακόρυφες στήλες. Δύο τέτοιοι άξονες για παράδειγμα θα μπορούσαν να είναι ένας αριθμός πεπτιδίων στον άξονα  $x$  και ένας αριθμός διαφορετικών κυττάρων στον άξονα  $\gamma$ . Το συγκεκριμένο παράδειγμα είναι αρκετά απλό αλλά αντανακλά τη βάση λειτουργίας του προγράμματος. Στη συνέχεια, το λογισμικό μετρά τις τιμές που προκύπτουν στα διάφορα κελιά του excel με συντεταγμένες  $x, \gamma$  έως ότου εξαντλήσει το σύνολο των τιμών. Έπειτα, προσπαθεί να τις αξιολογήσει συγκρίνοντας τις τιμές αυτές ως προς την

αξία τους και να κρατήσει στην τελική παρουσίαση μόνο τις σημαντικές.[49] Εδώ μπορούν να συμβαίνουν διαφορετικά γεγονότα. Μια πιθανότητα θα μπορούσε να είναι ότι τα δεδομένα έχουν σχέση ανεστραμμένης συσχέτισης, δηλαδή εάν το πεπτίδιο ένα έχει πολύ υψηλότερη τιμή στο κύτταρο ένα από ότι στο δύο και η σχέση αυτή φθίνει καθώς κινούμαστε προς τα κάτω δηλαδή μέχρι το τελευταίο πεπτίδιο όπου και τελικά αναστρέφεται.(επομένως το τελευταίο πεπτίδιο να παρουσιάζει αναστροφes κατά προσέγγιση τιμές από το πρώτο ως προς τα δύο κύτταρα)Το αποτέλεσμα αυτό θα μπορούσε να ερμηνευτεί από το σύστημα ως συγκράτηση μόνο των αντιπροσωπευτικών τιμών για κάθε πεπτίδιο σε κάθε κύτταρο(τιμές που διαφέρουν αρκετά μεταξύ τους) και πιθανή εξήγηση ότι τα συγκεκριμένα κύτταρα πρόκειται είτε για διαφορετικά είτε για ένα φυσιολογικό και ένα μη φυσιολογικό, με βάση το διαφορετικό προφίλ των πεπτιδίων που περιέχουν. Αυτή είναι η πιο απλή περίπτωση και σπάνια ρεαλιστική. Θα μπορούσε σε άλλο ενδεχόμενο οι σχέσεις πεπτιδίων-κυττάρων να μην εμφανίζουν κάποια συγκεκριμένη γραμμική σχέση αλλά πλήρη τυχαιότητα. Σε ένα παράδειγμα που περιέχει κύτταρα και πεπτίδια ως μεταβλητές, το PCA θα προσπαθήσει σαν πρώτο βήμα να δημιουργήσει clusters(ομαδοποιήσεις) των τιμών πεπτιδίων στα εκάστοτε κύτταρα, χρησιμοποιώντας μόνο τις περισσότερο αντιπροσωπευτικές τιμές. Ειδικότερα, εάν έχουμε μετρήσεις από έξι κύτταρα για δύο γονίδια, το PCA υπολογίζει σε ένα διάγραμμα με άξονες τα δύο πεπτίδια, τη μέση τιμή για το ένα πεπτίδιο από τις τιμές του και στα έξι κύτταρα και την αποδίδει στον αντίστοιχο άξονα του και ακριβώς το ίδιο για το δεύτερο πεπτίδιο. Αυτές οι δύο μέσες τιμές χαρακτηρίζονται κατά συνέπεια από δύο συντετεγμένες η κάθεμια οι οποίες σχεδιάζονται σαν σημείο στο διάγραμμα μας. Έπειτα αυτό το σημείο μεταφέρεται από την αρχική του θέση, στην αρχή των αξόνων(0,0)ως ένα ενιαίο σύστημα μαζί με τις αρχικές τιμές μας για τα έξι πεπτίδια, δηλαδή έχουμε μια μετατόπιση θέσης συστήματος χωρίς να χάνεται όμως πληροφορία καθώς οι σχετικές αποστάσεις μεταξύ των σημείων μας διατηρούνται στο ακέραιο. Σε επόμενο χρόνο, το PCA χαράσσει μια ευθεία γραμμή που περνά από την αρχή των αξόνων. Η γραμμή αυτή περιστρέφεται και μετράται συνεχώς η απόσταση των τιμών δεδομένων μας δηλαδή των τιμών των πεπτιδίων στα δύο κύτταρα. Στόχος είναι η ελαχιστοποίηση αυτών των αποστάσεων και ως εκ τούτου η ελαχιστοποίηση της απόκλισης από αυτήν. Έτσι καταληκτικά αποκτούμε τη λεγόμενη PC1 ευθεία η κλίση της οποίας η σημασία έγκειται στο ότι ερμηνεύει κατά πόσο «ανεβαίνουμε» καθώς προχωράμε δηλαδή το ρυθμό που συναντάμε το πεπτίδιο ένα προς το πεπτίδιο δύο σε κάθε στιγμιότυπο του διαγράμματος ή αλλιώς το βαθμό διασποράς των δύο πεπτιδίων ως προς την ευθεία που σχεδιάσαμε. Ύστερα, το σύστημα χαράσσει μια κάθετη γραμμή σε αυτήν που χάραξε προηγουμένως και μεταφέρει το σχηματισμό σταυρού μαζί με τα μετατοπισμένα πεπτίδια σε ένα νέο πεδίο μέσα από το οποίο πραγματοποιεί υπολογισμούς αποστάσεων από τις δύο αυτές γραμμές και αποδίδει τους αντίστοιχους βαθμούς σημαντικότητας. Κατ'αυτόν τον τρόπο πετύχαμε την απο-μάκρυνση από τα αρχικά δεδομένα κυττάρων -πεπτιδίων και πραγματοποιήσαμε τη μετάβαση στη συσχέτιση των πεπτιδίων χωρίς την κυτταρική παράμετρο, καταφέραμε δηλαδή μείωση ενός βαθμού διάστασης.(dimension downgrade) Αξίζει να αναφερθεί πως η παραπάνω εξήγηση είναι αρκετά εκλαϊκευμένη καθώς στην πραγματικότητα τα χρησιμοποιούμενα μαθηματικά είναι αρκετά περίπλοκα, ανάλογα βέβαια και με το μέγεθος και την ποικιλία των δειγμάτων μας. Μέσω του PCA μπορούμε να αξιολογήσουμε μια συγκεκριμένη παράμετρο που μας ενδιαφέρει, απαλλαγμένη από τις μη επιθυμητές με ταυτόχρονη διατήρηση κάποιων από τα χαρακτηριστικά της αφαιρεμένης πληροφορίας. Αναλύσεις PCA οι οποίες σχετίζονται με σημασία πεπτιδίων ή γονιδίων, ιδίως σε δείγματα με υψηλό

αριθμό δεδομένων, μπορούν επίσης να αποδοθούν σε ραβδογράμματα που εκφράζουν τη γενική εικόνα των αποστάσεων των πεπτιδίων από τους άξονες PC1 και PC2.[50] Ακολουθεί μια εικόνα τυπικής μορφής ανάλυσης PCA:



**Εικόνα 2.11** *Principal Components Analysis. Computational Toxicology, 527–547.* Στο εικονίδιο *a* παρουσιάζονται τα PCA ραβδογράμματα, στο εικονίδιο *b* η συνήθης εικόνα της τελικής επεξεργασίας χωρίς να εμφανίζονται οι άξονες τύπου PCA ενώ στο εικονίδιο *c* έχει γίνει επιστοιβαξη όλων των αρχικών δεδομένων στους PCA άξονες. Για να καταλάβει κανείς το πόσο «καθάρισε» η εικόνα, αρκεί να συγκρίνει τα εικονίδια *b* και *c* μεταξύ τους. [50]

## ΚΕΦΑΛΑΙΟ 3

## ΑΠΟΤΕΛΕΣΜΑΤΑ

### 3.1 Excel Dataset

Η βιοπληροφορική ανάλυση αφορά σε δείγμα που προήλθε από εννέα διαφορετικούς ανθρώπους που πάσχουν από πολλαπλό μυέλωμα στους οποίους μετρήθηκε η συγκέντρωση 59 πρωτεϊνών ύστερα από τη χρήση του φαρμάκου μπορτεζομίμη. Το συγκεκριμένο φάρμακο έχει δράση αναστολέα πρωτεασώματος και επομένως από τις συγκεντρώσεις των πρωτεϊνών μετά τη χορήγησή του, θα μπορούσε κάποιος να αποφασίσει για το ποια είναι η έκβαση της ασθένειας του κάθε αρρώστου. Επειδή χρησιμοποιήσαμε αναστολέα πρωτεασώματος, το αναμενόμενο αποτέλεσμα θα ήταν η πτώση της συγκέντρωσης όλων των πρωτεϊνών του δείγματος. Αυτό ωστόσο δε συμβαίνει. Ακολουθεί το dataset από το αρχικό αρχείο excel που παρουσιάζει την παραπάνω κατάσταση:

**Πίνακας 3.1** *Το dataset των 59 πρωτεϊνών. Πέραν των ονομάτων τους, περιλαμβάνονται στήλες που σχετίζονται με τις συγκεντρώσεις τους μετά τη χορήγηση φαρμάκου, το λόγο  $dr/nr$ , (deep responders/non responders) τη στατιστική σημαντικότητα της κάθεμιας ως προς το σύνολο των αρρώστων αλλά και η συγκέντρωση της κάθεμιας ξεχωριστά στα εννέα διαφορετικά δείγματα.*





του γονιδίου από το οποίο προήλθαν. Η στήλη D παρουσιάζει το όνομα των πρωτεϊνών ενώ η στήλη F τα αποτελέσματα του Mann-Whitney U test, ενός test δηλαδή που χρησιμοποιείται σε μελέτες για μη παραμετρικά δεδομένα που η χρησιμότητά του έγκειται στη διαπίστωση για το αν τα δεδομένα της έρευνάς μας προέρχονται από πληθυσμούς που εμφανίζουν παρόμοιες κατανομές. Από το σύνολο των αρχικών πρωτεϊνών, μόνο αυτές οι 59 είχαν scores από το mw-test που να αξιολογηθούν ως στατιστικά σημαντικές και γι'αυτό και παρουσιάζονται αυτές μόνο. Οι στήλες J έως R εμφανίζουν τις συγκεντρώσεις των πρωτεϊνών στους ασθενείς ενώ οι H και I αναφέρονται στο μέσο όρο DR και NR. Για να γίνει αυτό καλύτερα αντιληπτό, η στήλη I περιέχει τιμές που προκύπτουν από το μέσο όρο της συγκέντρωσης της εκάστοτε πρωτεΐνης στους Deep Responders(δηλαδή σε αυτούς που ανταποκρίθηκαν στη θεραπεία) και υπολογίζεται από το άθροισμα των τιμών των στηλών O,P,Q,R και στη συνέχεια ακολουθεί διαίρεση με το τέσσερα και ακριβώς η ίδια διαδικασία επαναλαμβάνεται για τη στήλη H αλλά αυτή τη φορά με τιμές από τις στήλες J,K,L,M,N και διαίρεση με το πέντε στη συνέχεια. Άρα, σύμφωνα με το προφίλ αυτών των υπολογισμένων συγκεντρώσεων αλλά και με βάση την κλινική τους έκβαση, οι ασθενείς έχουν διαχωριστεί σε τέσσερις deep responders και πέντε non responders. Τέλος, η στήλη E παρουσιάζει τους λόγους που προκύπτουν από τη διαίρεση των τιμών των στηλών H και I ,προσφέροντας μια εικόνα όχι πλέον με επίκεντρο την κατάσταση του ασθενούς αλλά του κατά πόσο η κάθε πρωτεΐνη ξεχωριστά εμφανίζεται ή μόνο σε deep responders ή μόνο σε non responders ή και στα δύο είδη ασθενών αλλά με διαφορετική συγκέντρωση. Μετά την ανάλυση τύπου PCA θα δοθεί μια ενδελεχής εξήγηση για το πως η στήλη E(DR/NR) έχει μεγάλη σημασία για την αναζήτηση βιοδεικτών καθώς οι πρωτεΐνες που αντιστάθηκαν στον αναστολέα πρωτεασώματος και συνεχίζουν να βρίσκονται σε σημαντικές συγκεντρώσεις στους ασθενείς πιθανότατα να μπορούσαν να διερευνηθούν περαιτέρω ως δυνητικοί βιοδείκτες. Εάν και εκ προοιμίου γνωρίζουμε πως η προσοχή μας θα πρέπει να εστιαστεί σε πρωτεΐνες που εμφανίζονται κυρίως σε non responders, θα πρέπει να προηγηθεί στατιστική ανάλυση PCA για να εξακριβώσουμε ποιες από αυτές έχουν τη μεγαλύτερη στατιστική αξία και επομένως τη μεγαλύτερη πιθανότητα να είναι πραγματικά πιθανοί βιοδείκτες.

### 3.2 Λειτουργίες πρωτεϊνών

Στη συνέχεια παρουσιάζονται συνοπτικά οι λειτουργίες και των 59 πρωτεϊνών ούτως ώστε να υπάρχει η δυνατότητα της εύκολης ανάτρεξης σε αυτές αλλά και μια σφαιρική εικόνα του ρόλου τους.

- 1) **ID: U3IP2\_HUMAN / Gene Name: RRP9 / Protein name: U3 small nucleolar RNA interacting protein** - Πρόκειται για πρωτεΐνη η οποία αποτελεί τμήμα ενός μικρού πηρυνικού ριβονουκλεοπρωτεϊνικού σωματιδίου(snoRNP) και πιθανώς σχετίζεται με την τροποποίηση του προριβοσωματικού RNA.(pre-RNA)
- 2) **ID: SYHC\_HUMAN / Gene Name: Hars / Protein name: Histidine tRNA ligase,cytoplasmic** – Πρωτεΐνη που χρησιμοποιεί ATP ως διαμεσολαβητικό μόριο για να συνδέσει ιστιδίνη με μόρια tRNA.
- 3) **ID: ERG7\_HUMAN / Gene Name: LSS / Protein name: Lanosterol synthase** – Καταλύει την κυκλοποίηση του 2,3 οξειδοσκουαλενίου προς λανοστερόλη.

- 4) **ID: RM12\_HUMAN / Gene Name: MRPL12 / Protein name: 39S ribosomal protein L12, mitochondrial** - Πρωτεΐνη που διαμεσολαβεί στη σύνδεση της RNA πολυμεράσης στον υποκινητή για την έναρξη της μεταγραφής.
- 5) **ID: RAB1A\_HUMAN / Gene Name: RAB1A / Protein name: Ras-related protein Rab-1A** - GTPάση η οποία δρα τόσο στο Golgi όσο και στο ενδοπλασματικό δίκτυο με κύριο στόχο το πακετάρισμα βιολογικών μορίων, τη μεταφορά τους μέσω κυστιδίων επικαλυμμένων με πρωτεϊνικά πλέγματα και τη σύντηξη των κυστιδίων αυτών με τις κυτταρικές μεμβράνες στόχους.
- 6) **ID: RS24\_HUMAN / Gene Name: RPS24 / Protein name: 40S ribosomal protein S24** - Πρωτεΐνη απαραίτητη για την επεξεργασία του preRNA και ωρίμανση των υπομονάδων της περιοχής 40s των ριβοσωμάτων.
- 7) **ID: MBLC2\_HUMAN / Gene Name: MBLAC2 / Protein name: Metallo-beta-lactamase domain-containing protein 2-** Μεταλλοπρωτεΐνη με κύριο ρόλο την προσθήκη μέχρι δύο ιόντων ψευδαργύρου σε ποικίλες αντιδράσεις.
- 8) **ID: EH1L1\_HUMAN / Gene Name: EHBP1L1 / Protein name: EH domain-binding protein 1-like protein 1-** Πιθανώς να συνδέεται με μεταφορά κυστιδίων.
- 9) **ID: ARMX3\_HUMAN / Gene Name: ARMCX / Protein name: Armadillo repeat-containing X-linked protein 3** - Πρωτεΐνη που σχετίζεται με τη μεταφορά αλλά και συσσώρευση μιτοχονδρίων στα νευρικά κύτταρα ενώ επίσης μάλλον εμπλέκεται και στο σύστημα κινεσίνης – TPAK 2 όπου αφορά στη μετακίνηση κυστιδίων κατά μήκος του νευράξονα.
- 10) **ID: RAP2C\_HUMAN / Gene Name: RAP2C / Protein name: Ras-related protein Rap-2c-** Μικρό μόριο με δράση GTPάσης που εμπλέκεται στην ανασύσταση της δομής του κυτταροσκελετού και πιθανώς σε μεταγραφή γονιδίων.
- 11) **ID: BDH\_HUMAN / Gene Name: BDH1 / Protein name: D-beta-hydroxybutyrate dehydrogenase, mito-chondrial** - Πρωτεΐνη που παρουσία φωσφατιδυλοχολίνης καταλύει την αντίδραση μετατροπής του υδρόξυβουτυρικού οξέος σε ακετοακετικό οξύ(ανιόν) με παράλληλη αποθήκευση ενέργειας σε ένα μόριο νικοτινάμιδουαδένινοουκλεοτιδίου.(NAD<sup>+</sup>)
- 12) **ID: MTCH2\_HUMAN / Gene Name: MTCH2 / Protein name: Mitochondrial carrier homolog 2 –** Πρωτεΐνη που μάλλον σχετίζεται με τα μιτοχόνδρια.
- 13) **ID: MAP2\_HUMAN / Gene Name: METAP2 / Protein name: Methionine aminopeptidase 2 –** Πρωτεΐνη που αφαιρεί ένα αμινοξύ μεθειονίνης σε νεοσυντιθέμενες πρωτεΐνες ενώ επίσης ασκεί προστατευτική δράση ως προς τον ευκαρυωτικό παράγοντα EIF2S1 μέσω ρύθμισης σειράς φωσφορυλίσεων, ο οποίος συμμετέχει στην πρωτεϊνική σύνθεση.
- 14) **ID: EI2BD\_HUMAN / Gene Name: EIF2B4 / Protein name: Translation initiation factor eIF-2B subunit delta** - Πρωτεΐνη που επιτρέπει στον ευκαρυωτικό μεταγραφικό παράγοντα εκκίνησης τύπου 2 να ανταλλάσει το GDP για GTP και έτσι να τον ενεργοποιεί.

- 15) **ID: SSRA\_HUMAN / Gene Name: SSR1 / Protein name: Translation initiation factor eIF-2B subunit delta** – Πρωτεΐνη που ανήκει στην κατηγορία πρωτεϊνών τύπου TRAP οι οποίες σχετίζονται με τη κατακράτηση ασβεστίου στο ενδοπλασματικό δίκτυο και κατ'επέκταση τη ρύθμιση πρωτεϊνών που διαθέτουν ασβε-στοεξαρτώμενη δράση.
- 16) **ID: RS9\_HUMAN / Gene Name: RPS9 / Protein name: 40S ribosomal protein S9** – Πρωτεΐνη που σχετίζεται με τη δομή της υπομονάδας 40s του ριβοσώματος.
- 17) **ID: NB5R1\_HUMAN / Gene Name: CYB5R1 / Protein name: NADH-cytochrome b5 reductase 1** – NADH αφυ-δρογονάση που αφορά στο κυτόχρωμα 5b του μιτοχονδρίου και σχετίζεται με το μεταβολισμό των λιπαρών οξέων, της σύνθεσης χοληστερόλης και το μεταβολισμό ηπατικών φαρμάκων.
- 18) **ID: SDHB\_HUMAN / Gene Name: SDHB / Protein name: Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial** – Σιδηροπρωτεΐνη τμήμα του συμπλόκου της ηλεκτρικής δεϋδρογονάσης η οποία αποτελεί μέρος της αναπνευστικής αλυσίδας μεταφοράς ηλεκτρονίων στο μιτοχόνδριο κατά την παραγωγή ενέργειας.
- 19) **ID: MPPA\_HUMAN / Gene Name: PMPCA / Protein name: Mitochondrial-processing peptidase subunit alpha** – Πρωτεΐνη που αποτελεί τμήμα της μιτοχονδριακής πρωτεάσης MPP η οποία σχετίζεται με την αφαίρεση του σήματος εισαγωγής πρωτεϊνών στο μιτοχόνδριο.
- 20) **ID: TIM44\_HUMAN / Gene Name: TIMM44 / Protein name: Mitochondrial import inner membrane translocase subunit TIM44** – Τμήμα του πρωτεϊνικού συμπλόκου PAM το οποίο εμπλέκεται στη μεταφορά πρωτεϊνών από τη εσωτερική κυτταρική μεμβράνη στη μήτρα του μιτοχονδρίου μέσω αξιοποίησης της ενέργειας υδρόλυσης του ATP.
- 21) **ID: APOL2\_HUMAN / Gene Name: APOL2 / Protein name: Apolipoprotein L2** – Πρωτεΐνη που σχετίζεται με τη μεταφορά λιπιδίων αλλά και τη δέσμευση λιπιδίων.
- 22) **ID: TMX3\_HUMAN / Gene Name: TMX3 / Protein name: Protein disulfide-isomerase TMX3** – Πρωτεΐνη που σχετίζεται με την αλλαγή θέσης των δισουλφιδικών δεσμών σε πρωτεΐνες.
- 23) **ID: ODB2\_HUMAN / Gene Name: DBT / Protein name: Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial** – Δεϋδρογονάση η οποία μετατρέπει τα άλφα κετοξέα σε άκυλοCoA και CO<sub>2</sub>.
- 24) **ID: IMDH2\_HUMAN / Gene Name: IMPDH2 / Protein name: Inosine-5'-monophosphate dehydrogenase 2** – Καταλύει την μετατροπή της 5 φωσφορικής ινοσίνης σε 5 φωσφορική ξανθινοσίνη και συμμετέχει στο σχηματισμό των νουκλεοτιδίων που περιέχουν γουανίνη.
- 25) **ID: DDX21\_HUMAN / Gene Name: DDX21 / Protein name: Nucleolar RNA helicase 2** – Η συγκεκριμένη πρωτεΐνη βοηθά στο ξετύλιγμα του DNA ούτως ώστε οι θηλιές που σχηματίζονται να αποτελέσουν εφалτήριο για τη δράση της RNA πολυμεράσης 1 και 2.

- 26) **ID: RL12\_HUMAN / Gene Name: RRP9 / Protein name: 60S ribosomal protein L12** – Πρωτεΐνη που συνδέεται στην υπομονάδα 26 του ριβοσώματος.
- 27) **ID: FAF2\_HUMAN / Gene Name: FAF2 / Protein name: FAS-associated factor 2** – Συμμετέχει στο μεταβολικό μονοπάτι δράσης απόπτωσης κακοδυλωμένων ή χημικά μεταλλαγμένων πρωτεϊνών που λαμβάνει χώρα στο ενδοπλασματικό δίκτυο.
- 28) **ID: GLSK\_HUMAN / Gene Name: GLS / Protein name: Glutaminase kidney isoform, mitochondrial** – Πρωτεΐνη που συνδέεται με την οξεοβασική ισορροπία αλλά και με το πρώτο βήμα της αντίδρασης μεταβολισμού της γλουταμίνης στο νεφρό.
- 29) **ID: COX6C\_HUMAN / Gene Name: COX6C / Protein name: Cytochrome c oxidase subunit 6C** – Αποτελεί μια από τις πρωτεΐνες της οξειδάσης του κυτοχρώματος, του τελευταίου συμπλόκου της αναπνευστικής αλυσίδας.
- 30) **ID: RL10\_HUMAN / Gene Name: RPL10 / Protein name: 60S ribosomal protein L10** – Συστατικό της μεγάλης ριβοσωμικής υπομονάδας και σχετίζεται με την πρωτεϊνοσύνθεση.
- 31) **ID: SRPRB\_HUMAN / Gene Name: SRPRB / Protein name: Signal recognition particle receptor subunit beta** – Πρωτεΐνη κομμάτι του SRP (signal recognition particle) που αναγνωρίζει το σήμα για το ενδοπλασματικό δίκτυο, συνδέεται με αυτό και οδηγεί σε ξεγλύστριμα της πρωτεΐνης στο εσωτερικό του.
- 32) **ID: TBA1C\_HUMAN / Gene Name: RRP9 / Protein name: Tubulin alpha-1C chain** – Αποτελεί τμήμα της τουμπουλίνης α, ενός εκ των δομικών λίθων του πολυμερισμού των μικροσωληνίσκων.
- 33) **ID: SAR1A\_HUMAN / Gene Name: SAR1A / Protein name: GTP-binding protein SAR1a** – Πρωτεΐνη που υποβοηθά τη μετακίνηση βιομορίων από το ενδοπλασματικό δίκτυο στο Golgi και επιπλέον διαδραματίζει ρόλο στη διαμερισματοποίηση των παραπάνω οργάνων.
- 34) **ID: TBA4A\_HUMAN / Gene Name: TUBA4A / Protein name: Tubulin alpha-4A chain protein** – Αποτελεί τμήμα της τουμπουλίνης α, ενός εκ των δομικών λίθων του πολυμερισμού των μικροσωληνίσκων.
- 35) **ID: ECHA\_HUMAN / Gene Name: HADHA / Protein name: Trifunctional enzyme subunit alpha, mitochondrial** – Το ένζυμο αυτό καταλύει τις τελευταίες τρεις από τις συνολικά τέσσερις αντιδράσεις κατά τη β οξείδωση λιπαρών οξέων.
- 36) **ID: TBB4B\_HUMAN / Gene Name: TUBB4B / Protein name: Tubulin beta-4B chain** – Αποτελεί τμήμα της τουμπουλίνης β, ενός εκ των δομικών λίθων του πολυμερισμού των μικροσωληνίσκων.
- 37) **ID: TBB5\_HUMAN / Gene Name: TUBB / Protein name: Tubulin beta chain** – Αποτελεί τμήμα της τουμπουλίνης β, ενός εκ των δομικών λίθων του πολυμερισμού των μικροσωληνίσκων.
- 38) **ID: IQGA1\_HUMAN / Gene Name: IQGAP1 / Protein name: Ras GTPase-activating-like protein IQGAP1** – Πρωτεΐνη που σχετίζεται με τη δομή της της ακτίνης του κυτταροσκελετού ενώ έχει

πιθανολογείται ότι συνδέεται με την καλμοδουλίνη(ασβεστοεξαρτώμενη πρωτεΐνη) και την ενεργοποίηση του αντιγόνου CDC42 χωρίς όμως να έχει γνωστή δράση στα δύο τελευταία.

39) **ID: SRSF1\_HUMAN / Gene Name: SRSF1 / Protein name: Serine/arginine-rich splicing factor 1** – Εμπλέκεται με στοιχεία του σπлайισώματος επιτρέποντας τη συγκράτηση των εξονίων στη σωστή τους θέση πριν το splicing.

40) **ID: ACAD8\_HUMAN / Gene Name: ACAD8 / Protein name: Isobutyryl-CoA dehydrogenase, mitochondrial** – Δεϋδρογονάση του 2-μέθυλοπροπάνυλο CoA σε 2-μέθυλοπροπένυλο CoA με ταυτόχρονη αναγωγή ενός νουκλεοτιδίου FAD σε FADH<sub>2</sub>.

41) **ID: HSDL2\_HUMAN / Gene Name: HSDL2 / Protein name: Hydroxysteroid dehydrogenase-like protein 2** – Πρωτεΐνη άγνωστης δράσης.

42) **ID: ING5\_HUMAN / Gene Name: ING5 / Protein name: Inhibitor of growth protein 5** – Συστατικό του συμπλόκου HBO1 το οποίο έχει δραστηριότητα ακετυλοτρανσφεράσης ειδικής σε ιστόνη H4, μειωμένη δραστηριότητα προς την ιστόνη H3 και είναι υπεύθυνη για το μεγαλύτερο μέρος της ακετυλίωσης H4 ιστόνης ή νίνο. Συστατικό του συμπλόκου MOZ / MOPP το οποίο έχει δραστηριότητα ακετυλοτρανσφεράσης H3 ιστόνης.

43) **ID: PLEC\_HUMAN / Gene Name: PLEC / Protein name: Plectin** – Πρωτεΐνη η οποία επιτρέπει τη διασύνδεση μικροϊνιδίων είτε μεταξύ τους είτε με μικροσωληνίσκους και προσδίδει στο κύτταρο μηχανική αντοχή.

44) **ID: LPAR3\_HUMAN / Gene Name: LPAR3 / Protein name: Lysophosphatidic acid receptor 3** – Πρωτεΐνη με δράση υποδοχέα λυσοφωσφατιδικού οξέος που ίσως να σχετίζεται με καρκίνο ωοθηκών.

45) **ID: BST2\_HUMAN / Gene Name: BST2 / Protein name: Bone marrow stromal antigen 2** – Πρωτεΐνη με δράση IFN-επαγόμενου παράγοντα περιορισμού κατά του ιού ξενιστή, ο οποίος εμποδίζει αποτελεσματικά την απελευθέρωση διαφόρων ιών θηλαστικών μέσω της άμεσης προσκόλλησης στις μεμβράνες των μολυσμένων κυττάρων. Οι δεσμευμένοι ιοί μπορούν να εσωτερικοποιηθούν με ενδοκύτωση και στη συνέχεια αποικοδομούνται ή μπορούν να παραμείνουν στην κυτταρική επιφάνεια. Έχει σχετιστεί μεταξύ άλλων και με δύο ορότυπους του HIV, τον HIV-1 και HIV-2.

46) **ID: K2C6B\_HUMAN / Gene Name: KRT6B / Protein name: Keratin, type II cytoskeletal 6B** - Μια από τα πολλά είδη κερατινών, ινώδων πρωτεϊνών τριπλής αλυσίδας που περιελίσσονται μεταξύ τους και σχηματίζουν μέσω των ισχυρών υδρόφοβων αλληλεπιδράσεων, ισχυρούς δεσμούς ενώ παράλληλα συνδέονται ομοιοπολικά με άλλα βιομορία συνεισφέροντας στη μηχανική αντοχή.

47) **ID: GNAI2\_HUMAN / Gene Name: GNAI2 / Protein name: Guanine nucleotide-binding protein G(i) subunit alpha-2** – Πρωτεΐνη που δεσμεύει νουκλεοτίδια γουανίνης και εμπλέκεται στη διαμεμβρανική σηματοδότηση.

48) **ID: IGK\_HUMAN / Gene Name: IGK / Protein name: Immunoglobulin kappa light chain** – Πρωτεΐνη απαραίτητη για τη σύνθεση της ελαφριάς αλυσού των αντισωμάτων η οποία από μόνη της δε φέρει

αυτόνομη λειτουργία ωστόσο η παρουσία της έχει συσχετιστεί με αρκετά είδη καρκίνων όπως εκείνο του καρκίνου της ουροδόχου κύστεως και του πολλαπλού μυελώματος.

49) **ID: SRP14\_HUMAN / Gene Name: RRP9 / Protein name: SRP14** – Πρωτεΐνη υποδοχέας του σήματος οδηγού των πρωτεϊνών στο ενδοπλασματικό δίκτυο. Μετά την πρόσδεση τους στις υποψήφιες ενδοπλασματικές πρωτεΐνες, τις βοηθά να ξετυλιχθούν και να εισαχθούν στο εσωτερικό του.

50) **ID: ANX11\_HUMAN / Gene Name: ANXA11 / Protein name: Annexin A11** – Πρωτεΐνη που δρα μέσω της σύνδεσής της σε μια ασβεστοεξαρτώμενη πρωτεΐνη ονόματι καλκυκλίνη και είναι απαραίτητη στο τελικό στάδιο της κυτταροκίνησης.

51) **ID: UBE2N\_HUMAN / Gene Name: UBE2N / Protein name: Ubiquitin-conjugating enzyme E2 N** – Το δεύτερο κατά σειρά ένζυμο που εμπλέκεται στην ουμπικιτινοποίηση των πρωτεϊνών.

52) **ID: CSN1\_HUMAN / Gene Name: GPS1 / Protein name: COP9 signalosome complex subunit 1** - Βασική πρωτεΐνη του συμπλόκου σήματος COS9 (CSN), ένα σύμπλεγμα που εμπλέκεται σε διάφορες κυτταρικές και αναπτυξιακές διεργασίες. Το σύμπλοκο CSN είναι ένας βασικός ρυθμιστής της οδού σύζευξης ουμπικουϊτίνης με τη μεσολάβηση της αποδενυλίωσης των υπομονάδων κουλίνης των συμπλεγμάτων λιγάσης ΕΡ τύπου SCF, οδηγώντας σε μείωση της δραστηριότητας της λιγκάσης Ubl συμπλόκων τύπου SCF όπως SCF, CSA ή DDB2 . Το σύμπλεγμα εμπλέκεται επίσης στη φωσφορυλίωση του p53 / TP53 και άλλων στόχων.

53) **ID: UB2V1\_HUMAN / Gene Name: UBE2V1 / Protein name: Ubiquitin-conjugating enzyme E2 variant 1** – Σχετίζεται με ένα μονοπάτι ουμπικιτινοποίησης ονόματι IKK που δε φαίνεται να σχετίζεται με αποδόμηση. Δε διαθέτει αυτόνομη δράση. Έχει επίσης συμμετοχή σε μεταγραφή διαφόρων κυτταροκινών όπως NF-Kappa, ιντερλευκινών και παραγόντων TNF.(tumor necrosis factor)

54) **ID: SC23A\_HUMAN / Gene Name: SEC23A / Protein name: Protein transport protein Sec23A** – Πρωτεΐνη που αποτελεί τμήμα του συστήματος COPII το οποίο εμπλέκεται με το σχηματισμό κυστιδίων που προορίζονται για διακυτταρική επικοινωνία.

55) **ID: CDC37\_HUMAN / Gene Name: CDC37 / Protein name: Hsp90 co-chaperone Cdc37** – Συμπράγοντας χαπερονινών οι οποίες εμπλέκονται στην ενεργοποίηση διαφορετικών κινασών και τελικά στη ρύθμιση πολυποίκιλων διεργασιών όπως τα συστήματα πρωτεϊνών θερμικού shock.(HSP90)

56) **ID: BOLA2\_HUMAN / Gene Name: BOLA2 / Protein name: Bola-like protein 2** – Πρωτεΐνη με δράση στο κυτταρόπλασμα όπου σχηματίζει συσώματα σιδήρου – θείου τα οποία και χρησιμοποιεί για να προσδεθεί σε άλλες κυτταροπλασματικές πρωτεΐνες και να μεταβάλει τη δραστηριότητα τους.

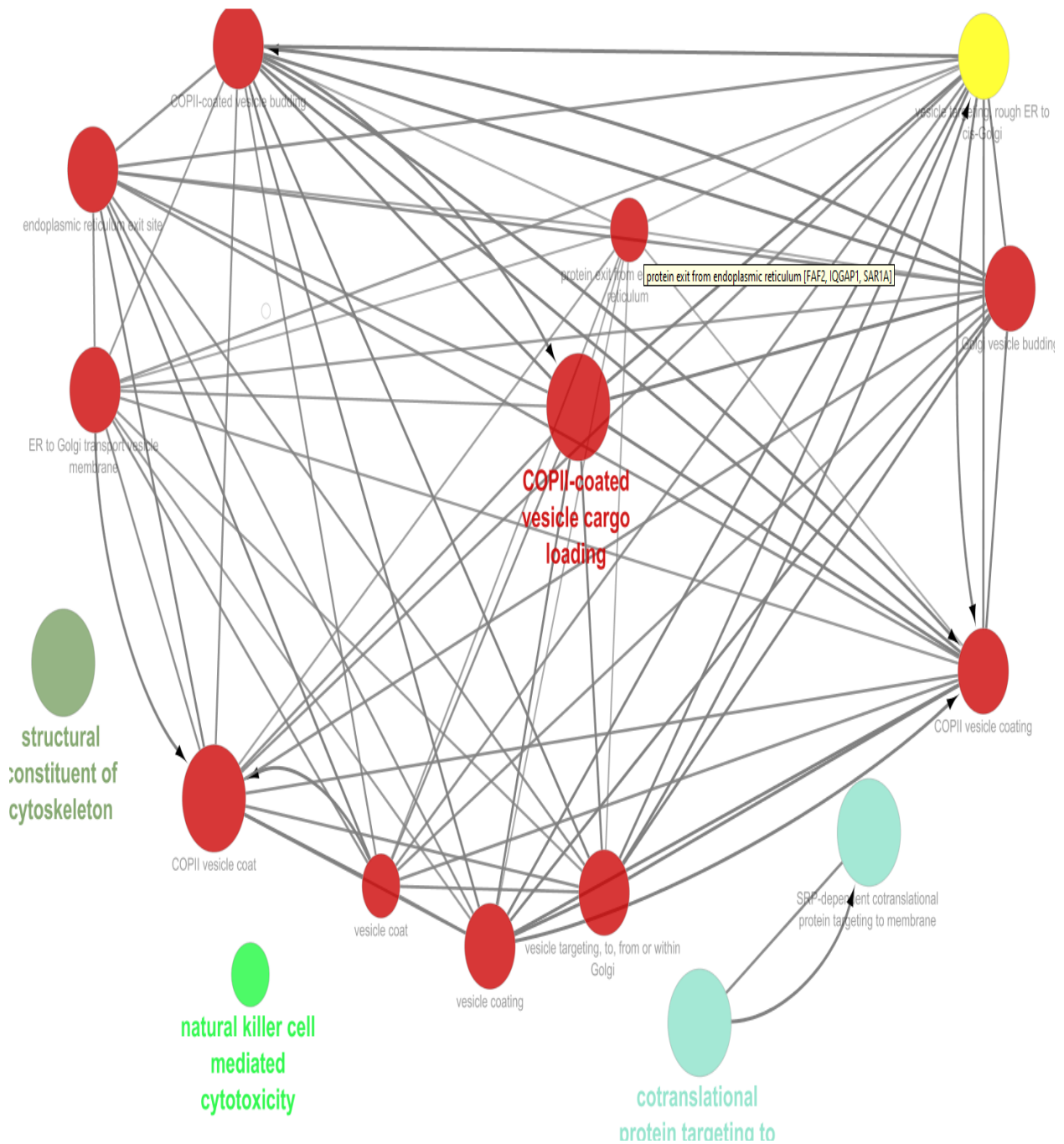
57) **ID: PPAL\_HUMAN / Gene Name: ACP2 / Protein name: Lysosomal acid phosphatase** – Ένζυμο που καταλύει τη μετατροπή φωσφορικών μονοεστέρων παρουσία νερού σε αλκοόλη και μια ομάδα φωσφορικού.

58) **ID: 1B08\_HUMAN / Gene Name: HLA-B / Protein name: HLA class I histocompatibility antigen, B-8 alpha chain** – Πρωτεΐνη με μη γνωστή λειτουργία.

59) ID: **ACSF3\_HUMAN** / Gene Name: **ACSF3** / Protein name: **Acyl-CoA synthetase family member 3, mitochondrial** - Καταλύει την αρχική αντίδραση στη σύνθεση μιτοχονδριακών λιπαρών οξέων, μετατρέποντας το μηλικό και μεθυλομηλικό στον αντίστοιχο θειοεστέρα του CoA.

### 3.3 Cluego Αποτελέσματα

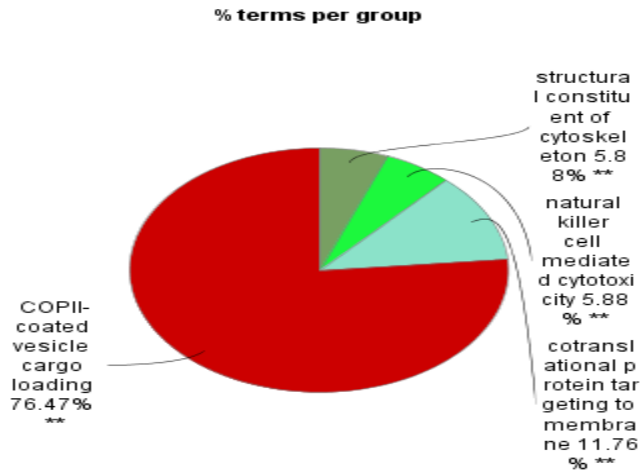
Στόχος της χρήσης του Cluego plugin ήταν αφενός η κατάταξη των πρωτεϊνών σε οικογένειες ως προς τη λειτουργία και την τοποθεσία τους και αφετέρου η συσχέτισή τους με τα αντίστοιχα γονίδια. Ακολουθούν δύο τύποι ανάλυσης, ο πρώτος αφορά στη λειτουργία και ο δεύτερος στην κυτταρική τους τοποθέσια. Όσον αφορά στην ανάλυση ως προς τις βιολογικές διεργασίες, ακολουθεί το διάγραμμα:



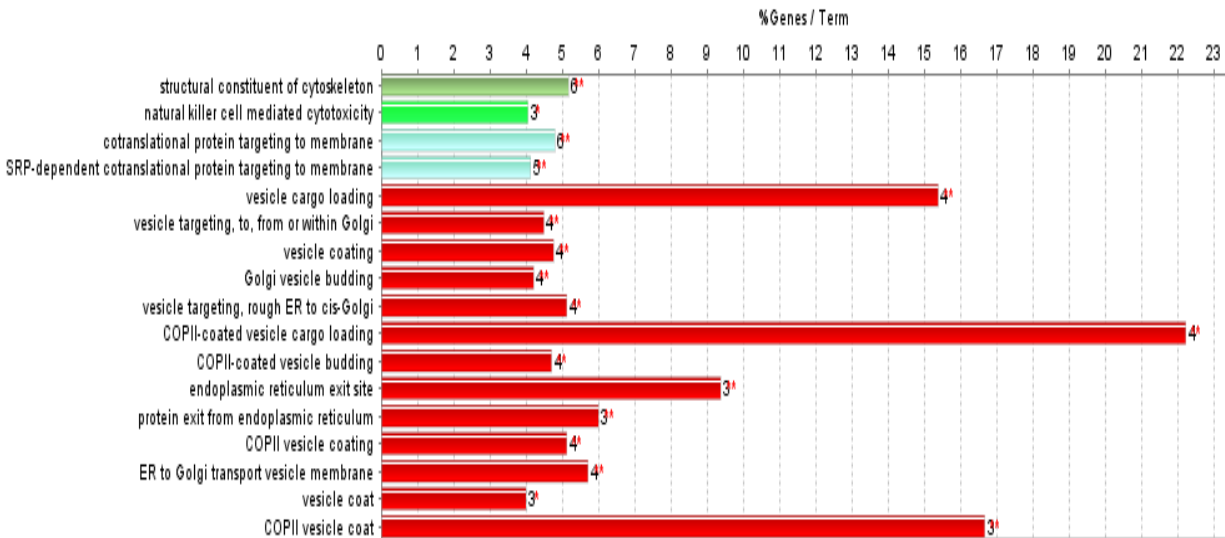
**Εικόνα 3.1** Ανάλυση τύπου *biological processes*. Αποτυπώνονται σαφώς οι κύριες κατηγορίες(τέσσερις) βιολογικών διαδικασιών στις οποίες συμμετέχουν οι πρωτεΐνες του δείγματός μας: α)πρωτεΐνες με δομικό ρόλο στον κυτταροσκελετού β)πρωτεΐνες που σχετίζονται με κύτταρα φυσικούς φονείς γ)πρωτεΐνες που σχετίζονται με μετακίνηση άλλων πρωτεϊνών διαμέσου μεμβράνων ενόσω αυτές μεταφράζονται και δ)πρωτεΐνες που σχετίζονται με το σύστημα COPII το οποίο αφορά σε κατασκευή καλύμματος των κυστιδίων που επρόκειτο να αλλάξουν κυτταρική θέση ή να εξωκυττωθούν με στόχο να εισαχθούν σε άλλα κύτταρα.



Από τις τέσσερις κατηγορίες, η μεγάλη πλειοψηφία των πρωτεϊνών φαίνεται να σχετίζεται με το σύστημα COPII το οποίο είναι υπεύθυνο για την επικάλυψη κυστιδίων που προορίζονται να μεταφέρουν το περιεχόμενό τους είτε σε κάποιο διαφορετικό κυτταρικό διαμέρισμα είτε σε άλλα κύτταρα. Στη συνέχεια, παρουσιάζονται δύο είδη εικόνων, μιά με τη μορφή πίτας και μια με τη μορφή διαγράμματος, οι οποίες αφορούν στο είδος της οικογένειας των γονιδίων από τις οποίες προέκυψαν οι 59 πρωτεΐνες:

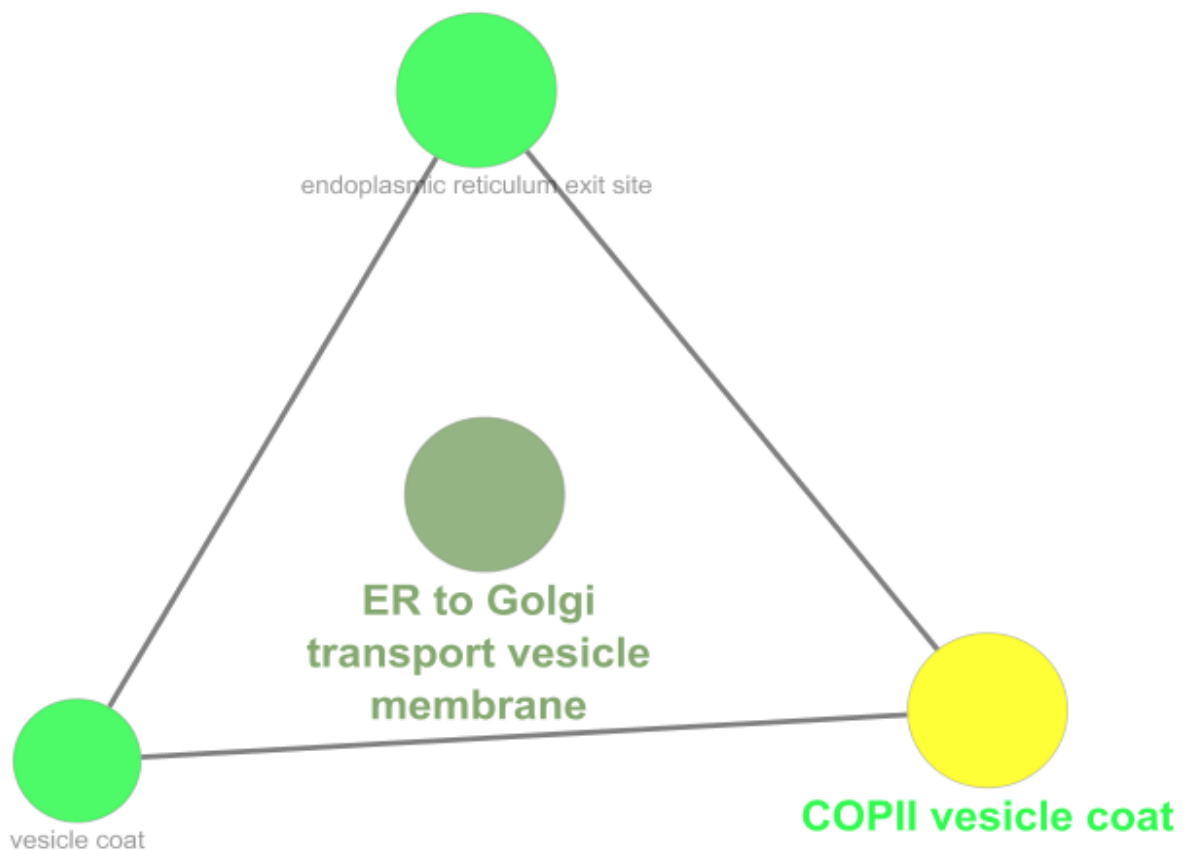


**Εικόνα 3.2** Οι τέσσερις κατηγορίες αποτυπώνονται ως ποσοστά συμμετοχής στις τέσσερις γονιδιακές οικογένειες από τις οποίες πηγάζουν οι 59 πρωτεΐνες. Είναι εμφανές πως οι περισσότερες πρωτεΐνες ανήκουν στην κατηγορία του COPII όπως έχει ήδη αναφερθεί.



**Εικόνα 3.3** Συμπληρωματική της εικόνας 3.2. Εδώ παρουσιάζονται αναλυτικά όλα τα είδη των υποοικογενειών της κάθε κατηγορίας γονιδίου. Επιπλέον, οι αριθμοί που συνοδεύουν την κάθε ράβδο αντιστοιχούν στο ποσοστό των γονιδίων του δείγματός μας ως προς την υπεροικογένεια που αυτά ανήκουν.

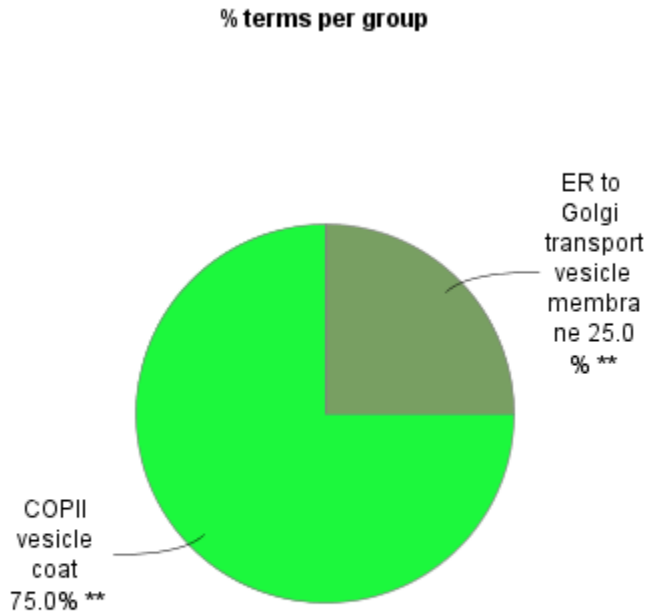
Για να γίνει η εικόνα 3.3 περισσότερο αντιληπτή, εάν θεωρήσουμε πως τα γονίδια των πρωτεϊνών του δείγματος μας ανήκουν στην κατηγορία που αναγράφεται στο αριστερό μέρος και αυτή η κατηγορία αποτελεί στο σύνολό της το 100%, ο ακριβώς από πάνω άξονας αντιπροσωπεύει το ποσοστό των γονιδίων του δείγματος μας που ανήκει στη συγκεκριμένη κατηγορία. Για παράδειγμα η πρώτη κατηγορία που αφορά στον κυτταροσκελετό, στο δείγμα μας υπάρχουν πρωτεΐνες που καλύπτουν ως προς το γονιδιακό τους προφίλ το 5.5% του συνόλου των γονιδίων που έχουν επισημειωθεί πως ανήκουν σε αυτήν την κατηγορία. Καθίσταται σαφές πως η συντριπτική πλειοψηφία των πρωτεϊνών μας αφορά στο σύστημα COPII της κατασκευής κυστιδίων. Αντίστοιχα, ο αριθμός δεξιά των ραβδογραμμάτων είναι ενδεικτικός του αριθμού των γονιδίων του δείγματος που ανήκουν σε αυτήν την κατηγορία (στο προηγούμενο παράδειγμα ισοδυναμεί με 6) ως προς το σύνολο των γονιδίων του δείγματος.(59) Ακολουθεί η ανάλυση κυτταρικής τοποθεσίας:



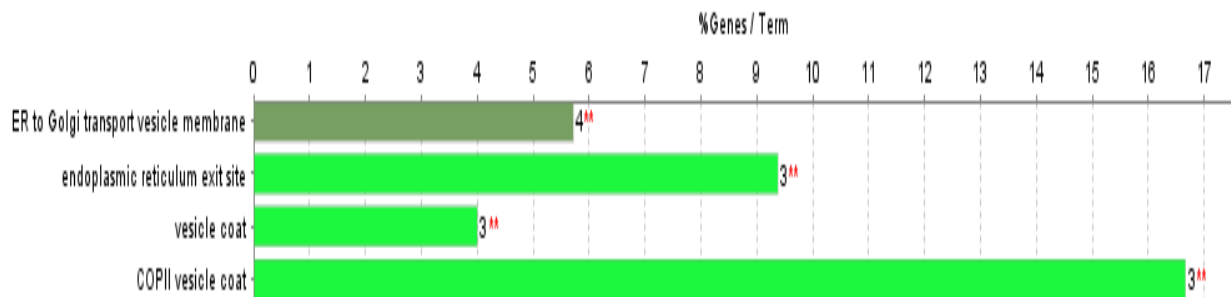
**Εικόνα 3.4** Παρουσιάζονται οι κυτταρικές τοποθεσίες δράσης των πρωτεϊνών: α)πρωτεϊνικό πλέγμα επένδυσης κυστιδίων του συστήματος COPII με τις υποκατηγορίες του(επένδυση κυστιδίων και έξοδος από το ενδοπλασματικό δίκτυο) και β)πρωτεΐνες σε κυτίδα που μεταφέρονται από το ενδοπλασματικό δίκτυο στο σύστημα Golgi.

Παρατηρούμε και σε αυτήν την περίπτωση πως οι πρωτεΐνες μας αφορούν αποκλειστικά στη μεταφορά κυστιδίων. Η ανάλυση αυτού του τύπου δεν έχει ως στόχο να παρουσιάσει μεμονωμένα τη

δράση και τοποθεσία της κάθε πρωτεΐνης αλλά να προσφέρει μια ολιστική προσέγγιση με βάση το εισαχθέν dataset και την αναζήτηση βιβλιογραφίας για το ποιες είναι οι περισσότερο πιθανές κυτταρικές περιοχές όπου θα μπορούσαμε να τις συναντήσουμε. Ακολουθεί η πίνα και το ραβδόγραμμα μα γονιδίων όπως και προηγουμένως:



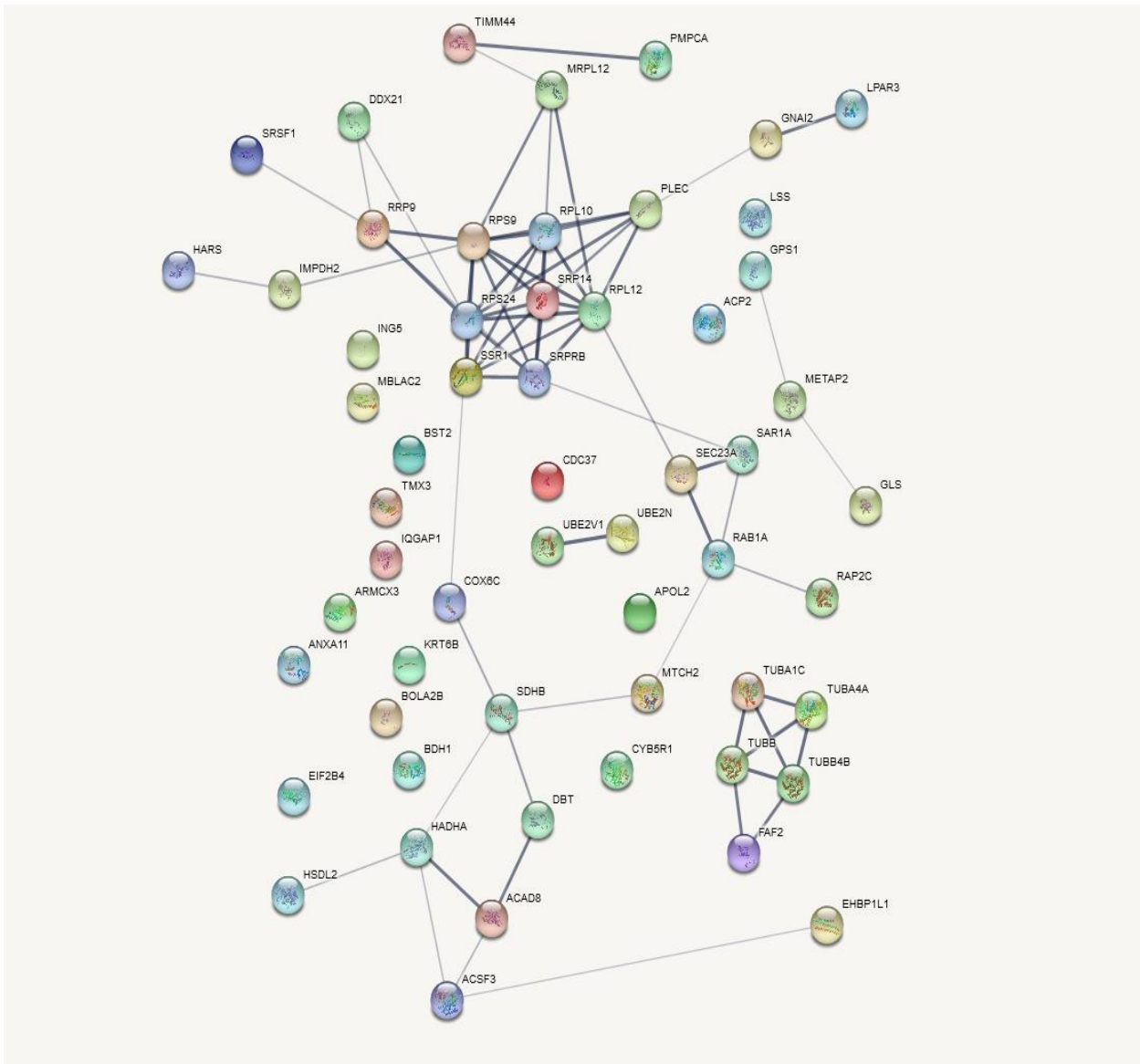
**Εικόνα 3.4** Παρουσιάζεται το ποσοστό των γονιδίων που συμμετέχουν στις δύο οικογένειες ως προς το σύνολο του δείγματος.



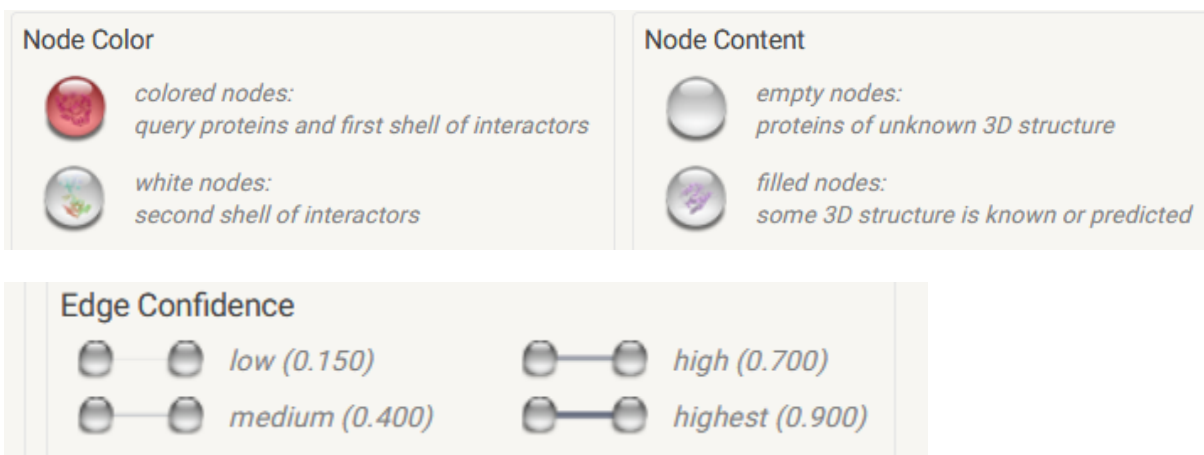
**Εικόνα 3.5** Το ραβδόγραμμα με τα ποσοστά συμμετοχής των γονιδίων μας στη συνολική γονιδιακή δεξαμενή της συγκεκριμένης υπερικογένειας και ως προς το σύνολο των γονιδίων του δείγματος αντιστοίχως.

### 3.4 STRING Αποτελέσματα

Συνεχίζοντας στο επίπεδο των λογισμικών οπτικοποίησης, το STRING έρχεται να προσφέρει την προοπτική των συνδέσεων μεταξύ των πρωτεϊνών ως κομμάτια μικρότερων δικτύων. Πιο συγκεκριμένα, ακολουθεί η εικόνα του STRING που δίνει μια αντίληψη των πρωτεϊνικών αλληλεπιδράσεων καθώς και τα αντίστοιχα legends:



**Εικόνα 3.6** Δίνεται μια αναπαράσταση των πιθανοτήτων οι πρωτεΐνες να συνδέονται μεταξύ τους σχηματίζοντας εσωτερικά δίκτυα πληροφορίας. Η ένταση των γραμμών αναπαριστά το score της πιθανότητας η συγκεκριμένη σύνδεση να ευσταθεί.



**Εικόνα 3.7** Επεξήγηση του περιεχομένου του κάθε κόμβου καθώς και βαθμονόμηση της πιθανότητας αληθείας των εικονιζόμενων σχέσεων που αποτυπώνεται ως διαφορετική ένταση του γκρι στις διαφορετικές συνδέσεις.

Από την εικόνα 3.6, παρατηρούμε πως εμφανίζονται κυρίως δύο δίκτυα που φαίνεται να χαρακτηρίζονται από υψηλά scores πρωτεϊνικών αλληλεπιδράσεων. Το πρώτο, εντοπίζεται κάτω δεξιά και περιλαμβάνει σχεδόν αποκλειστικά μέλη της οικογένειας των τουμπουλίνων, πρωτεϊνών οι οποίες φέρουν δομικό ρόλο στη σύνθεση των μικροσωληνίσκων. Η πυκνή αυτή σχέση δε θα πρέπει να δημιουργεί εντύπωση καθώς η δράση των τουμπουλίνων είναι σε άμεση σύνδεση αφού η τουμπουλίνη α και η τουμπουλίνη β συνδέονται μη ομοιοπολικά για να σχηματίσουν το πρωτοϊνίδιο. Ωστόσο δε φαίνεται να απαιτείται περαιτέρω διερεύνηση αυτού του δικτύου καθώς όλες οι πρωτεΐνες που περιλαμβάνει δεν ανήκουν στην κατηγορία με υψηλό score προς την κατεύθυνση του non responding που μας ενδιαφέρει. Το επόμενο δίκτυο εντοπίζεται έκκεντρα κυρίως. Αποτελείται από δέκα περίπου πρωτεΐνες οι οποίες ανήκουν σχεδόν αποκλειστικά στην κατηγορία RP και σχετίζονται με τη δομή των ριβοσωμάτων. Και σε αυτήν την περίπτωση αυτό ήταν αναμενόμενο καθώς διαδραματίζουν παρόμοιες δράσεις, ωστόσο και πάλι δε φαίνεται να έχουν κάποια ιδιαίτερη σημασία καθώς οι περισσότερες ανήκουν στην κατηγορία των deep responders πρωτεϊνών. Μια σύνδεση που θα μπορούσε να έχει κάποια αξία δυνητικά στο συγκεκριμένο δίκτυο είναι εκείνη μεταξύ των SRP14 και PLEC. Η SRP14 απαντάται μόνο σε non responders και η PLEC κυρίως σε non responders ενώ το score αλληλεπίδρασης τους είναι 0.758 και χαρακτηρίζεται ως υψηλό. Περαιτέρω ενδιαφέρον έχει το γεγονός πως η GNAI2 συνδέεται με την LPAR3 με πολύ υψηλό score ενώ και οι δύο ανήκουν στην κατηγορία των non responders. Επιπλέον, μεταξύ τους η PLEC και η GNAI2 συνδέονται με score 0.420 που χαρακτηρίζεται ως μέτριο αλλά είναι κοντά στο 50% και αν κάτι τέτοιο ευσταθεί τότε σίγουρα αξίζει να διερευνηθεί παραπάνω αφού αποτελεί το συνδυαστικό κρίκο για δύο συνδέσεις (PLEC-SRP14, GNAI2-LPAR3) αρκετά ισχυρές όπου αμφότερες αφορούν non responders αυξάνοντας τη σημασία τους ως δυνητικοί βιοδείκτες. Αυτό έγκειται στο ότι εάν υπάρχει μια αλυσίδα αλληλεπίδρασης που αφορά σε τέσσερις non responders πρωτεΐνες, πιθανώς να φέρουν κάποια χαρακτηριστικά που τους προσδίδουν λειτουργική σημασία στο πολλαπλό μείλωμα αλλά και αυξημένη αντίσταση στα φάρμακα εκλογής. Μια ακόμη σύνδεση που ίσως έχει σημασία είναι εκείνη μεταξύ των UBE2N και UBE2V1, δύο πρωτεΐνες

αποκλειστικά σε non responders που συνδέονται με score σχεδόν 1. Είναι πασιφανής η ισχύς της μεταξύ τους σχέσης και ίσως και αυτές και αυτές και εμφανίζουν κάποια καιρία χαρακτηριστικά για τη μελέτη μας σε περαιτέρω διερεύνηση. Οι περισσότερες από τις υπόλοιπες non responders πρωτεΐνες είτε δε συνδέονται με άλλες καθόλου είτε δεν εμφανίζουν ιδιαίτερα υψηλά scores αλληλεπιδράσεων ούτως ώστε να μπορεί να εξαχθεί κάποιο ασφαλές συμπέρασμα. Ακολουθούν οι δύο πίνακες του συνόλου των scores των πρωτεϊνικών αλληλεπιδράσεων:

**Πίνακας 3.2** Σελίδα 1 του πίνακα με scores των συνδέσεων που αποδόθηκαν στις 59 πρωτεΐνες μέσω της χρήσης του διαδικτυακού λογισμικού STRING.

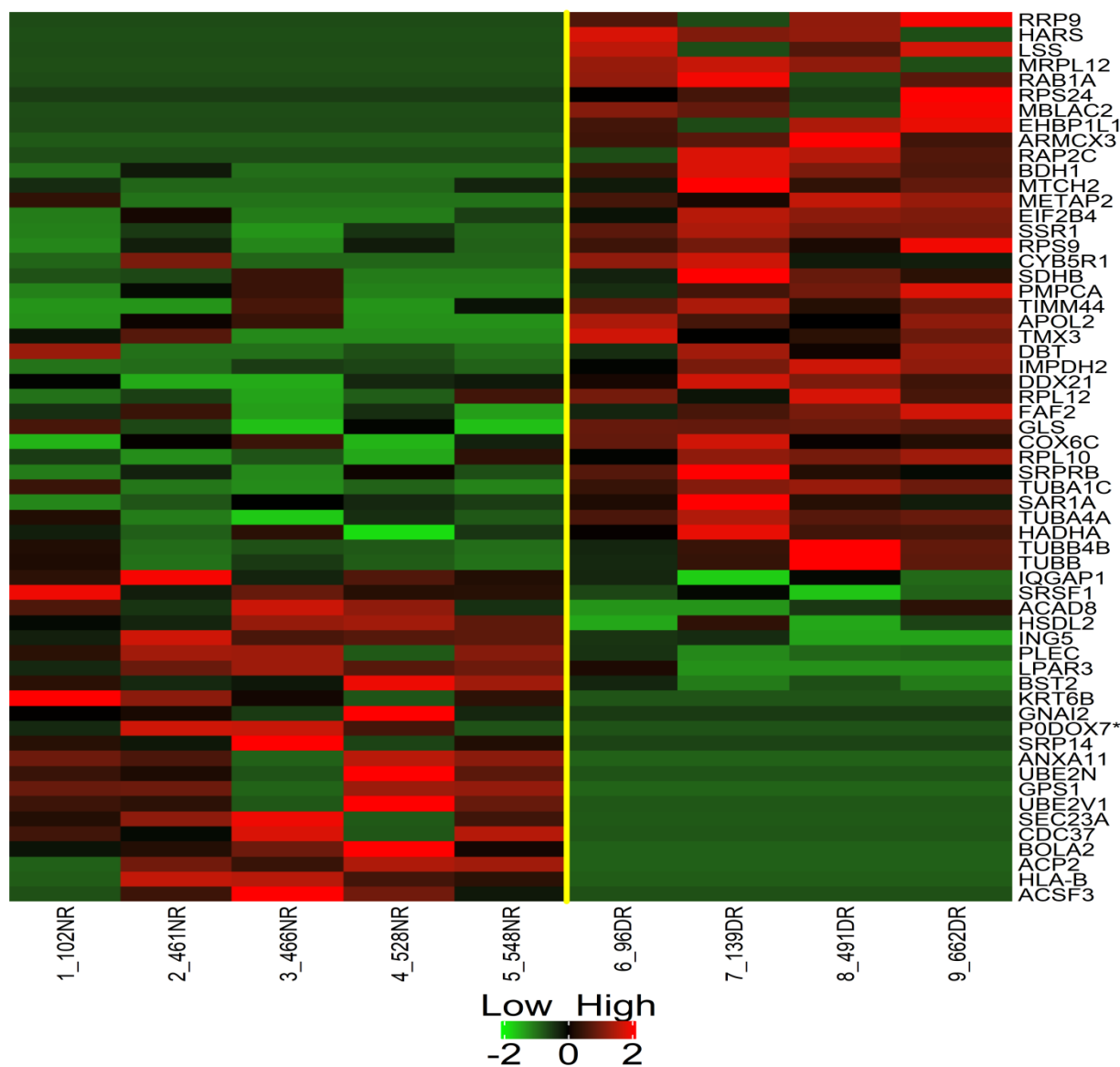
*node1	node2	node1_accession	node2_accession	node1_annotation	node2_annotation	score
ACAD8	ACSF3	ENSP00000281182	ENSP00000479130	Isobutyryl-CoA dehydrogenase...	Acyl-CoA synthetase family mem...	0.502
ACAD8	DBT	ENSP00000281182	ENSP00000359151	Isobutyryl-CoA dehydrogenase...	Dihydropolamide branched chal...	0.911
ACAD8	HADHA	ENSP00000281182	ENSP00000370023	Isobutyryl-CoA dehydrogenase...	Trifunctional enzyme subunit al...	0.931
ACSF3	ACAD8	ENSP00000479130	ENSP00000281182	Acyl-CoA synthetase family me...	Isobutyryl-CoA dehydrogenase...	0.502
ACSF3	EHBP1L1	ENSP00000479130	ENSP00000312671	Acyl-CoA synthetase family me...	EH domain-binding protein 1-lik...	0.409
ACSF3	HADHA	ENSP00000479130	ENSP00000370023	Acyl-CoA synthetase family me...	Trifunctional enzyme subunit al...	0.464
COX6C	RPS24	ENSP00000428895	ENSP00000414321	Cytochrome c oxidase subunit 6...	40S ribosomal protein S24; Requ...	0.413
COX6C	SDHB	ENSP00000428895	ENSP00000364649	Cytochrome c oxidase subunit 6...	Succinate dehydrogenase [ubiqui...	0.592
DBT	ACAD8	ENSP00000359151	ENSP00000281182	Dihydropolamide branched chal...	Isobutyryl-CoA dehydrogenase...	0.911
DBT	SDHB	ENSP00000359151	ENSP00000364649	Dihydropolamide branched chal...	Succinate dehydrogenase [ubiqui...	0.679
DDX21	RPS24	ENSP00000346120	ENSP00000414321	Nucleolar RNA helicase 2; RNA...	40S ribosomal protein S24; Requ...	0.489
DDX21	RRP9	ENSP00000346120	ENSP00000232888	Nucleolar RNA helicase 2; RNA...	US small nucleolar RNA-interacti...	0.502
EHBP1L1	ACSF3	ENSP00000312671	ENSP00000479130	EH domain-binding protein 1-lik...	Acyl-CoA synthetase family me...	0.409
FAF2	TUBB	ENSP00000261942	ENSP00000339001	FAS-associated factor 2; Plays a...	Tubulin beta chain; Tubulin is th...	0.900
FAF2	TUBB4B	ENSP00000261942	ENSP00000341289	FAS-associated factor 2; Plays a...	Tubulin beta-4B chain; Tubulin i...	0.900
GLS	METAP2	ENSP00000317379	ENSP00000325312	Glutaminase kidney isoform, mi...	Methionine aminopeptidase 2; C...	0.400
GNAI2	LPAR3	ENSP00000312999	ENSP00000395389	Guanine nucleotide-binding prot...	Lysophosphatidic acid receptor ...	0.945
GNAI2	PLEC	ENSP00000312999	ENSP00000323856	Guanine nucleotide-binding prot...	Plectin; Interlinks intermediate fi...	0.420
GPS1	METAP2	ENSP00000376167	ENSP00000325312	COP9 signalosome complex su...	Methionine aminopeptidase 2; C...	0.429
HADHA	ACAD8	ENSP00000370023	ENSP00000281182	Trifunctional enzyme subunit al...	Isobutyryl-CoA dehydrogenase...	0.931
HADHA	ACSF3	ENSP00000370023	ENSP00000479130	Trifunctional enzyme subunit alp...	Acyl-CoA synthetase family mem...	0.464
HADHA	HSDL2	ENSP00000370023	ENSP00000381785	Trifunctional enzyme subunit alp...	Hydroxysteroid dehydrogenase-li...	0.488
HADHA	SDHB	ENSP00000370023	ENSP00000364649	Trifunctional enzyme subunit alp...	Succinate dehydrogenase [ubiqui...	0.407
HARS	IMPDH2	ENSP00000425634	ENSP00000321584	Histidine-tRNA ligase, cytoplas...	Inosine-5'-monophosphate dehy...	0.505
HSDL2	HADHA	ENSP00000381785	ENSP00000370023	Hydroxysteroid dehydrogenase-li...	Trifunctional enzyme subunit alp...	0.488
IMPDH2	HARS	ENSP00000321584	ENSP00000425634	Inosine-5'-monophosphate dehy...	Histidine-tRNA ligase, cytoplas...	0.505
IMPDH2	RPS9	ENSP00000321584	ENSP00000302896	Inosine-5'-monophosphate dehy...	Ribosomal protein S9	0.551
LPAR3	GNAI2	ENSP00000395389	ENSP00000312999	Lysophosphatidic acid receptor 3...	Guanine nucleotide-binding prote...	0.945
METAP2	GLS	ENSP00000325312	ENSP00000317379	Methionine aminopeptidase 2; C...	Glutaminase kidney isoform, mi...	0.400
METAP2	GPS1	ENSP00000325312	ENSP00000376167	Methionine aminopeptidase 2; C...	COP9 signalosome complex sub...	0.429
MRPL12	RPL10	ENSP00000333837	ENSP00000413436	39S ribosomal protein L12, mitoc...	60S ribosomal protein L12; Comp...	0.658
MRPL12	RPL12	ENSP00000333837	ENSP00000354739	39S ribosomal protein L12, mitoc...	60S ribosomal protein L12; Binds...	0.900
MRPL12	RPS9	ENSP00000333837	ENSP00000302896	39S ribosomal protein L12, mitoc...	Ribosomal protein S9	0.837
MRPL12	TIMM44	ENSP00000333837	ENSP00000270538	39S ribosomal protein L12, mitoc...	Mitochondrial import inner mem...	0.461
MTCH2	RAB1A	ENSP00000303222	ENSP00000387286	Mitochondrial carrier homolog 2...	Ras-related protein Rab-1A; The s...	0.402
MTCH2	SDHB	ENSP00000303222	ENSP00000364649	Mitochondrial carrier homolog 2...	Succinate dehydrogenase [ubiqui...	0.469
PLEC	GNAI2	ENSP00000323856	ENSP00000312999	Plectin; Interlinks intermediate fil...	Guanine nucleotide-binding prote...	0.420
PLEC	RPL10	ENSP00000323856	ENSP00000413436	Plectin; Interlinks intermediate fil...	60S ribosomal protein L10; Comp...	0.834
PLEC	RPL12	ENSP00000323856	ENSP00000354739	Plectin; Interlinks intermediate fil...	60S ribosomal protein L12; Binds...	0.923
PLEC	RPS24	ENSP00000323856	ENSP00000414321	Plectin; Interlinks intermediate fil...	40S ribosomal protein S24; Requ...	0.919
PLEC	RPS9	ENSP00000323856	ENSP00000302896	Plectin; Interlinks intermediate fil...	Ribosomal protein S9	0.915
PLEC	SRP14	ENSP00000323856	ENSP00000267884	Plectin; Interlinks intermediate fil...	Signal recognition particle 14 kDa...	0.758
FMPCA	TIMM44	ENSP00000360782	ENSP00000270538	Mitochondrial-processing peptida...	Mitochondrial import inner memb...	0.925
RAB1A	MTCH2	ENSP00000387286	ENSP00000303222	Ras-related protein Rab-1A; The s...	Mitochondrial carrier homolog 2; ...	0.402
RAB1A	RAP2C	ENSP00000387286	ENSP00000340274	Ras-related protein Rab-1A; The s...	Ras-related protein Rap-2c; Small ...	0.557
RAB1A	SART1A	ENSP00000387286	ENSP00000362339	Ras-related protein Rab-1A; The s...	GTP-binding protein SART1a; Invol...	0.592
RAB1A	SEC23A	ENSP00000387286	ENSP00000306881	Ras-related protein Rab-1A; The s...	Protein transport protein Sec23A; ...	0.956
RAP2C	RAB1A	ENSP00000340274	ENSP00000387286	Ras-related protein Rap-2c; Small ...	Ras-related protein Rab-1A; The s...	0.557
RPL10	MRPL12	ENSP00000413436	ENSP00000333837	60S ribosomal protein L10; Comp...	39S ribosomal protein L12, mitoc...	0.658
RPL10	PLEC	ENSP00000413436	ENSP00000323856	60S ribosomal protein L10; Comp...	Plectin; Interlinks intermediate fil...	0.834
RPL10	RPL12	ENSP00000413436	ENSP00000354739	60S ribosomal protein L10; Comp...	60S ribosomal protein L12; Binds...	0.999
RPL10	RPS24	ENSP00000413436	ENSP00000414321	60S ribosomal protein L10; Comp...	40S ribosomal protein S24; Requ...	0.999
RPL10	RPS9	ENSP00000413436	ENSP00000302896	60S ribosomal protein L10; Comp...	Ribosomal protein S9	0.999
RPL10	SRP14	ENSP00000413436	ENSP00000267884	60S ribosomal protein L10; Comp...	Signal recognition particle 14 kDa...	0.915
RPL10	SRPRB	ENSP00000413436	ENSP00000418401	60S ribosomal protein L10; Comp...	Signal recognition particle recept...	0.904
RPL10	SSR1	ENSP00000413436	ENSP00000244763	60S ribosomal protein L10; Comp...	Translocon-associated protein su...	0.907
RPL12	MRPL12	ENSP00000354739	ENSP00000333837	60S ribosomal protein L12; Binds...	39S ribosomal protein L12, mitoc...	0.900
RPL12	PLEC	ENSP00000354739	ENSP00000323856	60S ribosomal protein L12; Binds...	Plectin; Interlinks intermediate fil...	0.923
RPL12	RPL10	ENSP00000354739	ENSP00000413436	60S ribosomal protein L12; Binds...	60S ribosomal protein L10; Comp...	0.999
RPL12	RPS24	ENSP00000354739	ENSP00000414321	60S ribosomal protein L12; Binds...	40S ribosomal protein S24; Requ...	0.999
RPL12	RPS9	ENSP00000354739	ENSP00000302896	60S ribosomal protein L12; Binds...	Ribosomal protein S9	0.999
RPL12	SEC23A	ENSP00000354739	ENSP00000306881	60S ribosomal protein L12; Binds...	Protein transport protein Sec23A; ...	0.557
RPL12	SRP14	ENSP00000354739	ENSP00000267884	60S ribosomal protein L12; Binds...	Signal recognition particle 14 kDa...	0.996
RPL12	SRPRB	ENSP00000354739	ENSP00000418401	60S ribosomal protein L12; Binds...	Signal recognition particle recept...	0.972
RPL12	SSR1	ENSP00000354739	ENSP00000244763	60S ribosomal protein L12; Binds...	Translocon-associated protein su...	0.904
RPS24	COX6C	ENSP00000414321	ENSP00000428895	40S ribosomal protein S24; Requir...	Cytochrome c oxidase subunit 6C...	0.413
RPS24	DDX21	ENSP00000414321	ENSP00000346120	40S ribosomal protein S24; Requir...	Nucleolar RNA helicase 2; RNA he...	0.489
RPS24	PLEC	ENSP00000414321	ENSP00000323856	40S ribosomal protein S24; Requir...	Plectin; Interlinks intermediate fil...	0.919
RPS24	RPL10	ENSP00000414321	ENSP00000413436	40S ribosomal protein S24; Requir...	60S ribosomal protein L10; Comp...	0.999
RPS24	RPL12	ENSP00000414321	ENSP00000354739	40S ribosomal protein S24; Requir...	60S ribosomal protein L12; Binds...	0.999
RPS24	RPS9	ENSP00000414321	ENSP00000302896	40S ribosomal protein S24; Requir...	Ribosomal protein S9	0.999
RPS24	RRP9	ENSP00000414321	ENSP00000232888	40S ribosomal protein S24; Requir...	US small nucleolar RNA-interacti...	0.979
RPS24	SRP14	ENSP00000414321	ENSP00000267884	40S ribosomal protein S24; Requir...	Signal recognition particle 14 kDa...	0.945
RPS24	SRPRB	ENSP00000414321	ENSP00000418401	40S ribosomal protein S24; Requir...	Signal recognition particle recept...	0.900
RPS24	SSR1	ENSP00000414321	ENSP00000244763	40S ribosomal protein S24; Requir...	Translocon-associated protein su...	0.902
RPS9	IMPDH2	ENSP00000302896	ENSP00000321584	Ribosomal protein S9	Inosine-5'-monophosphate dehydr...	0.551
RPS9	MRPL12	ENSP00000302896	ENSP00000333837	Ribosomal protein S9	39S ribosomal protein L12, mitoc...	0.837
RPS9	PLEC	ENSP00000302896	ENSP00000323856	Ribosomal protein S9	Plectin; Interlinks intermediate fil...	0.915
RPS9	RPL10	ENSP00000302896	ENSP00000413436	Ribosomal protein S9	60S ribosomal protein L10; Comp...	0.999
RPS9	RPL12	ENSP00000302896	ENSP00000354739	Ribosomal protein S9	60S ribosomal protein L12; Binds...	0.999

**Πίνακας 3.3** Σελίδα 2 του πίνακα με scores των συνδέσεων που αποδόθηκαν στις 59 πρωτεΐνες μέσω της χρήσης του διαδικτυακού λογισμικού STRING.

RPS9	RPS24	ENSP00000302896	ENSP00000414321	Ribosomal protein S9	40S ribosomal protein S24; Requ...	0.999
RPS9	RRP9	ENSP00000302896	ENSP00000232888	Ribosomal protein S9	U3 small nucleolar RNA-interactin...	0.985
RPS9	SRP14	ENSP00000302896	ENSP00000267884	Ribosomal protein S9	Signal recognition particle 14 kDa...	0.996
RPS9	SRPRB	ENSP00000302896	ENSP00000418401	Ribosomal protein S9	Signal recognition particle recept...	0.900
RPS9	SSR1	ENSP00000302896	ENSP00000244763	Ribosomal protein S9	Translocon-associated protein su...	0.908
RRP9	DDX21	ENSP00000232888	ENSP00000346120	U3 small nucleolar RNA-interactin...	Nucleolar RNA helicase 2; RNA h...	0.502
RRP9	RPS24	ENSP00000232888	ENSP00000414321	U3 small nucleolar RNA-interactin...	40S ribosomal protein S24; Requ...	0.979
RRP9	RPS9	ENSP00000232888	ENSP00000302896	U3 small nucleolar RNA-interactin...	Ribosomal protein S9	0.985
RRP9	SRPRB	ENSP00000232888	ENSP00000418401	U3 small nucleolar RNA-interactin...	Signal recognition particle recept...	0.451
RRP9	SRSF1	ENSP00000232888	ENSP00000258962	U3 small nucleolar RNA-interactin...	Serine/arginine-rich splicing facto...	0.474
SAR1A	RAB1A	ENSP00000362339	ENSP00000387286	GTP-binding protein SAR1a; Invol...	Ras-related protein Rab-1A; The s...	0.592
SAR1A	SEC23A	ENSP00000362339	ENSP00000306881	GTP-binding protein SAR1a; Invol...	Protein transport protein Sec23A...	0.994
SAR1A	SRPRB	ENSP00000362339	ENSP00000418401	GTP-binding protein SAR1a; Invol...	Signal recognition particle recept...	0.477
SDHB	COX6C	ENSP00000364649	ENSP00000428955	Succinate dehydrogenase [ubiqui...	Cytochrome c oxidase subunit 6C...	0.592
SDHB	DBT	ENSP00000364649	ENSP00000358191	Succinate dehydrogenase [ubiqui...	Dihydropyrimidine branched chain...	0.679
SDHB	HADHA	ENSP00000364649	ENSP00000370023	Succinate dehydrogenase [ubiqui...	Trifunctional enzyme subunit alp...	0.407
SDHB	MTCH2	ENSP00000364649	ENSP00000303222	Succinate dehydrogenase [ubiqui...	Mitochondrial carrier homolog 2...	0.469
SEC23A	RAB1A	ENSP00000306881	ENSP00000387286	Protein transport protein Sec23A...	Ras-related protein Rab-1A; The s...	0.956
SEC23A	RPL12	ENSP00000306881	ENSP00000354739	Protein transport protein Sec23A...	60S ribosomal protein L12; Binds...	0.557
SEC23A	SAR1A	ENSP00000306881	ENSP00000362339	Protein transport protein Sec23A...	GTP-binding protein SAR1a; Invol...	0.994
SRP14	PLEC	ENSP00000267884	ENSP00000323856	Signal recognition particle 14 kDa...	Plectin; Interlinks intermediate fila...	0.758
SRP14	RPL10	ENSP00000267884	ENSP00000413436	Signal recognition particle 14 kDa...	60S ribosomal protein L10; Compo...	0.915
SRP14	RPL12	ENSP00000267884	ENSP00000354739	Signal recognition particle 14 kDa...	60S ribosomal protein L12; Binds d...	0.996
SRP14	RPS24	ENSP00000267884	ENSP00000414321	Signal recognition particle 14 kDa...	40S ribosomal protein S24; Requ...	0.945
SRP14	RPS9	ENSP00000267884	ENSP00000302896	Signal recognition particle 14 kDa...	Ribosomal protein S9	0.996
SRP14	SRPRB	ENSP00000267884	ENSP00000418401	Signal recognition particle 14 kDa...	Signal recognition particle recepto...	0.963
SRP14	SSR1	ENSP00000267884	ENSP00000244763	Signal recognition particle 14 kDa...	Translocon-associated protein sub...	0.918
SRPRB	RPL10	ENSP00000418401	ENSP00000413436	Signal recognition particle recepto...	60S ribosomal protein L10; Compo...	0.904
SRPRB	RPL12	ENSP00000418401	ENSP00000354739	Signal recognition particle recepto...	60S ribosomal protein L12; Binds d...	0.972
SRPRB	RPS24	ENSP00000418401	ENSP00000414321	Signal recognition particle recepto...	40S ribosomal protein S24; Requ...	0.900
SRPRB	RPS9	ENSP00000418401	ENSP00000302896	Signal recognition particle recepto...	Ribosomal protein S9	0.900
SRPRB	RRP9	ENSP00000418401	ENSP00000232888	Signal recognition particle recepto...	U3 small nucleolar RNA-interacting...	0.451
SRPRB	SAR1A	ENSP00000418401	ENSP00000362339	Signal recognition particle recepto...	GTP-binding protein SAR1a; Involv...	0.477
SRPRB	SRP14	ENSP00000418401	ENSP00000267884	Signal recognition particle recepto...	Signal recognition particle 14 kDa...	0.963
SRPRB	SSR1	ENSP00000418401	ENSP00000244763	Signal recognition particle recepto...	Translocon-associated protein sub...	0.964
SRSF1	RRP9	ENSP00000258962	ENSP00000232888	Serine/arginine-rich splicing factor...	U3 small nucleolar RNA-interacting...	0.474
SSR1	RPL10	ENSP00000244763	ENSP00000413436	Translocon-associated protein sub...	60S ribosomal protein L10; Compo...	0.907
SSR1	RPL12	ENSP00000244763	ENSP00000354739	Translocon-associated protein sub...	60S ribosomal protein L12; Binds d...	0.904
SSR1	RPS24	ENSP00000244763	ENSP00000414321	Translocon-associated protein sub...	40S ribosomal protein S24; Requir...	0.902
SSR1	RPS9	ENSP00000244763	ENSP00000302896	Translocon-associated protein sub...	Ribosomal protein S9	0.908
SSR1	SRP14	ENSP00000244763	ENSP00000267884	Translocon-associated protein s...	Signal recognition particle 14 kD...	0.918
SSR1	SRPRB	ENSP00000244763	ENSP00000418401	Translocon-associated protein s...	Signal recognition particle recept...	0.964
TIMM44	MRPL12	ENSP00000270538	ENSP00000333837	Mitochondrial import inner mem...	39S ribosomal protein L12, mitoc...	0.461
TIMM44	PMPCA	ENSP00000270538	ENSP00000360782	Mitochondrial import inner mem...	Mitochondrial-processing peptid...	0.925
TUBA1C	TUBA4A	ENSP00000301072	ENSP00000301072	Tubulin alpha-1C chain; Tubulin I...	Tubulin alpha-4A chain; Tubulin I...	0.950
TUBA1C	TUBB	ENSP00000301072	ENSP00000339001	Tubulin alpha-1C chain; Tubulin I...	Tubulin beta chain; Tubulin is the...	0.978
TUBA1C	TUBB4B	ENSP00000301072	ENSP00000341289	Tubulin alpha-1C chain; Tubulin I...	Tubulin beta-4B chain; Tubulin is...	0.978
TUBA4A	TUBA1C	ENSP00000248437	ENSP00000301072	Tubulin alpha-4A chain; Tubulin I...	Tubulin alpha-1C chain; Tubulin I...	0.950
TUBA4A	TUBB	ENSP00000248437	ENSP00000339001	Tubulin alpha-4A chain; Tubulin I...	Tubulin beta chain; Tubulin is the...	0.990
TUBA4A	TUBB4B	ENSP00000248437	ENSP00000341289	Tubulin alpha-4A chain; Tubulin I...	Tubulin beta-4B chain; Tubulin is...	0.989
TUBB	FAF2	ENSP00000339001	ENSP00000261942	Tubulin beta chain; Tubulin is the...	FAS-associated factor 2; Plays an...	0.900
TUBB	TUBA1C	ENSP00000339001	ENSP00000301072	Tubulin beta chain; Tubulin is the...	Tubulin alpha-1C chain; Tubulin I...	0.978
TUBB	TUBA4A	ENSP00000339001	ENSP00000248437	Tubulin beta chain; Tubulin is the...	Tubulin alpha-4A chain; Tubulin I...	0.990
TUBB	TUBB4B	ENSP00000339001	ENSP00000341289	Tubulin beta chain; Tubulin is the...	Tubulin beta-4B chain; Tubulin is...	0.998
TUBB4B	FAF2	ENSP00000341289	ENSP00000261942	Tubulin beta-4B chain; Tubulin is...	FAS-associated factor 2; Plays an...	0.900
TUBB4B	TUBA1C	ENSP00000341289	ENSP00000301072	Tubulin beta-4B chain; Tubulin is...	Tubulin alpha-1C chain; Tubulin I...	0.978
TUBB4B	TUBA4A	ENSP00000341289	ENSP00000248437	Tubulin beta-4B chain; Tubulin is...	Tubulin alpha-4A chain; Tubulin I...	0.989
TUBB4B	TUBB	ENSP00000341289	ENSP00000339001	Tubulin beta-4B chain; Tubulin is...	Tubulin beta chain; Tubulin is the...	0.998
UBE2N	UBE2V1	ENSP00000316176	ENSP00000340305	Ubiquitin-conjugating enzyme E2...	Ubiquitin-conjugating enzyme E2...	0.998
UBE2V1	UBE2N	ENSP00000316176	ENSP00000340305	Ubiquitin-conjugating enzyme E2...	Ubiquitin-conjugating enzyme E2...	0.998
SSR1	SRP14	ENSP00000244763	ENSP00000267884	Translocon-associated protein s...	Signal recognition particle 14 kD...	0.918
SSR1	SRPRB	ENSP00000244763	ENSP00000418401	Translocon-associated protein s...	Signal recognition particle recept...	0.964
TIMM44	MRPL12	ENSP00000270538	ENSP00000333837	Mitochondrial import inner mem...	39S ribosomal protein L12, mitoc...	0.461
TIMM44	PMPCA	ENSP00000270538	ENSP00000360782	Mitochondrial import inner mem...	Mitochondrial-processing peptid...	0.925
TUBA1C	TUBA4A	ENSP00000301072	ENSP00000301072	Tubulin alpha-1C chain; Tubulin I...	Tubulin alpha-4A chain; Tubulin I...	0.950
TUBA1C	TUBB	ENSP00000301072	ENSP00000339001	Tubulin alpha-1C chain; Tubulin I...	Tubulin beta chain; Tubulin is the...	0.978
TUBA1C	TUBB4B	ENSP00000301072	ENSP00000341289	Tubulin alpha-1C chain; Tubulin I...	Tubulin beta-4B chain; Tubulin is...	0.978
TUBA4A	TUBA1C	ENSP00000248437	ENSP00000301072	Tubulin alpha-4A chain; Tubulin I...	Tubulin alpha-1C chain; Tubulin I...	0.950
TUBA4A	TUBB	ENSP00000248437	ENSP00000339001	Tubulin alpha-4A chain; Tubulin I...	Tubulin beta chain; Tubulin is the...	0.990
TUBA4A	TUBB4B	ENSP00000248437	ENSP00000341289	Tubulin alpha-4A chain; Tubulin I...	Tubulin beta-4B chain; Tubulin is...	0.989
TUBB	FAF2	ENSP00000339001	ENSP00000261942	Tubulin beta chain; Tubulin is the...	FAS-associated factor 2; Plays an...	0.900
TUBB	TUBA1C	ENSP00000339001	ENSP00000301072	Tubulin beta chain; Tubulin is the...	Tubulin alpha-1C chain; Tubulin I...	0.978
TUBB	TUBA4A	ENSP00000339001	ENSP00000248437	Tubulin beta chain; Tubulin is the...	Tubulin alpha-4A chain; Tubulin I...	0.990
TUBB	TUBB4B	ENSP00000339001	ENSP00000341289	Tubulin beta chain; Tubulin is the...	Tubulin beta-4B chain; Tubulin is...	0.998
TUBB4B	FAF2	ENSP00000341289	ENSP00000261942	Tubulin beta-4B chain; Tubulin is...	FAS-associated factor 2; Plays an...	0.900
TUBB4B	TUBA1C	ENSP00000341289	ENSP00000301072	Tubulin beta-4B chain; Tubulin is...	Tubulin alpha-1C chain; Tubulin I...	0.978
TUBB4B	TUBA4A	ENSP00000341289	ENSP00000248437	Tubulin beta-4B chain; Tubulin is...	Tubulin alpha-4A chain; Tubulin I...	0.989
TUBB4B	TUBB	ENSP00000341289	ENSP00000339001	Tubulin beta-4B chain; Tubulin is...	Tubulin beta chain; Tubulin is the...	0.998
UBE2N	UBE2V1	ENSP00000316176	ENSP00000340305	Ubiquitin-conjugating enzyme E2...	Ubiquitin-conjugating enzyme E2...	0.998
UBE2V1	UBE2N	ENSP00000316176	ENSP00000340305	Ubiquitin-conjugating enzyme E2...	Ubiquitin-conjugating enzyme E2...	0.998

### 3.5 Heatmap Αποτελέσματα

Η ανάλυση Heatmap αφορά κυρίως σε οπτικοποίηση των ποσοτικών δεδομένων συγκέντρωσης των πρωτεϊνών του δείγματος και φέρει διπλή σημασία καθώς βοηθά στην εύκολη πλοήγηση τόσο όσον αφορά στις συγκεντρώσεις της εκάστοτε πρωτεΐνης όσο και στο προφίλ αυτών των συγκεντρώσεων ως προς τον κάθε ασθενή. Ειδικότερα έχουμε:



**Εικόνα 3.8** Heatmap παρουσίαση της συγκέντρωσης των πρωτεϊνών στα διαφορετικά δείγματα με αποχρώσεις που εκτείνονται από το πράσινο(χαμηλή συγκέντρωση) μέχρι το κόκκινο.(υψηλή συγκέντρωση)

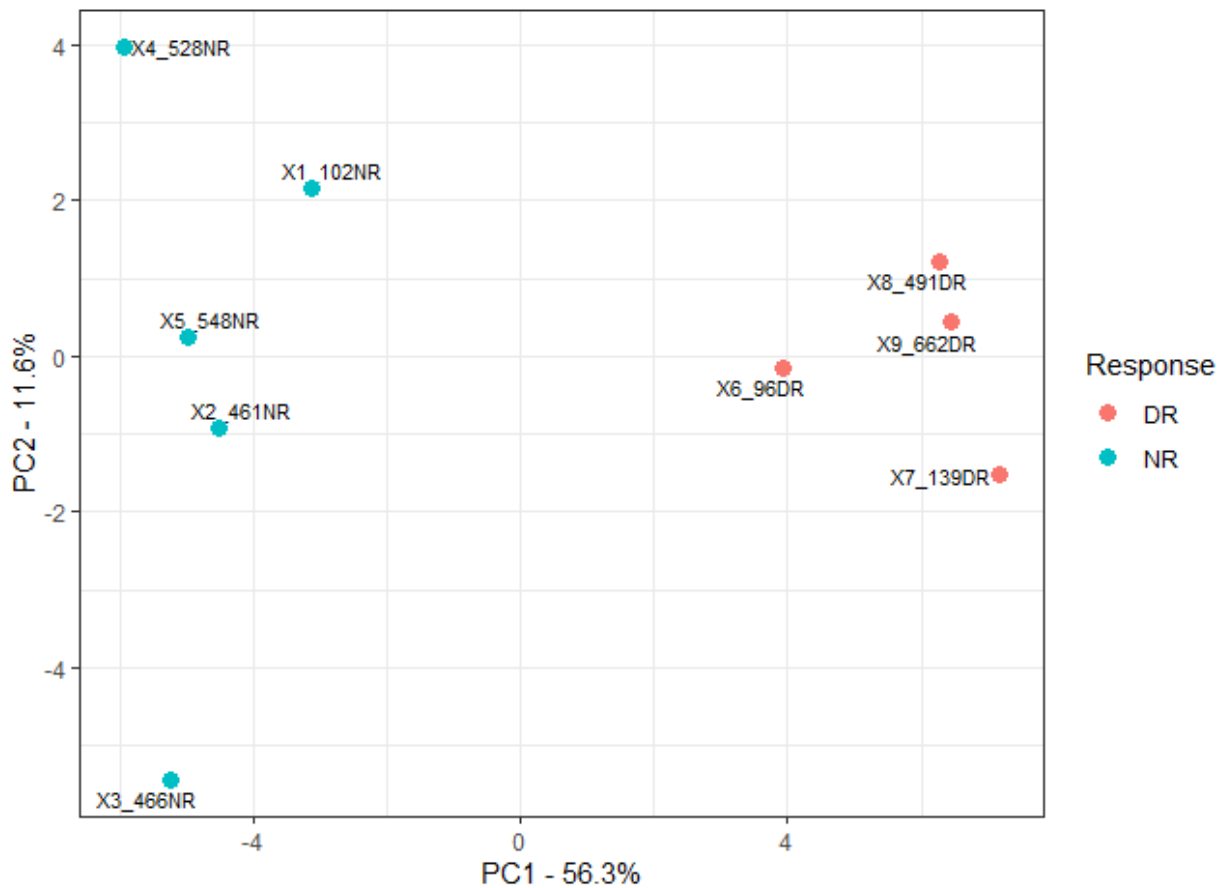


Όπως βλέπουμε το Heatmap είναι χωρισμένο με μια κίτρινη γραμμή σε δύο μέρη με το αριστερό να αφορά στους non responders και το δεξιό τους deep responders. Από την εικόνα μπορούμε να καταλάβουμε πως υπάρχει μια σχέση αναλογίας χιαστί μεταξύ των συγκεντρώσεων των πρωτεϊνών των deep responders και των non responders. Επιπλέον, είναι εύκολα παρατηρήσιμο πως γενικά οι συγκεντρώσεις των πρωτεϊνών εμφανίζουν παρόμοια επίπεδα πλην εξαιρέσεων μεταξύ των deep responders και αντίστοιχα παρόμοια επίπεδα μεταξύ των non responders. Η σχέση σχεδόν χιαστί που βλέπουμε υποδεικνύει με σαφήνεια πως η περιοχή με αξία αναζήτησης βιοδεικτών είναι η κάτω αριστερά και δικαιολογεί οπτικά τα δεδομένα του αρχικού dataset.

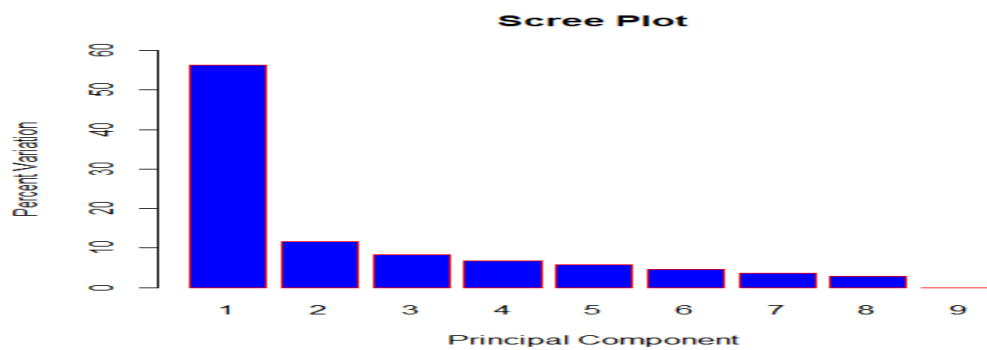
### 3.6 PCA Αποτελέσματα

Σε αντίθεση με τις προηγούμενες αναλύσεις που χρησιμοποίησαν κυρίως ως δεδομένα τα ονόματα των πρωτεϊνών και τις συγκεντρώσεις τους για να βγάλουν συμπέρασμα, το PCA αποτελεί μια αρκετά πιο περίπλοκη ανάλυση που ως τελικό στόχο έχει να αξιολογήσει τις πρωτεΐνες ως προς τη στατιστική τους αξία στην εμφάνιση της ασθένειας και κατ'έπекταση τη συνεισφορά τους ως δυνητικοί βιοδείκτες. Για να γίνει αυτό, αρχικά λαμβάνονται από το λογισμικό τόσα PCAs όσα και το σύνολο των ασθενών μας, δηλαδή 9. Ως PCAs ορίζουμε την ευθεία γραμμή που ενώνει τα δύο σημεία που απέχουν τη μεγαλύτερη απόσταση μεταξύ τους και στη συνέχεια όλα τα υπόλοιπα σημεία αξιολογούνται με βάση την απόστασή τους από τον άξονα που ορίζει αυτή η ευθεία. Για να γίνει αυτό περισσότερο αντιληπτό, εάν έχουμε 9 ασθενείς και από τον καθένα έχουμε δείγμα 59 πρωτεϊνών, τότε σε ένα διάγραμμα με συντεταγμένες τις συγκεντρώσεις των πρωτεϊνών και το δείγμα στο οποίο ανήκουν, θα είχαμε  $59 \times 9 = 531$  διαφορετικά σημεία. Ως PC1 ορίζεται η απόσταση μεταξύ των δύο σημείων από τα 531 που απέχουν τη μέγιστη απόσταση. Στη συνέχεια, όλα τα υπόλοιπα σημεία κρίνονται ως προς τη συνεισφορά τους στη διακύμανση του PC1 δηλαδή του πόσο κοντά ή μακριά είναι στα δύο αυτά ακριανά σημεία. Όσο πιο κοντά είναι τόσο μεγαλύτερο score συνεισφοράς έχουν ως προς PC1 και το αντίστροφο. Ακολούθως, η PC2 είναι η δεύτερη διακύμανση η οποία ορίζεται ως μια κάθετη στην PC1 ευθεία και αυτή τη φορά χαράσσεται μεταξύ των δύο σημείων που απέχουν τη μέγιστη απόσταση μεταξύ τους αλλά βρίσκονται σε διεύθυνση κάθετη σε αυτήν της PC1. Κατά αυτόν τον τρόπο, δημιουργείται ένα σύστημα αξόνων  $x, y$  του οποίου ως  $x$  ορίζεται η συνεισφορά της διακύμανσης από την PC1 και ως  $y$  η συνεισφορά της διακύμανσης ως προς PC2. Η διαδικασία αυτή συνεχίζεται και για τα υπόλοιπα 7 PCAs, χαράσσοντας κάθε φορά την αμέσως επόμενη μεγαλύτερη απόσταση αλλά σε διαφορετικό επίπεδο του αρχικού. Με αυτόν τον τρόπο καταφέρνουμε να μεταφερόμαστε από ένα διάγραμμα που θα περιελάμβανε 531 σημεία σε ένα άλλο με μόλις 9, το οποίο διαχωρίζει τους ασθενείς μας σε deep responders και non responders ενώ ταυτόχρονα τους αξιολογεί ως προς τη στατιστική αξία που προσφέρει η συνεισφορά τους ως προς PC1 και PC2. Είναι εμφανές πως  $PC1 > PC2 > PC3$  και ούτω καθεξής. Εκτός των παραπάνω, η PCA ανάλυση προσφέρει και scores για την κάθε πρωτεΐνη ξεχωριστά ανάλογα με τη συνεισφορά της κυρίως ως προς PC1 και PC2 όπου μετουσιώνεται σε στατιστική σημαντικότητα των πρωτεϊνών αυτών. Όσο μεγαλύτερα είναι αυτά τα scores, τόσο πιο κοντά είναι στις ακραίες τιμές των PCAs και τόσο μεγαλύτερη στατιστική αξία έχουν ως

πιθανοί βιοδείκτες. Ακολουθούν τα διαγράμματα που αφορούν στην ανάλυση PCA αλλά και τα scores αυτών:



**Εικόνα 3.9** Παρουσίαση του PCA διαγράμματος στο οποίο ως άξονες x,y χρησιμοποιούνται οι PC1 και PC2 διακυμάνσεις. Τα ποσοστά που αναγράφονται δίπλα στους άξονες αφορούν στο ποσοστό που καταλαμβάνει η διακύμανση PC1 και PC2 αντίστοιχα ως προς το σύνολο των 9 διακυμάνσεων. Αριστερά του 0 είναι όλοι οι non responders και δεξιά οι deep responders.



**Εικόνα 3.10** Παρουσίαση των ποσοστών των 9 PCA διακυμάνσεων που χρησιμοποιήθηκαν προκειμένου να αποδοθούν τα scores αξιολόγησης της στατιστικής σημασίας όλων των πρωτεϊνών.



Από όλα τα παραπάνω μπορούμε να βγάλουμε κάποια σημαντικά συμπεράσματα. Αρχικά κοιτώντας την εικόνα 3.9 η οποία δείχνει το scree plot, μπορούμε να καταλάβουμε πως το λογισμικό υπολόγισε τα scores του κυρίως με βάση το PC1 που λαμβάνει ποσοστό 56.3% των συνολικών διακυμάνσεων. Σαν δεύτερος άξονας χρησιμοποιήθηκε ο PC2 ως η επόμενη μεγαλύτερη διακύμανση μετά την PC1. Από την εικόνα 3.8, παρατηρούμε πως υπάρχουν συνολικά 9 σημεία συντεταγμένων, όσος είναι και ο αριθμός των αρχικών μας δειγμάτων. Επομένως το PCA διενήργησε μελέτη στατιστικής σημαντικότητας η οποία προκύπτει από την τοποθέτηση όχι και των 59 πρωτεϊνών του κάθε δείγματος αλλά με την αποτύπωση ενός score με βάση τη συνεισφορά της κάθε πρωτεΐνης(που εξαρτάται από τη συγκέντρωσή της και την απόσταση από τη διακύμανση PC1 και PC2 αντιστοίχως) στο κάθε δείγμα ξεχωριστά. Με άλλα λόγια, επιτελέστηκε μείωση διαστάσεων (dimension reduction) από το απλό διάγραμμα με άξονες την συγκέντρωση πρωτεϊνών και τον αριθμό των δειγμάτων σε τελικό διάγραμμα που παρουσιάζει του ασθενείς χωρίς τις πρωτεΐνες τους αλλά με ταυτόχρονη διατήρηση της στατιστικής σημασίας των προφίλ συγκέντρωσης της κάθε πρωτεΐνης σε αυτούς. Η στατιστική σημαντικότητα μετράται ως η μεγαλύτερη απόλυτη τιμή, δηλαδή η μεγαλύτερη απόσταση από το 0. Επειδή κάθε PCA άξονας ορίζεται από δύο σημεία, αυτό είναι απολύτως λογικό. Αριστερά του 0 βλέπουμε ότι τοποθετούνται οι non responders και δεξιά οι deep responders. Αυτό πρακτικά σημαίνει πως και οι πρωτεΐνες που αφορούν τους non responders αναμένουμε να έχουν αρνητικό πρόσημο ενώ εκείνες που αφορούν τους deep responders θετικό, κάτι που επιβεβαιώνεται από τον πίνακα 3.4. Μελετώντας την εικόνα 3.8 δεν μπορούμε όμως να κατανοήσουμε με σαφήνεια τη συνεισφορά των πρωτεϊνών στο PCA παρά μόνο να αποκτήσουμε μια αδρή εικόνα για τη θέση που τοποθετούνται οι deep responders και οι non responders με βάση την ανάλυση PCA. Για να υπεισέλθουμε σε μεγαλύτερο βάθος στο θέμα, μεγάλη σημασία εμφανίζει ο πίνακας 3.4 που περιλαμβάνει τα PCA scores και θα μας επιτρέψουν να εξετάσουμε τη στατιστική σημασία της κάθε πρωτεΐνης στις περιπτώσεις πολλαπλού μυελώματος. Από τις 59 πρωτεΐνες, οι πιο σημαντικές στατιστικά είναι εκείνες που έχουν τη μεγαλύτερη απόλυτη τιμή, δηλαδή οι δέκα πρώτες και οι δέκα τελευταίες του πίνακα. Ωστόσο, οι δέκα πρώτες δεν είναι σημαντικές για την έρευνά μας διότι αφορούν σε συγκεντρώσεις πρωτεϊνών οι οποίες βρέθηκαν σε deep responders και άρα ανταποκρινόμενους στη θεραπεία ασθενείς και επομένως δεν έχουν πιθανότατα σχέση με την ασθένεια. Άρα οι πρωτεΐνες LPAR3, ING5, PPAL, BOLA2, GPS1, HLA-B, BST2, ANX11, PLEC, UB2V1 είναι εκείνες με τα υψηλότερα κατά απόλυτη τιμή scores που αφορούν σε non responders. Εδώ παρατηρείται κάτι πολύ ενδιαφέρον. Συγκρίνοντας αυτά τα αποτελέσματα σε σχέση με το αρχικό dataset, βλέπουμε πως η πρωτεΐνη LPAR3, η οποία δεν ανιχνεύεται μόνο στους non responders και θα ήταν αναμενόμενο πρωτεΐνες που απαντώνται αποκλειστικά στους non responders να έχουν μεγαλύτερη στατιστική αξία έναντι εκείνων που ανιχνεύονται τόσο στους non responders όσο και στους deep responders, να έχει το μεγαλύτερο score από όλες. Επίσης αυτό ισχύει και για άλλες πρωτεΐνες όπως η πλεκτίνη η οποία φέρει score μεγαλύτερο από άλλες που συναντώνται μόνο σε non responders. Αυτό ερμηνεύεται από το γεγονός πως το PCA λαμβάνει υπόψιν όχι μόνο τις κατανομές συγκεντρώσεων αλλά και τη διακύμανσή τους ως προς τις κεντρικές γραμμές διακύμανσης PC1 και PC2. Συνεπώς, όλες αυτές πρωτεΐνες φέρουν τη μεγαλύτερη στατιστική αξία από τις υπόλοιπες 59 κατά τη φθίνουσα σειρά με την οποία ονομάστηκαν και κατά συνέπεια έχουν τη μεγαλύτερη αξία ως δυνητικοί βιοδείκτες.

## ΚΕΦΑΛΑΙΟ 4

## ΣΥΜΠΕΡΑΣΜΑΤΑ

Μέσα από τα διαφορετικά είδη αναλύσεων που επιτελέσαμε, καταφέραμε να αποκτήσουμε μια καλύτερη εικόνα για τις πρωτεΐνες του δείγματός μας. Κύριος στόχος ήταν η οπτικοποίηση, η συσχέτιση αλλά και η εξαγωγή της στατιστικής αξίας της κάθε πρωτεΐνης ούτως ώστε να αποφασίσουμε για το ποιες από αυτές θα μπορούσε να ήταν χρήσιμο να μελετήσουμε ως πιθανούς βιοδείκτες. Παράλληλα, καταφέραμε να πραγματοποιήσουμε και κάποιας μορφής κανονικοποίησης των δεδομένων μας, γεγονός που μας οδήγησε σε περισσότερο κατανοητά συμπεράσματα σε σχέση με τη άγουρη και ανεπεξέργαστη μορφή της πληροφορίας του αρχικού dataset. Μέσω του Cluego, κατατάξαμε τις πρωτεΐνες στις κύριες οικογένειες που ανήκουν με βάση το γονιδιακό τους υπόβαθρο ενώ καταλήξαμε πως η πλειοψηφία σχετίζεται με το σύστημα μεταφοράς κυστίδιων. Αξίζει να αναφερθεί σε αυτό το σημείο πως η ανάλυση μέσω Cluego είναι εξαιρετικά δύσκολη σε ένα τόσο μικρό δείγμα και τα αποτελέσματα μας δεν επαρκούν για να είμαστε σίγουροι με απόλυτη ακρίβεια καθώς ο μικρός αριθμός πρωτεϊνών δεν επιτρέπει την εύκολη ομαδοποίησή τους. Όσον αφορά στο πρόγραμμα STRING και σε αυτήν την περίπτωση το μικρό δείγμα φέρει μεγάλους περιορισμούς ως προς τις δυνατότητες επεξεργασίας μιας και όσο μικρότερος είναι ο αριθμός των πρωτεϊνών τόσο μικρότερη και η πιθανότητα να αλληλεπιδρούν μεταξύ τους εκτός αν αφορά σε ένα πολύ συγκεκριμένο δείγμα που φέρει πρωτεΐνες με υψηλή συγγένεια. Παρόλα αυτά, καταφέραμε να προτείνουμε κυρίως δύο οδούς που θα μπορούσαν να διερευνηθούν περαιτέρω για την ανεύρεση βιοδεικτών λόγω των υψηλών scores πιθανότητας των σχέσεων που συνδέουν αυτές τις non responders πρωτεΐνες. Από εκεί και ύστερα, μέσω της Heatmap ανάλυσης, αποκτήσαμε οπτική πληροφορία για τη συγκέντρωση των πρωτεϊνών του δείγματος, καταφέροντας μάλιστα να διαχωρίσουμε τους deep responders και τους non responders με αποκλειστικά και μόνο αυτήν την παράμετρο. Ταυτόχρονα, το Heatmap επιτρέπει την εύκολη πλοήγηση του αναγνώστη ως προς το προφίλ των πρωτεϊνών μας, είτε κοιτώντας μεμονωμένα τη συγκέντρωση της κάθε πρωτεΐνης στο εκάστοτε δείγμα είτε μέσω της οπτικοποίησης της περιοχής των πρωτεϊνών που εμφανίζονται κυρίως στους non responders και επομένως είναι καλοί υποψήφιοι βιοδείκτες. Τέλος, η PCA ανάλυση διετελέστηκε με σκοπό την απόδοση στατιστικής αξίας σε κάθε πρωτεΐνη και προσέφερε μια πολύ συγκεκριμένη εικόνα για το ποιες από τις non responders πρωτεΐνες είναι οι πλέον σημαντικές. Η PCA ανάλυση φέρει αρκετές λεπτομέρειες που δεν άπτονται του επιπέδου της ανάλυσης που επιτελέσαμε. Ωστόσο, ακόμα και σε ανθρώπους με μέτρια γνώση στατιστικής, προσφέρει μια γενική αίσθηση ως προς το σε τι έχει σημασία να εστιάσουμε στο δείγμα μας στατιστικά και παρέχει μια κατεύθυνση για την έρευνά μας την οποία η καθαρή βιολογική μελέτη δεν μπορεί να δώσει από μόνη της. Είναι βέβαιο πως η αναζήτηση βιοδεικτών αποτελεί ένα τεράστιας σημασίας αλλά και δυσκολίας πόνημα που πέραν της βιοστατιστικής ανάλυσης απαιτεί πληθώρα πειραμάτων και επαληθεύσιμων δεδομένων. Πολλές φορές όμως η στατιστική και βιοπληροφορική έρευνα ανοίγει οδούς που δε φανταζόμασταν εκ προοιμίου και δίνει λύση σε προβλήματα που φαίνονται αδιέξοδα, δικαιολογώντας απόλυτα την αναγκαιότητα της και τον επικουρικό της ρόλο στη σύγχρονη βιολογική έρευνα.

## ΚΕΦΑΛΑΙΟ 5

## ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- 1)Santi, A., Kugeratski, F. G., & Zanivan, S. (2018). Cancer Associated Fibroblasts: The Architects of Stroma Remodeling. *PROTEOMICS*, 18(5-6), 1700167. doi:10.1002/pmic.201700167
- 2)Saha, A., Kaul, R., Murakami, M., & Robertson, E. S. (2010). Tumor viruses and cancer biology: Modulating signaling pathways for therapeutic intervention. *Cancer Biology & Therapy*, 10(10), 961–978. doi:10.4161/cbt.10.10.13923
- 3)Richard G. Margolese,, MD, Gabriel N. Hortobagyi,, MD, and Thomas A. Buchholz,, MD.Holland-Frei Cancer Medicine. 6th edition.
- 4)Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. doi:10.3322/caac.21492
- 5) Kumar, S. K., Rajkumar, V., Kyle, R. A., van Duin, M., Sonneveld, P., Mateos, M.-V., ... Anderson, K. C. (2017). Multiple myeloma. *Nature Reviews Disease Primers*, 3, 17046. doi:10.1038/nrdp.2017.46
- 6)Leung, N., Bridoux, F., Hutchison, C. A., Nasr, S. H., Cockwell, P., ... Fermand, J.-P. (2012). Monoclonal gammopathy of renal significance: when MGUS is no longer undetermined or insignificant. *Blood*, 120(22), 4292–4295. doi:10.1182/blood-2012-07-445304
- 7)Becker, N. (2011). Epidemiology of Multiple Myeloma. *Multiple Myeloma*, 25–35. doi:10.1007/978-3-540-85772-3\_2
- 8)Kazandjian, D. (2016). Multiple myeloma epidemiology and survival: A unique malignancy. *Seminars in Oncology*, 43(6), 676–681. doi:10.1053/j.seminoncol.2016.11.004
- 9)Kuehl, W. M., & Bergsagel, P. L. (2012). Molecular pathogenesis of multiple myeloma and its premalignant precursor. *Journal of Clinical Investigation*, 122(10), 3456–3463. doi:10.1172/jci61188
- 10)Fairfield, H., Falank, C., Avery, L., & Reagan, M. R. (2016). Multiple myeloma in the marrow: pathogenesis and treatments. *Annals of the New York Academy of Sciences*, 1364(1), 32–51. doi:10.1111/nyas.13038
- 11) Terpos, E., Ntanasis-Stathopoulos, I., Gavriatopoulou, M., & Dimopoulos, M. A. (2018). Pathogenesis of bone disease in multiple myeloma: from bench to bedside. *Blood Cancer Journal*, 8(1). doi:10.1038/s41408-017-0037-4

- 12) Al-Farsi, K. (2013). Multiple Myeloma: An Update. *Oman Medical Journal*, 28(1), 3–11. doi:10.5001/omj.2013.02
- 13) Klein, B., Seckinger, A., Moehler, T., & Hose, D. (2011). Molecular Pathogenesis of Multiple Myeloma: Chromosomal Aberrations, Changes in Gene Expression, Cytokine Networks, and the Bone Marrow Microenvironment. *Multiple Myeloma*, 39–86. doi:10.1007/978-3-540-85772-3\_3
- 14) Boyce, B. F., & Xing, L. (2007). Biology of RANK, RANKL, and osteoprotegerin. *Arthritis Research & Therapy*, 9(Suppl 1), S1. doi:10.1186/ar2165
- 15) Lohr, J. G., Stojanov, P., Carter, S. L., Cruz-Gordillo, P., Lawrence, M. S., Auclair, D., ... Stockerl-Goldstein, K. (2014). Widespread Genetic Heterogeneity in Multiple Myeloma: Implications for Targeted Therapy. *Cancer Cell*, 25(1), 91–101. doi:10.1016/j.ccr.2013.12.015
- 16) Schroeder, H. W., & Cavacini, L. (2010). Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(2), S41–S52. doi:10.1016/j.jaci.2009.09.046
- 17) Corre, J., Munshi, N., & Avet-Loiseau, H. (2015). Genetics of multiple myeloma: another heterogeneity level? *Blood*, 125(12), 1870–1876. doi:10.1182/blood-2014-10-567370
- 18) Prideaux, S. M., Conway O'Brien, E., & Chevassut, T. J. (2014). The Genetic Architecture of Multiple Myeloma. *Advances in Hematology*, 2014, 1–16. doi:10.1155/2014/864058
- 19) Chng, W. J., Glebov, O., Bergsagel, P. L., & Kuehl, W. M. (2007). Genetic events in the pathogenesis of multiple myeloma. *Best Practice & Research Clinical Haematology*, 20(4), 571–596. doi:10.1016/j.beha.2007.08.004
- 20) Walker, B. A., Wardell, C. P., Johnson, D. C., Kaiser, M. F., Begum, D. B., Dahir, N. B., ... Morgan, G. J. (2013). Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. *Blood*, 121(17), 3413–3419. doi:10.1182/blood-2012-12-471888
- 21) Robiou du Pont, S., Cleyne, A., Fontan, C., Attal, M., Munshi, N., Corre, J., & Avet-Loiseau, H. (2017). Genomics of Multiple Myeloma. *Journal of Clinical Oncology*, 35(9), 963–967. doi:10.1200/jco.2016.70.6705
- 22) Rajkumar, S. V. (2016). Updated Diagnostic Criteria and Staging System for Multiple Myeloma. *American Society of Clinical Oncology Educational Book*, (36), e418–e423. doi:10.1200/edbk\_159009
- 23) Rajkumar, S. V., & Kumar, S. (2016). Multiple Myeloma: Diagnosis and Treatment. *Mayo Clinic Proceedings*, 91(1), 101–119. doi:10.1016/j.mayocp.2015.11.007
- 24) Gerecke, C., Fuhrmann, S., Striffler, S., Schmidt-Hieber, M., Einsele, H., & Knop, S. (2016). The Diagnosis and Treatment of Multiple Myeloma. *Deutsches Ärzteblatt Online*. doi:10.3238/arztebl.2016.0470

- 25)Selleck, M. J., Senthil, M., & Wall, N. R. (2017). Making Meaningful Clinical Use of Biomarkers. *Biomarker Insights*, 12, 117727191771523. doi:10.1177/1177271917715236
- 26)Lê, G. N., Bones, J., Coyne, M., Bazou, D., Dowling, P., O’gorman, P., & Larkin, A. (2019). Current and future biomarkers for risk-stratification and treatment personalisation in multiple myeloma. *Molecular Omics*. doi:10.1039/c8mo00193f
- 27)Patel, S. (2014). Role of Proteomics in Biomarker Discovery. *Advances in Protein Chemistry and Structural Biology*, 39–75. doi:10.1016/b978-0-12-800168-4.00003-2
- 28)Rinaldi, A. (2011). Teaming up for biomarker future. *EMBO Reports*, 12(6), 500–504. doi:10.1038/embor.2011.90
- 29) Bayat, A. (2002). *Science, medicine, and the future: Bioinformatics*. *BMJ*, 324(7344), 1018–1022. doi:10.1136/bmj.324.7344.1018
- 30) Can, T. (2013). *Introduction to Bioinformatics. Methods in Molecular Biology*, 51–71. doi:10.1007/978-1-62703-748-8\_4
- 31)White, C. N., Chan, D. W., & Zhang, Z. (2004). Bioinformatics strategies for proteomic profiling. *Clinical Biochemistry*, 37(7), 636–641. doi:10.1016/j.clinbiochem.2004.05.004
- 32)Haoudi, A., & Bensmail, H. (2006). Bioinformatics and data mining in proteomics. *Expert Review of Proteomics*, 3(3), 333–343. doi:10.1586/14789450.3.3.333
- 33)Schmidt, A., Forne, I., & Imhof, A. (2014). Bioinformatic analysis of proteomics data. *BMC Systems Biology*, 8(Suppl 2), S3. doi:10.1186/1752-0509-8-s2-s3
- 34)Fung, E. T., Weinberger, S. R., Gavin, E., & Zhang, F. (2005). Bioinformatics approaches in clinical proteomics. *Expert Review of Proteomics*, 2(6), 847–862. doi:10.1586/14789450.2.6.847
- 35)Shannon, P. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13(11), 2498–2504. doi:10.1101/gr.1239303
- 36)Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., ... Tang, Y. (2012). Prediction of Drug-Target Interactions and Drug Repositioning via Network-Based Inference. *PLoS Computational Biology*, 8(5), e1002503. doi:10.1371/journal.pcbi.1002503
- 37)Su, G., Morris, J. H., Demchak, B., & Bader, G. D. (2014). Biological Network Exploration with Cytoscape 3. *Current Protocols in Bioinformatics*, 47(1), 8.13.1–8.13.24. doi:10.1002/0471250953.bi0813s47
- 38)Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., ... Ideker, T. (2012). A travel guide to Cytoscape plugins. *Nature Methods*, 9(11), 1069–1076. doi:10.1038/nmeth.2212



- 39) Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., ... Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8), 1091–1093. doi:10.1093/bioinformatics/btp101
- 40) Mlecnik, B., Galon, J., & Bindea, G. (2018). Comprehensive functional analysis of large lists of genes and proteins. *Journal of Proteomics*, 171, 2–10. doi:10.1016/j.jprot.2017.03.016
- 41) A Hypergeometric Probability Model for Protein Identification and Validation Using Tandem Mass Spectral Data and Protein Sequence Databases. *Analytical Chemistry*, 75(15), 3792–3798. doi:10.1021/ac034157w
- 42) <http://proteomicsnews.blogspot.com/2018/08/cluego-another-promising-downstream.html>
- 43) Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., ... von Mering, C. (2016). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research*, 45(D1), D362–D368. doi:10.1093/nar/gkw937
- 44) Mering, C. v. (2003). STRING: a database of predicted functional associations between proteins. *Nucleic Acids Research*, 31(1), 258–261. doi:10.1093/nar/gkg034
- 45) Von Mering, C. (2004). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Research*, 33(Database issue), D433–D437. doi:10.1093/nar/gki005
- 46) Pleil, J. D., Stiegel, M. A., Madden, M. C., & Sobus, J. R. (2011). Heat map visualization of complex environmental and biomarker measurements. *Chemosphere*, 84(5), 716–723. doi:10.1016/j.chemosphere.2011.03.017
- 47) Key, M. (2012). A tutorial in displaying mass spectrometry-based proteomic data using heat maps. *BMC Bioinformatics*, 13(Suppl 16), S10. doi:10.1186/1471-2105-13-s16-s1
- 48) Zhao, S., Guo, Y., Sheng, Q., & Shyr, Y. (2014). Advanced Heat Map and Clustering Analysis Using Heatmap3. *BioMed Research International*, 2014, 1–6. doi:10.1155/2014/986048
- 49) Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology*, 26(3), 303–304. doi:10.1038/nbt0308-303
- 50) Groth, D., Hartmann, S., Klie, S., & Selbig, J. (2012). Principal Components Analysis. *Computational Toxicology*, 527–547. doi:10.1007/978-1-62703-059-5\_22