



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Εφαρμογή βιοπληροφορικών τεχνικών και
αλγορίθμων μηχανικής μάθησης με σκοπό την ανεύρεση
διαγνωστικών, προγνωστικών και προβλεπτικών βιοδεικτών
για τον καρκίνο της ουροδόχου κύστης**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Μιχαήλ Ε. Σαραφίδης

Αθήνα, Νοέμβριος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Εφαρμογή βιοπληροφορικών τεχνικών και
αλγορίθμων μηχανικής μάθησης με σκοπό την ανεύρεση
διαγνωστικών, προγνωστικών και προβλεπτικών βιοδεικτών
για τον καρκίνο της ουροδόχου κύστης**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Μιχαήλ Ε. Σαραφίδης

Συμβουλευτική Επιτροπή : Δημήτριος – Διονύσιος Κουτσούρης (Επιβλέπων)
Γεώργιος Ματσόπουλος
Κωνσταντίνα Νικήτα

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 25^η Νοεμβρίου 2022.

.....

Δημήτριος Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π.

.....

Παναγιώτης Μπαμίδης
Καθηγητής Α.Π.Θ.

.....

Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

.....

Βασίλειος Ζουμπουρλής
Διευθυντής Ερευνών Ε.Ι.Ε.

.....

Θεμιστοκλής Έξαρχος
Επίκ. Καθηγητής Ιονίου Παν/μίου

.....

Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

.....

Παναγιώτης Τσανάκας
Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2022

.....

Μιχαήλ Ε. Σαραφίδης

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Μιχαήλ Ε. Σαραφίδης, 2022.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η συγκεκριμένη διδακτορική διατριβή αφορά στην εφαρμογή βιοπληροφορικών τεχνικών και αλγορίθμων μηχανικής μάθησης με σκοπό την ανεύρεση διαγνωστικών, προγνωστικών και προβλεπτικών – ως προς την απόκριση στη θεραπεία – βιοδεικτών για τον καρκίνο της ουροδόχου κύστης. Ο καρκίνος της ουροδόχου κύστης αποτελεί μία ετερογενή ασθένεια με μεγάλη επίπτωση και επιπολασμό παγκοσμίως, η οποία ευθύνεται για σημαντική νοσηρότητα και θνητότητα. Στο πλαίσιο αυτής της εργασίας, πραγματοποιήθηκε μια συστηματική αναζήτηση η οποία συγκέντρωσε όλα τα δεδομένα γονιδιακής έκφρασης από μικροσυστοιχίες DNA που έχουν καταχωρηθεί στη βάση δεδομένων Gene Expression Omnibus (GEO) του Εθνικού Κέντρου Πληροφοριών Βιοτεχνολογίας (NCBI) των Η.Π.Α. και αφορούσαν στη μελέτη και σύγκριση υγιών και καρκινικών ιστών για τη συγκεκριμένη ασθένεια. Από τη συστηματική ανασκόπηση εντοπίστηκαν 18 σύνολα δεδομένων που πληρούσαν τα κριτήρια της αναζήτησης και τα οποία συμπεριλήφθηκαν στην ενοποιητική μετα-ανάλυση. Για τα σύνολα αυτά, λήφθηκαν τα πρωτογενή δεδομένα, προ-επεξεργάστηκαν ανάλογα με τον τύπο της μικροσυστοιχίας και, αφού ελέγχθηκαν ποιοτικά, ενσωματώθηκαν, έπειτα από κανονικοποίηση, σε ένα ενοποιημένο μετα-σύνολο δεδομένων. Το ενοποιημένο αυτό μετα-σύνολο δεδομένων χρησιμοποιήθηκε για τον προσδιορισμό των διαφορικά εκφραζόμενων γονιδίων μεταξύ των υγιών και των καρκινικών δειγμάτων. Έπειτα, πραγματοποιήθηκε η ανάλυση του δικτύου των πρωτεϊνών που παράγονται από αυτά τα γονίδια και ανιχνεύθηκαν τα σημαντικά γονίδια – κόμβοι του δικτύου. Ακόμη, εφαρμόστηκε η ανάλυση του δικτύου της σταθμισμένης γονιδιακής συνέκφρασης, η οποία αποτελεί μία μη επιβλεπόμενη τεχνική, και ανιχνεύθηκαν τα γονίδια – κόμβοι, τα οποία παρουσίαζαν μεγάλη συσχέτιση με το φαινότυπο των δειγμάτων. Στη συνέχεια, εντοπίστηκαν τα κοινά γονίδια – κόμβοι των δύο παραπάνω μεθόδων, τα οποία και αποτέλεσαν τα γονίδια – κόμβους «κλειδιά» της παρούσας έρευνας. Τα συγκεκριμένα γονίδια μελετήθηκαν αρχικά για τη διαφορική έκφρασή τους σε δείγματα από ούρα και πλάσμα αίματος ασθενών με καρκίνο ουροδόχου κύστης και υγιών μαρτύρων. Ακολούθως, αναλύθηκε η προγνωστική ικανότητα των γονιδίων αυτών με χρήση μονο-μεταβλητής, πολυ-μεταβλητής και LASSO ανάλυσης παλινδρόμησης. Επίσης, υλοποιήθηκε η ανάλυση των καμπυλών επιβίωσης Kaplan – Meier και του λειτουργικού χαρακτηριστικού δέκτη (ROC) για τον προσδιορισμό των γονιδίων με προγνωστική αξία και κατασκευάστηκε ένα μοντέλο πρόγνωσης με βάση την έκφραση τριών γονιδίων. Το μοντέλο αυτό ελέγχθηκε ως προς την απόδοσή του σε δύο ανεξάρτητα σύνολα δεδομένων, παρουσιάζοντας ιδιαίτερα καλή απόδοση. Ακόμη, με την εφαρμογή των ανωτέρω μεθόδων, αναλύθηκε η προβλεπτική ικανότητα των γονιδίων αυτών όσον αφορά την πρόβλεψη της ανταπόκρισης των ασθενών με διηθητικό καρκίνο ουροδόχου κύστης στην προ-εγχειρητική χημειοθεραπεία. Έτσι, δημιουργήθηκε ένα μοντέλο πρόβλεψης με βάση την έκφραση έξι γονιδίων, το οποίο ελέγχθηκε σε δύο ανεξάρτητα σύνολα δεδομένων, παρουσιάζοντας καλή απόδοση. Από τις ανωτέρω αναλύσεις, προσδιορίστηκε ένα σύνολο εννέα γονιδίων – βιοδεικτών τα οποία βρέθηκε να εκφράζονται διαφορικά στα ούρα ή στο πλάσμα αίματος μεταξύ των ασθενών με καρκίνο ουροδόχου κύστης και των υγιών μαρτύρων, και, παράλληλα, φάνηκε να κατέχουν κάποια προγνωστική ή προβλεπτική ικανότητα. Η έκφραση των συγκεκριμένων βιοδεικτών στον ιστό της ουροδόχου κύστης ασθενών αλλά και υγιών μαρτύρων επιβεβαιώθηκε με χρήση εικόνων ανοσοϊστοχημείας και με την αξιοποίηση δημόσιων πλατφορμών βιοπληροφορικής. Τέλος, οι συγκεκριμένοι εννέα βιοδείκτες χρησιμοποιήθηκαν ως χαρακτηριστικά για την υλοποίηση μοντέλων ταξινόμησης, τα οποία φάνηκε να παρουσιάζουν ιδιαίτερα υψηλή απόδοση ως προς την διακριτοποίηση των δειγμάτων σε καρκινικά και υγιή, αναδεικνύοντας τη διαγνωστική αξία των συγκεκριμένων βιοδεικτών.

Λέξεις – Κλειδιά: καρκίνος ουροδόχου κύστης, ουρολογικός καρκίνος, πρόγνωση, διάγνωση, πρόβλεψη, βιοδείκτες, γονίδια – κόμβοι, βιοπληροφορική, μηχανική μάθηση, μικροσυστοιχίες, μετα-ανάλυση

Αυτή η σελίδα είναι σκόπιμα λευκή.

Abstract

This PhD thesis concerns the application of bioinformatics techniques and machine learning algorithms towards the identification of diagnostic, prognostic and predictive - in terms of patients' response to treatment - biomarkers for bladder cancer. Bladder cancer (BCa) is a heterogeneous disease accounting for high incidence and prevalence worldwide, and is responsible for significant morbidity and mortality. In the context of this study, a systematic search was performed and all the gene expression data from DNA microarrays registered in the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI), were collected in order to study and compare healthy and cancerous tissues for this disease. The systematic search identified 18 datasets that fulfilled the inclusion criteria and were included in the integrated meta-analysis. For these datasets, the raw data were obtained, pre-processed according to the microarray platform and, after the quality control and normalization, were integrated into a merged meta-dataset. This merged meta-dataset was utilized to determine the differentially expressed genes between control and cancer samples. Then, the protein – protein interaction network analysis was performed and the hub genes were detected. Furthermore, the weighted gene co-expression network analysis, which is an unsupervised technique, was conducted and the hub genes, which showed a high correlation with the phenotype of the samples, were detected. Subsequently, the common hub genes of the above two methods were identified, which were determined as the key hub genes of the present study. These genes were first studied for their differential expression in urine and blood plasma samples from bladder cancer patients and healthy controls. Subsequently, the predictive value of these genes was analyzed using univariate Cox, multivariate Cox and LASSO regression analyses. Kaplan-Meier survival curve and functional receiver operating characteristic (ROC) analyses were also implemented to identify genes with prognostic value and a prognostic model was constructed based on the expression of three genes. This model was tested for its performance on two independent datasets, showing high performance. Furthermore, by applying the above methods, the predictive ability of the key hub genes – in terms of predicting the response of patients with invasive bladder cancer to pre-operative chemotherapy – was analyzed. Thus, a prediction model based on the expression of six genes was created and tested on two independent datasets, showing good performance. From the above analyses, a set of nine biomarker genes were identified, which were found to be differentially expressed in urine or blood plasma between bladder cancer patients and healthy controls, and, at the same time, appeared to possess some predictive or prognostic ability. The expression of these biomarkers in bladder tissue of patients and healthy controls was confirmed using immunohistochemistry images and utilizing public bioinformatics platforms. Finally, these nine biomarkers were used as features to implement classification models, which showed a particularly high performance, in terms of discriminating samples into cancerous and healthy ones, highlighting the diagnostic value of these biomarkers.

Keywords: bladder cancer; urologic cancer; diagnosis; prognosis; prediction; biomarkers; hub genes; bioinformatics; machine learning; microarrays; meta-analysis

Αυτή η σελίδα είναι σκόπιμα λευκή.

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα Καθηγητή μου κ. Δημήτριο – Διονύσιο Κουτσούρη για την ευκαιρία που μου έδωσε να πραγματοποιήσω διδακτορικές σπουδές στο Ε.Μ.Π. στο αντικείμενο της Βιοϊατρικής Τεχνολογίας. Οι συμβουλές, οι υποδείξεις και η καθοδήγησή του αποτέλεσε για εμένα σημαντικό αρωγό στην εκπόνηση αυτής της διδακτορικής διατριβής. Επίσης, μου προσέφερε τη δυνατότητα να ασχοληθώ με ένα θέμα που άπτεται στην επιστήμη της Βιοπληροφορικής και, πιο συγκεκριμένα, την εφαρμογή αυτής στην ασθένεια του καρκίνου, γεγονός το οποίο αποτελούσε για εμένα προσωπικό όραμα όταν ξεκίνησα τις διδακτορικές μου σπουδές. Κατά τη διάρκεια των διδακτορικών μου σπουδών, η εμπειρία που απέκτησα από τις εκπαιδευτικές και ερευνητικές δραστηριότητες του εργαστηρίου είναι πολύτιμη και θεωρώ πως θα αποτελέσει ισχυρό εφόδιο για την μετέπειτα επαγγελματική μου πρόοδο.

Συνεχίζοντας, θα ήθελα να ευχαριστήσω τον Καθηγητή κ. Γεώργιο Ματσόπουλο, το νέο Διευθυντή του Εργαστηρίου Βιοϊατρικής Τεχνολογίας, ο οποίος με τις συμβουλές και τις υποδείξεις του συνέβαλε στην προσπάθειά μου αυτή και με κατεύθυνε κατάλληλα. Επιπρόσθετα, θα ήθελα να ευχαριστώ θερμά και τα υπόλοιπα πέντε μέλη της επταμελούς συμβουλευτικής επιτροπής, την Καθηγήτρια κ. Κωνσταντίνα Νικήτα, τον Καθηγητή κ. Παναγιώτη Μπαμίδη, τον Διευθυντή Ερευνών κ. Βασίλειο Ζουμπουρλή, τον Καθηγητή κ. Παναγιώτη Τσανάκα και τον Επίκουρο Καθηγητή κ. Θεμιστοκλή Έξαρχο. Ένα ιδιαίτερο ευχαριστώ οφείλω στον φίλο και συνεργάτη Δρ. Γεώργιο Λάμπρου, ο οποίος βοήθησε ιδιαίτερα σε όλα τα στάδια εκπόνησης της παρούσας διδακτορικής διατριβής, από τη σύλληψη της ιδέας μέχρι και την ερμηνεία των βιολογικών συμπερασμάτων. Η συμβολή του ήταν καθοριστική και καίρια καθ' όλη τη διάρκεια αυτής της ερευνητικής προσπάθειας.

Ακόμη, θα ήθελα να ευχαριστήσω όλα τα μέλη του Εργαστηρίου Βιοϊατρικής Τεχνολογίας του Ε.Μ.Π., συναδέλφους και – τολμώ να πω – σημαντικούς φίλους, με τους οποίους μοιραστήκαμε όλα αυτά τα χρόνια πολλές όμορφες στιγμές, σημαντικές εμπειρίες, κοινές αγωνίες, παραγωγική ανταλλαγή ιδεών, ταξίδια και πολύ γέλιο.

Κλείνοντας, θα ήθελα να ευχαριστήσω μέσα από την καρδιά μου τους αγαπημένους μου ανθρώπους από το οικογενειακό και φιλικό μου περιβάλλον, οι οποίοι ήταν συνοδοιπόροι, κάποιιοι πιο άμεσα και άλλοι πιο έμμεσα, σε όλη αυτή την προσπάθεια και το ερευνητικό ταξίδι. Από αυτούς, επιθυμώ να κάνω ιδιαίτερη μνεία στη μητέρα μου, η οποία στέκεται πάντα δίπλα μου με αμέριστη αγάπη, συμπαράσταση και κατανόηση, συμβάλλοντας συναισθηματικά και ηθικά στη στήριξή μου σε κάθε βήμα της ζωής μου.

Αυτή η σελίδα είναι σκόπιμα λευκή.

Περιεχόμενα

Περίληψη.....	5
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	11
Λίστα Εικόνων	13
Λίστα Πινάκων.....	19
Ευρετήριο όρων.....	21
1. Εισαγωγή.....	23
1.1 Καρκίνος	23
Επιδημιολογία	23
Αίτια και Παράγοντες Κινδύνου	25
Αντιμετώπιση	29
1.2 Καρκίνος ουροδόχου κύστης	31
Ορισμός.....	31
Επιδημιολογία και Παράγοντες Κινδύνου	31
Ανατομία και Φυσιολογία της Ουροδόχου Κύστης	36
Μηχανισμοί και Παθοφυσιολογία του Καρκίνου της Ουροδόχου Κύστης	38
Διάγνωση, προληπτικός έλεγχος και πρόληψη	40
Διαχείριση – Αντιμετώπιση – Θεραπεία	42
Ποιότητα ζωής.....	48
Προοπτικές.....	50
1.3 Βιοδείκτες για τον καρκίνο ουροδόχου κύστης	52
1.4 Γονιδιακή έκφραση	54
1.5 Μικροσυστοιχίες	56
Προετοιμασία δείγματος και σήμανση.....	58
Υβριδισμός.....	59
Πλύσιμο.....	59
Λήψη εικόνας.....	59
Κανονικοποίηση.....	59
1.6 Επεξεργασία και Ανάλυση δεδομένων μικροσυστοιχιών	60
1.7 Ομικές τεχνολογίες και βιολογικά δεδομένα	68
1.8 Επαναχρησιμοποίηση Δημόσιων Δεδομένων Γονιδιακής Έκφρασης	70
Επαναχρησιμοποίηση δεδομένων ως μέρος της επανάστασης των «μεγάλων δεδομένων».....	70
Τύποι επαναχρησιμοποιήσιμων δεδομένων	72
1.9 Βιοπληροφορικές μετα-αναλύσεις για την ανεύρεση γονιδίων – κόμβων στον ΚΟΚ.....	74

1.10 Περιγραφή της παρούσας μελέτης	75
2. Υλικό και Μέθοδοι.....	77
2.1 Σχεδιασμός Μελέτης και Ροή Εργασίας	77
2.2 Πηγή Δεδομένων, Συστηματική Αναζήτηση και Επιλογή Συνόλων Δεδομένων	79
2.3 Προ-επεξεργασία με βάση τον κατασκευαστή της πλατφόρμας.....	80
2.4 Ποιοτικός Έλεγχος Δειγμάτων	81
2.5 Αντιστοίχιση Ανιχνευτών – Γονιδίων	82
2.6 Φαινόμενα Παρτίδας και Διατεχνολογική Κανονικοποίηση	82
2.7 Ανάλυση Διαφορικής Έκφρασης Γονιδίων.....	83
2.8 Λειτουργική Ανάλυση Εμπλουτισμού της Γονιδιακής Έκφρασης	84
2.9 Ανάλυση του Δικτύου Αλληλεπιδράσεων Πρωτεΐνης – Πρωτεΐνης.....	85
2.10 Ανάλυση Δικτύων Συνέκφρασης Γονιδίων.....	86
2.11 Διαφορική Έκφραση Γονιδίων στα Ούρα και στο Πλάσμα Αίματος	87
2.12 Εύρεση Προγνωστικών Γονιδίων για τον Καρκίνο της Ουροδόχου Κύστης.....	88
2.13 Εύρεση Προβλεπτικών Γονιδίων για τον Καρκίνο της Ουροδόχου Κύστης	90
2.14 Ταυτοποίηση της Έκφρασης των Βιοδεικτών και Ανοσοϊστοχημεία	91
2.15 Ανάλυση Διαγνωστικής Απόδοσης των Βιοδεικτών.....	92
3. Αποτελέσματα.....	93
3.1 Συστηματική αναζήτηση και Επιλογή Συνόλων Δεδομένων	93
3.2 Ποιοτικός Έλεγχος Δειγμάτων	94
3.3 Αντιστοίχιση Ανιχνευτών – Γονιδίων	95
3.4 Φαινόμενα Παρτίδας και Διατεχνολογική Κανονικοποίηση	95
3.5 Ανάλυση Διαφορικής Έκφρασης Γονιδίων.....	97
3.6 Λειτουργική Ανάλυση Εμπλουτισμού της Γονιδιακής Έκφρασης	99
3.7 Ανάλυση του Δικτύου Αλληλεπιδράσεων Πρωτεΐνης – Πρωτεΐνης.....	103
3.8 Ανάλυση Δικτύων Συνέκφρασης Γονιδίων.....	106
3.9 Διαφορική Έκφραση Γονιδίων στα Ούρα και στο Πλάσμα Αίματος	111
3.10 Προγνωστικά Γονίδια για τον Καρκίνο της Ουροδόχου Κύστης.....	114
3.11 Προβλεπτικά Γονίδια για τον Καρκίνο της Ουροδόχου Κύστης	121
3.12 Ταυτοποίηση της Έκφρασης των Βιοδεικτών και Ανοσοϊστοχημεία	128
3.13 Διαγνωστική Απόδοση των Γονιδίων / Βιοδεικτών για τον Καρκίνο της Ουροδόχου Κύστης	134
4. Συζήτηση.....	137
5. Συμπεράσματα.....	149
6. Βιβλιογραφικές Αναφορές	151
Λίστα δημοσιεύσεων.....	173

Λίστα Εικόνων

Εικόνα 1: Ο καρκίνος είναι μια ασθένεια που προκαλείται όταν τα κύτταρα διαιρούνται ανεξέλεγκτα και εξαπλώνονται στους γύρω ιστούς.	23
Εικόνα 2: Εθνική κατάταξη του καρκίνου ως αιτία θανάτου σε ηλικίες < 70 ετών για το έτος 2019.	24
Εικόνα 3: (Α) Ο 4-βάθμιος δείκτης ανθρώπινης ανάπτυξης (HDI) και (Β) 20 περιοχές του κόσμου με τα αντίστοιχα ποσοστά πληθυσμού.	25
Εικόνα 4: Ο καρκίνος προκαλείται από αλλαγές στο DNA. Οι περισσότερες αλλαγές στο DNA που προκαλούν καρκίνο συμβαίνουν σε τμήματα του DNA που ονομάζονται γονίδια. Οι αλλαγές αυτές ονομάζονται επίσης γενετικές αλλαγές.	26
Εικόνα 5: Οι γενετικές αλλαγές που προκαλούν καρκίνο μπορεί να είναι κληρονομικές ή να οφείλονται σε ορισμένους περιβαλλοντικούς παράγοντες. Οι γενετικές αλλαγές μπορούν επίσης να συμβούν λόγω σφαλμάτων που συμβαίνουν κατά τη διαίρεση των κυττάρων.	27
Εικόνα 6: Προτυποποιημένη κατά ηλικία επίπτωση του ΚΟΚ στους άντρες παγκοσμίως.	32
Εικόνα 7: Προτυποποιημένη κατά ηλικία επίπτωση του ΚΟΚ στις γυναίκες παγκοσμίως.	32
Εικόνα 8: Μέσος αριθμός νέων περιπτώσεων και επίπτωση του ΚΟΚ ανά ηλικιακή ομάδα.	33
Εικόνα 9: Αναμενόμενος αριθμός νέων διαγνώσεων και θανάτων ΚΟΚ για το έτος 2030 συγκριτικά με το 2020.	34
Εικόνα 10: Το ουροποιητικό σύστημα του ανθρώπου.	36
Εικόνα 11: Ανατομία της ουροδόχου κύστης και της ουρήθρας σε άνδρες και γυναίκες.	37
Εικόνα 12: Ο καρκίνος ουροδόχου κύστης (ΚΟΚ).	38
Εικόνα 13: Τύποι και στάδια του ΚΟΚ.	39
Εικόνα 14: Παραδείγματα εικόνων κυστεοσκόπησης που λαμβάνονται με τη χρήση λευκού (αριστερά) και μπλε φωτός (δεξιά) για την καλύτερη διάγνωση του ΚΟΚ.	43
Εικόνα 15: Ανατομική έκταση της τυπικής λεμφαδενεκτομής στον καρκίνο της ουροδόχου κύστης.	46
Εικόνα 16: Το κεντρικό δόγμα της μοριακής βιολογίας.	55
Εικόνα 17: Τα κύρια στάδια ενός πειράματος με τη χρήση μικροσυστοιχιών.	57
Εικόνα 18: Ενδεικτική εικόνα μικροσυστοιχίας γονιδίων.	60
Εικόνα 19: Παραδείγματα γραφικών παραστάσεων ελέγχου ποιότητας κατά την ανάλυση δεδομένων διαφορικής έκφρασης. Από αριστερά προς τα δεξιά: Κατανομές έντασης συστοιχιών, διάγραμμα PCA, εκτιμήσεις πυκνότητας.	62
Εικόνα 20: Διάγραμμα διασποράς κόκκινου (R) έναντι πράσινου (G) καναλιού (αριστερά) Διάγραμμα MA: μέση ένταση (A) έναντι λογαριθμικού λόγου (M) (δεξιά).	63
Εικόνα 21: Διάγραμμα MA στις αρχικές (πρωτογενείς) τιμές (αριστερά) Διάγραμμα MA στις κανονικοποιημένες τιμές (δεξιά).	63
Εικόνα 22: Ένα παράδειγμα ενός θερμικού χάρτη στον οποίο τα γονίδια έχουν ομαδοποιηθεί με βάση το μοτίβο της γονιδιακής τους έκφρασης.	65
Εικόνα 23: Ένα παράδειγμα ανάλυσης μονοπατιών από το Wikipathways.	66
Εικόνα 24: Επισκόπηση της διοχέτευσης (pipeline) ανάλυσης δεδομένων μικροσυστοιχιών.	67
Εικόνα 25: Οι σύγχρονες ομικές (-omics) τεχνολογίες αποτελούν τις υψηλού επιπέδου τεχνολογίες που χρησιμοποιούνται για την ολιστική ανάλυση των μορίων που συνθέτουν τα κύτταρα των ζωντανών οργανισμών.	68
Εικόνα 26: Η πληθώρα των σύγχρονων τεχνολογιών παραγωγής βιολογικών δεδομένων δημιουργηθεί συμφόρηση (bottleneck) ως προς την ανάλυση, ερμηνεία και κλινική αξιοποίηση των δεδομένων αυτών.	69
Εικόνα 27: Η εξέλιξη της συμπεριφοράς διαμοιρασμού δεδομένων. (1) Η τεχνική πρόοδος καθιστά δυνατή τον παγκόσμιο διαμοιρασμό μεγάλων συνόλων δεδομένων, (2) η αυξημένη προσβασιμότητα στην απαιτούμενη τεχνολογία τον καθιστά ευρέως διαθέσιμο, (3) οι υποχρεώσεις και τα οφέλη για τους ερευνητές καθιερώνουν μία συμπεριφορά / κουλτούρα διαμοιρασμού, (4) ο αριθμός των	

συνόλων δεδομένων αυξάνεται και τα καθιστά ελκυστικά, (5) η επαναχρησιμοποίηση αναπτύσσεται με την πάροδο του χρόνου - γεγονός που οδηγεί σε θετικό βρόχο ανατροφοδότησης και στη συνήθεια να μοιράζονται όλα τα δεδομένα.....	71
Εικόνα 28: Βασικοί τύποι επαναχρησιμοποιούμενων βιολογικών δεδομένων	73
Εικόνα 29: Η συνολική μεθοδολογία σχεδιασμού της μελέτης και η ροή εργασίας της ενοποιητικής μετα-ανάλυσης της παρούσας εργασίας.....	78
Εικόνα 30: Διάγραμμα ροής της συστηματικής αναζήτησης βάσει του πρωτοκόλλου Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) Μια πλατφόρμα μικροσυστοιχίας η οποία είναι πανομοιότυπη με μια εμπορική πλατφόρμα και στην οποία χρησιμοποιείται ένα προσαρμοσμένο, αναδιαμορφωμένο περιβάλλον CDF για την εξαγωγή δεδομένων, θεωρήθηκε ως εμπορική πλατφόρμα.	80
Εικόνα 31: Εφαρμογή της μεθόδου PCA σε καθένα από τα 18 σύνολα δεδομένων, χρησιμοποιώντας την ευκλείδεια απόσταση για τη μέτρηση των ανομοιοτήτων στις τιμές γονιδιακής έκφρασης μεταξύ των δειγμάτων. Τα κόκκινα σημεία υποδηλώνουν δείγματα KOK και τα μπλε σημεία υποδηλώνουν δείγματα ελέγχου.....	97
Εικόνα 32: Εφαρμογή της μεθόδου PCA στα σύνολα δεδομένων GSE13507A και GSE13507B. Τα κόκκινα σημεία υποδηλώνουν δείγματα KOK και τα μπλε σημεία υποδηλώνουν δείγματα ελέγχου..	97
Εικόνα 33: Διάγραμμα ηφαιστείου των διαφορικά εκφραζόμενων γονιδίων μεταξύ των δειγμάτων KOK και των δειγμάτων ελέγχου στο ενοποιημένο μετα-σύνολο δεδομένων. Τα διαφορικά εκφραζόμενα γονίδια προσδιορίστηκαν με βάση τα κριτήρια $ \log_2(FC) \geq 1,3$ και προσαρμοσμένη τιμή- $p < 0,01$, όπως απεικονίζονται με γκριζες γραμμές. Τα μπλε και τα κόκκινα σημεία υποδηλώνουν τα γονίδια που υπο-εκφράζονται και υπερ-εκφράζονται, αντίστοιχα. Τα μαύρα σημεία υποδηλώνουν γονίδια που δεν παρουσιάζουν στατιστικά σημαντική διαφορά στην έκφραση μεταξύ των δύο φαινοτύπων.....	98
Εικόνα 34: Διάγραμμα θερμικού χάρτη των 100 κορυφαίων διαφορικά εκφραζόμενων γονιδίων μεταξύ των δειγμάτων KOK και των υγιών δειγμάτων ελέγχου στο ενοποιημένο μετα-σύνολο δεδομένων. Το μπλε και το κόκκινο χρώμα αντιπροσωπεύουν τη σχετική μείωση και αύξηση της έκφρασης, αντίστοιχα, ενώ το λευκό χρώμα αντιπροσωπεύει ότι το γονίδιο δεν παρουσιάζει κάποια σημαντική αλλαγή στη γονιδιακή έκφραση. Οι οριζόντιοι και κάθετοι άξονες είναι ομαδοποιημένοι ανά γονίδια και δείγματα, αντίστοιχα.....	99
Εικόνα 35: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού με βάση την Οντολογία Γονιδίων (GO) όσον αφορά:	101
Εικόνα 36: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού μονοπατιών με βάση την KEGG.	101
Εικόνα 37: (A) Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού μονοπατιών με βάση τη Reactome (B) Χάρτης των εμπλουτισμένων όρων της βάσης Reactome που παρουσιάζονται σε μορφή δικτύου.	102
Εικόνα 38: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού με βάση την Οντολογία Ασθενειών (DO).	103
Εικόνα 39: Το κατασκευασμένο δίκτυο ΑΠΠ, όπως οπτικοποιήθηκε από τη βάση δεδομένων STRING. Οι κόμβοι αντιπροσωπεύουν πρωτεΐνες και οι ακμές αντιπροσωπεύουν συσχετίσεις μεταξύ πρωτεΐνης-πρωτεΐνης. Το πάχος της γραμμής αντιπροσωπεύει το σκορ εμπιστοσύνης της εκάστοτε λειτουργικής συσχέτισης.....	104
Εικόνα 40: Το δίκτυο ΑΠΠ των τελικών 87 βασικών γονιδίων – κόμβων. Οι κόμβοι αντιπροσωπεύουν τις πρωτεΐνες και οι ακμές αντιπροσωπεύουν τις συσχετίσεις μεταξύ πρωτεΐνης-πρωτεΐνης. Το πάχος της γραμμής αντιπροσωπεύει το σκορ εμπιστοσύνης μιας λειτουργικής συσχέτισης.....	105
Εικόνα 41: Σύνοψη των βασικών δεικτών του δικτύου σε συνάρτηση με τη δύναμη ελαφριού καταωφλιού για κάθε ένα από τα οκτώ επιλεγμένα σύνολα δεδομένων τα οποία παρουσιάζονται με διαφορετικά χρώματα.....	107

Εικόνα 42: Το δενδρόγραμμα της ιεραρχικής ομαδοποίησης των 8.201 γονιδίων με σκοπό τον προσδιορισμό συναινετικών ομάδων με βάση τη συναινετική τοπολογική επικάλυψη. Τα γονίδια μιας κοινής ομάδας λαμβάνουν το ίδιο χρώμα, όπως παρουσιάζεται στη χρωματική ζώνη κάτω από το δενδρόγραμμα. Τα γονίδια που δεν αντιστοιχούν σε καμία από τις ομάδες έχουν γκριζό χρώμα.....	107
Εικόνα 43: Οι θερμικοί χάρτες των σχέσεων των ιδιοδιανυσμάτων των συναινετικών ομάδων και των φαινοτυπικών χαρακτηριστικών για καθένα από τα οκτώ σύνολα δεδομένων. Κάθε γραμμή αντιστοιχεί στο ιδιοδιάνυσμα της συναινετικής ομάδας και κάθε στήλη αντιστοιχεί σε ένα φαινοτυπικό χαρακτηριστικό. Κάθε κελί περιέχει την αντίστοιχη συσχέτιση (που κυμαίνεται από μπλε σε κόκκινο) και την τιμή-p.....	109
Εικόνα 44: Ο συναινετικός θερμικός χάρτης των σχέσεων των ιδιοδιανυσμάτων των συναινετικών ομάδων και των φαινοτυπικών χαρακτηριστικών. Κάθε γραμμή αντιστοιχεί στο ιδιοδιάνυσμα της συναινετικής ομάδας και κάθε στήλη αντιστοιχεί σε ένα φαινοτυπικό χαρακτηριστικό. Κάθε κελί περιέχει την αντίστοιχη συσχέτιση (που κυμαίνεται από μπλε σε κόκκινο) και την τιμή-p.....	109
Εικόνα 45: Διαγράμματα διασποράς της γονιδιακής σημαντικότητας (GS) για τον φαινότυπο του καρκίνου και της συμμετοχής στην ομάδα (MM) για τις βασικές ομάδες. Οι γραμμές υποδεικνύουν το ανώτερο και το κατώτερο τεταρτημόριο.....	110
Εικόνα 46: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα ούρων με τη χρήση του μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE51843 ($n = 11$).....	111
Εικόνα 47: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα ούρων με τη χρήση μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE68020 ($n = 50$).....	112
Εικόνα 48: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα πλάσματος αίματος με τη χρήση του μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE138118 ($n = 57$).....	114
Εικόνα 49: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE13507 ($n = 165$) (σύνολο εκπαίδευσης) (Α) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με KOK, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με τον διάμεσο προγνωστικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, το τρίτο, το πέμπτο και το δέκατο έτος μετά τη διάγνωση) με σκοπό την αξιολόγηση της απόδοσης του προγνωστικού μοντέλου των τριών γονιδίων. (Γ) Διάγραμμα δάσους για την πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox στο σύνολο δεδομένων GSE13507. Η εικόνα ενσωματώνει την τιμή του λόγου κινδύνου (e^{coef}) μαζί με το 95% διάστημα εμπιστοσύνης (CI) και την τιμή-p για κάθε γονίδιο.....	117
Εικόνα 50: Ανάλυση συσχέτισης της γονιδιακής έκφρασης των τριών γονιδίων του προγνωστικού μοντέλου, που προέκυψαν από την πλατφόρμα GEPIA2. Χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson και η ανάλυση έδειξε ότι δεν υπάρχει στατιστικά σημαντικός συντελεστής συσχέτισης μεταξύ αυτών των γονιδίων για τον KOK (μέγιστη τιμή 0,6).....	117
Εικόνα 51: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE32894 ($n = 224$) (πρώτο σύνολο δοκιμής) (Α) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με KOK, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προγνωστικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προγνωστικής απόδοσης του μοντέλου των τριών γονιδίων.	118
Εικόνα 52: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE32548 ($n = 131$) (δεύτερο σύνολο δοκιμής) (Α) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με KOK, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προγνωστικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προγνωστικής απόδοσης του μοντέλου των τριών γονιδίων.	118
Εικόνα 53: Διαγράμματα επιβίωσης Kaplan – Meier για την προγνωστική υπογραφή των τριών γονιδίων, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών υψηλού και χαμηλού κινδύνου, αντίστοιχα. Οι ασθενείς	

ομαδοποιήθηκαν σύμφωνα με τη διάμεση τιμή έκφρασης των τριών γονιδίων, όσον αφορά τη συνολική επιβίωση, και σύμφωνα με μια προσαρμοσμένη υψηλή και χαμηλή τιμή αποκοπής 85% και 15%, αντίστοιχα, όσον αφορά την επιβίωση ελεύθερης νόσου.	119
Εικόνα 54: Διαγράμματα επιβίωσης Kaplan – Meier για τα βασικά γονίδια – κόμβους, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών υψηλού και χαμηλού κινδύνου, αντίστοιχα. Οι ασθενείς ομαδοποιήθηκαν σύμφωνα με τις διάμεσες τιμές αποκοπής ή τις τιμές τεταρτημόριου.	121
Εικόνα 55: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά μεταξύ των ομάδων «Καμία ανταπόκριση», «Μερική ανταπόκριση» και «Πλήρης ανταπόκριση» όσον αφορά την προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη των ασθενών με μυοδιηθητικό ΚΟΚ.	122
Εικόνα 56: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά μεταξύ των ομάδων «Καμία ανταπόκριση» και «Μερική ή πλήρης ανταπόκριση» όσον αφορά την προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη των ασθενών με μυοδιηθητικό ΚΟΚ.	123
Εικόνα 57: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE169455 (n = 148) (σύνολο εκπαίδευσης) (Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων. (Γ) Διάγραμμα δάσους για την πολυμεταβλητή ανάλυση παλινδρόμησης Cox στο σύνολο δεδομένων GSE169455. Η Εικόνα ενσωματώνει την τιμή του λόγου κινδύνου (e^{coef}) μαζί με το 95% διάστημα εμπιστοσύνης (CI) και την τιμή-p για κάθε γονίδιο.	126
Εικόνα 58: Ανάλυση συσχέτισης της γονιδιακής έκφρασης των έξι γονιδίων του προβλεπτικού μοντέλου, που προέκυψαν από την πλατφόρμα GEPIA2. Χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson και έδειξε ότι δεν υπάρχει στατιστικά σημαντικός συντελεστής συσχέτισης μεταξύ αυτών των γονιδίων για τον ΚΟΚ (μέγιστη τιμή 0,44).	126
Εικόνα 59: Ανάλυση επιβίωσης με χρήση του συνόλου δεδομένων GSE87304 (n = 258) (πρώτο σύνολο δοκιμής) (Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το τρίτο, πέμπτο και έκτο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων.	127
Εικόνα 60: Ανάλυση επιβίωσης με χρήση του συνόλου δεδομένων GSE69795 (n = 19) (δεύτερο σύνολο δοκιμής) (Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη. (Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και έβδομο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων.	127
Εικόνα 61: Διαγράμματα επιβίωσης Kaplan – Meier της προβλεπτικής υπογραφής των έξι γονιδίων, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών με υψηλή και χαμηλή έκφραση για τα έξι αυτά γονίδια, αντίστοιχα. Οι ασθενείς ομαδοποιήθηκαν (n = 121 σε κάθε ομάδα) σύμφωνα με τα προσαρμοσμένα χαμηλά και υψηλά κατώφλια αποκοπής, 30% και 70%, αντίστοιχα, όσον αφορά τη σύγκριση της συνολικής τους επιβίωσης και της ελεύθερης νόσου επιβίωσής τους.	128
Εικόνα 62: Η ανάλυση των επιπέδων της γονιδιακής έκφρασης των εννέα προτεινόμενων βιοδεικτών σε ασθενείς με ΚΟΚ, όπως δημιουργήθηκε με τη χρήση της πλατφόρμας GEPIA2. Τα κόκκινα πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης mRNA σε ιστούς ΚΟΚ και τα γκρι πλαίσια	

αντιπροσωπεύουν τα επίπεδα έκφρασης σε υγιής ιστούς ελέγχου της ουροδόχου κύστης από ασθενείς των κοορτών TCGA και GTEx. Ως κριτήρια αποκοπής για τη στατιστική σημαντικότητα εφαρμόστηκαν τα: τιμή- $p < 0,05$ και $ \log_2FC < 1$. BLCA: καρκίνος της ουροδόχου κύστης, TPM: Αριθμός μεταγραφών ανά εκατομμύριο.....	130
Εικόνα 63: Η ανάλυση του επιπέδου της γονιδιακής έκφρασης των γονιδίων <i>ANXA5</i> και <i>COL3A1</i> σε ασθενείς με KOK για τους μη-θηλώδεις και θηλώδεις υπότυπους, που δημιουργήθηκε με τη χρήση της πλατφόρμας GEPIA2. Τα κόκκινα πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης mRNA σε ιστούς υποτύπων KOK και τα γκρι πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης σε υγιής ιστούς ελέγχου της ουροδόχου κύστης από ασθενείς των κοορτών TCGA και GTEx. Ως κριτήρια αποκοπής για τη στατιστική σημαντικότητα εφαρμόστηκαν τα: τιμή- $p < 0,05$ και $ \log_2FC < 1$. BLCA: καρκίνος της ουροδόχου κύστης, TPM: Αριθμός μεταγραφών ανά εκατομμύριο.	130
Εικόνα 64: Διαγράμματα βιολιού που αντιπροσωπεύουν τη συσχέτιση μεταξύ των επιπέδων έκφρασης των εννέα προτεινόμενων βιοδεικτών – γονιδίων και των τριών παθολογικών σταδίων του KOK σε ασθενείς με KOK, τα οποία δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2 και με βάση την κοόρτη TCGA. Οι τιμές F και $Pr(>F)$ υποδηλώνουν την τιμή του στατιστικού F-τεστ και την τιμή- p του τεστ, αντίστοιχα.....	131
Εικόνα 65: Ανοσοϊστοχημική (IHC) επικύρωση των εννέα προτεινόμενων βιοδεικτών – γονιδίων σε δείγματα ανθρώπινου καρκινικού και φυσιολογικού ιστού της ουροδόχου κύστης, που προέρχονται από το Human Protein Atlas (www.proteinatlas.org). Η χρώση έδειξε ότι η πρωτεϊνική έκφραση των εννέα βιοδεικτών – γονιδίων ήταν υψηλότερη στους ιστούς με KOK σε σύγκριση με δείγματα φυσιολογικών ιστών ουροδόχου κύστης.	133
Εικόνα 66: Επικύρωση των διαγνωστικών μοντέλων, χρησιμοποιώντας τα εννέα γονίδια – βιοδείκτες ως χαρακτηριστικά, με ανάλυση καμπύλης ROC σε κάθε ένα από τα σύνολα δεδομένων που συμπεριλήφθηκαν στην ενοποιητική μετα-ανάλυση και περιείχαν τουλάχιστον 10 δείγματα (Πίνακας 2).....	135
Εικόνα 67: Επικύρωση των διαγνωστικών μοντέλων, χρησιμοποιώντας τα εννέα γονίδια – βιοδείκτες ως χαρακτηριστικά, με ανάλυση καμπύλης ROC στο ενοποιημένο μετα-σύνολο δεδομένων ($n = 606$) και στο εξωτερικό σύνολο δεδομένων επικύρωσης E-MTAB-1650 ($n = 30$).....	135

Αυτή η σελίδα είναι σκόπιμα λευκή.

Λίστα Πινάκων

Πίνακας 1: Κοινοί τύποι αρχείων πρωτογενών δεδομένων μικροσυστοιχιών.	61
Πίνακας 2: Τα χαρακτηριστικά των 18 σειρών / συνόλων δεδομένων που συμπεριλήφθηκαν στην παρούσα ενοποιητική μετα-ανάλυση.	93
Πίνακας 3: Παράμετροι απόδοσης των διαφόρων μοντέλων ταξινόμησης για τα εκάστοτε σύνολα διαφορικά εκφραζόμενων γονιδίων.....	97
Πίνακας 4: Τα γονίδια – κόμβοι με βάση τις μεθόδους ανάλυσης των προσθέτων <i>cytoHubba</i> και <i>MCODE</i> του λογισμικού <i>Cytoscape</i>	105
Πίνακας 5: Τα γονίδια – κόμβοι – κλειδιά (key hub genes) της παρούσας μελέτης, τα οποία αποτελούν την τομή μεταξύ των γονιδίων – κόμβων που προέκυψαν από την ανάλυση του δικτύου ΑΠΠ και των γονιδίων – κόμβων που προέκυψαν από την ανάλυση των δικτύων συνέκφρασης των γονιδίων.....	110
Πίνακας 6: Μονο-μεταβλητή και πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox των γονιδίων – κλειδιών που σχετίζονται με την επιβίωση των ασθενών με ΚΟΚ HR: λόγος κινδύνου (Hazard ratio), CI: Διάστημα εμπιστοσύνης (Confidence interval) *τιμή-p < 0.05, **τιμή-p < 0.01, ***τιμή-p < 0.001, ****τιμή-p < 0.0001	115
Πίνακας 7: Μονο-μεταβλητή και πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox των γονιδίων – κλειδιών που σχετίζονται με την επιβίωση των ασθενών με μυοδιηθητικό ΚΟΚ που λάμβαναν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη HR: λόγος κινδύνου (Hazard ratio), CI: Διάστημα εμπιστοσύνης (Confidence interval) *τιμή-p < 0.05, **τιμή-p < 0.01, ***τιμή-p < 0.001, ****τιμή-p < 0.0001	125

Αυτή η σελίδα είναι σκόπιμα λευκή.

Ευρετήριο όρων

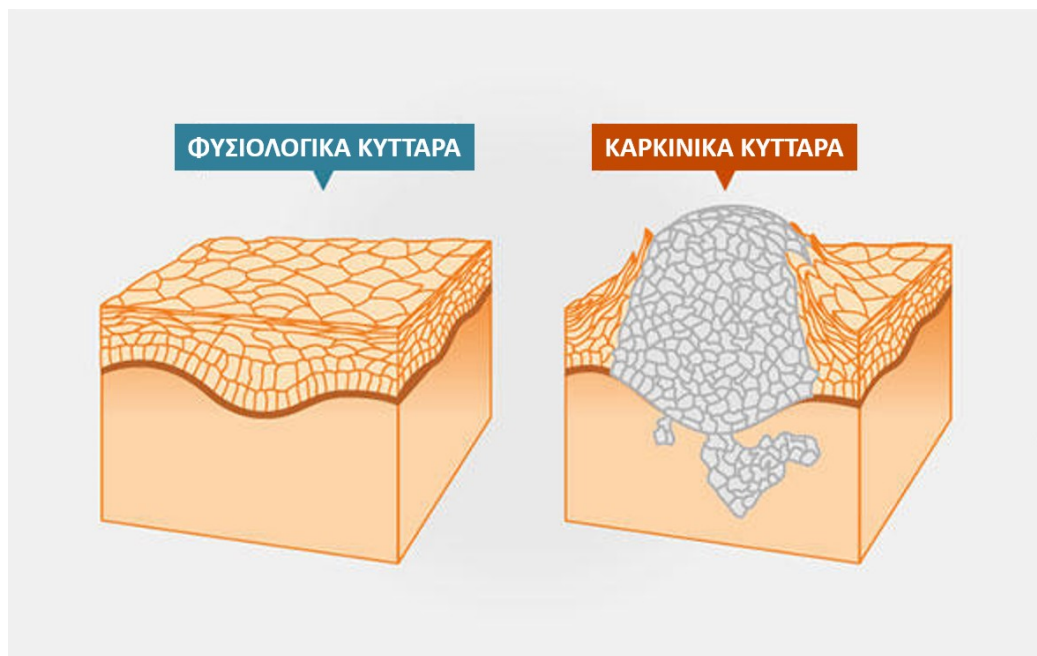
Ελληνικός όρος		Αγγλικός όρος	
ΑΔΣΓΣ	Ανάλυση Δικτύων Σταθμισμένης Γονιδιακής Συνέκφρασης	WGCNA	Weighted gene co-expression network analysis
ΑΚΣ	Ανάλυση Κύριων Συνιστωσών	PCA	Principal Component Analysis
ΑΣΜ	Ανθεκτική Συνθετική Μέθοδος	RRA	Robust Rank Aggregation
ΑΜ – ΑΑΠ	Αλυσιδωτή αντίδραση πολυμεράσης με αντίστροφη μεταγραφή	RT – PCR	Reverse Transcription - Polymerase Chain Reaction
ΑΑΠ	Ατλαντας των Ανθρώπινων Πρωτεϊνών	HPA	Human Protein Atlas
--	Βάκιλος Calmette-Guérin	BCG	Bacillus Calmette-Guérin
ΓΑΚ	Γονιδιακός Άτλαντας του Καρκίνου	TCGA	The Cancer Genome Atlas
ΔΑΑ	Δείκτη Ανθρώπινης Ανάπτυξης	HDI	Human Development Index
--	Διαδραστική Ανάλυση του Προφίλ της Γονιδιακής Έκφρασης	GEPIA2	Gene Expression Profiling Interactive Analysis
ΔΕ	Διάστημα Εμπιστοσύνης	CI	Confidence Interval
ΑΠΠ	Αλληλεπίδραση Πρωτεΐνης – Πρωτεΐνης	PPI	Protein – Protein Interaction
ΔΕΟΟΚ	Διουρηθρική εκτομή του όγκου της ουροδόχου κύστης	TURBT	TransUrethral Resection of a Bladder Tumor
ΕΚΓΤ	Εγκυκλοπαίδεια του Κιότο για τα Γονίδια και τα Γονιδιώματα	KEGG	Kyoto Encyclopedia of Genes and Genomes
ΕΝΕ	Ειδική της Νόσου Επιβίωση	CSS	Cancer Specific Survival
ΕΑΧΣΕ	Ελάχιστος Απόλυτος Χειριστής Συρρίκνωσης και Επιλογής	LASSO	Least Absolute Shrinkage and Selection Operator
ΕΥΕ	Ελεύθερη Υποτροπής Επιβίωση	RFS	Recurrence Free Survival
ΚΟΚ	Καρκίνος Ουροδόχου Κύστης	BCa	Urinary Bladder Cancer
ΚΕΠΝ	Κέντρο Ελέγχου και Πρόληψης Νοσημάτων	CDC	Centers for Disease Control and Prevention
ΛΧΕ	Λειτουργικός Χαρακτηριστικός Δέκτης	ROC	Receiver Operating Characteristic
ΛΚ	Λόγος Κινδύνου	HR	Hazard Ratio
ΜΣΚ	Μέθοδος Συνταιριασμού Καμπύλης	LOWESS	LOcally WEighted Scatterplot Smoothing
ΜΣΟΓ	Μελέτη Συσχέτισης σε Ολόκληρο το Γονιδίωμα	GWAS	Genome-Wide Association Study
ΜΜΚΟΚ	Μη Μυοδιηθητικός Καρκίνος της Ουροδόχου Κύστης	NMIBC	Non-Muscle Invasive Bladder Cancer
ΜΔΥ	Μηχανές Διανυσμάτων Υποστήριξης	SVM	Support Vector Machines

ΜΚΟΚ	Μυοδιθητικός Καρκίνος της Ουροδόχου Κύστης	MIBC	Muscle Invasive Bladder Cancer
NEX	Νεοεπικουρική Χημειοθεραπεία	NAC	Neoadjuvant chemotherapy
ΟΑ	Οντολογία Ασθενειών	DO	Disease Ontology
ΟΓ	Οντολογία Γονιδίων	GO	Gene Ontology
ΠΟΥ	Παγκόσμιος Οργανισμός Υγείας	WHO	World Health Organization
ΠΚΚ	Περιοχής Κάτω από την Καμπύλη	AUC	Area Under the Curve
ΠΤΕ	Πίνακας Τοπολογικής Επικάλυψης	TOM	Topological Overlap Matrix
ΠΗΕ	Προτυποποιημένη κατά Ηλικία Επίπτωση	ASI	Age-Standardized Incidence
--	Συμπληρωματικό DNA	cDNA	complementary DNA
ΣΕ	Συνολική Επιβίωση	OS	Overall Survival
ΤΕΜ	Τυχαιοποιημένη Ελεγχόμενη Μελέτη	RCT	Randomized Controlled Trial
ΨΠΑ	Ψευδές Ποσοστό Ανακάλυψης	FDR	False Discovery Rate
--	--	BH	Benjamini – Hochberg
--	--	GEO	Gene Expression Omnibus

1. Εισαγωγή

1.1 Καρκίνος

Ο **καρκίνος** είναι ένας γενικός όρος ο οποίος περιγράφει μια μεγάλη ομάδα ασθενειών που μπορεί να επηρεάσουν οποιοδήποτε μέρος του σώματος. Εναλλακτικοί όροι που χρησιμοποιούνται για να περιγράψουν το σύνολο αυτών των ασθενειών είναι οι κακοήθεις όγκοι και τα νεοπλάσματα. Ένα καθοριστικό χαρακτηριστικό του καρκίνου είναι η **ταχεία δημιουργία μη φυσιολογικών κυττάρων** που αναπτύσσονται πέρα από τα συνηθισμένα όριά τους και τα οποία μπορούν, στη συνέχεια, να εισβάλουν σε παρακείμενα μέρη του σώματος και να εξαπλωθούν σε άλλα όργανα, διαδικασία που χαρακτηρίζεται ως **μετάσταση** (Εικόνα 1). Οι ευρέως διαδεδομένες μεταστάσεις στο ανθρώπινο σώμα είναι η κύρια αιτία θανάτου από καρκίνο.



Εικόνα 1: Ο καρκίνος είναι μια ασθένεια που προκαλείται όταν τα κύτταρα διαιρούνται ανεξέλεγκτα και εξαπλώνονται στους γύρω ιστούς.

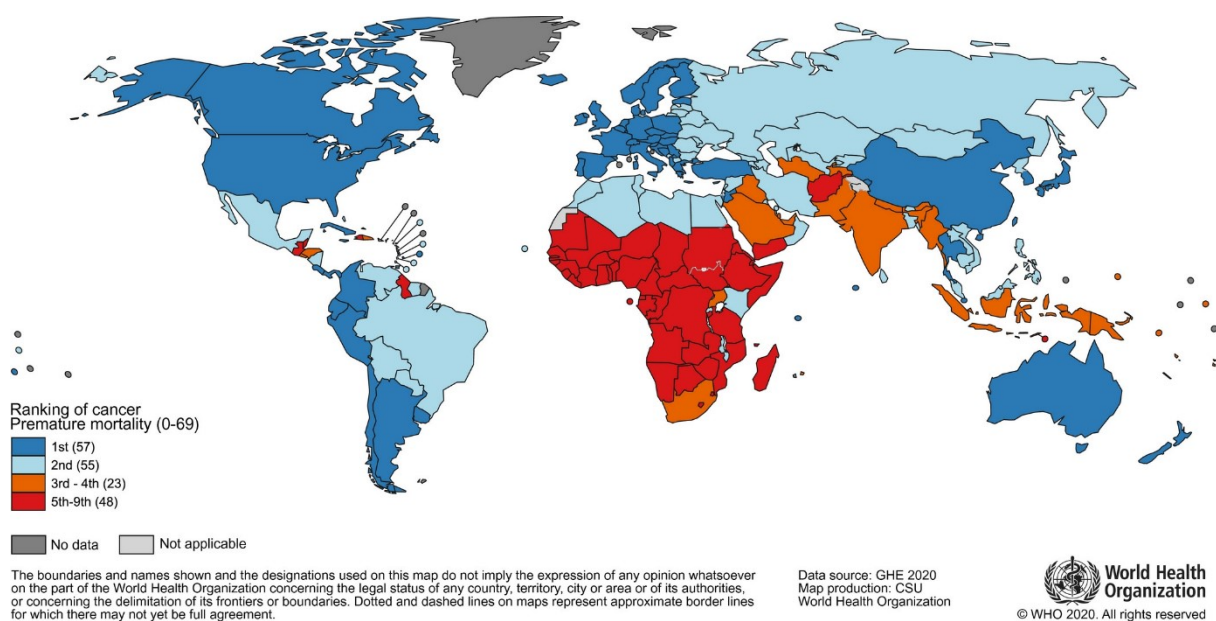
(Πηγή: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>)

Επιδημιολογία

Ο καρκίνος αποτελεί μία από τις **κύριες αιτίες θανάτου παγκοσμίως** και σημαντικό εμπόδιο για την αύξηση του προσδόκιμου ζωής σε κάθε χώρα του σύγχρονου κόσμου. Σύμφωνα με εκτιμήσεις του Παγκόσμιου Οργανισμού Υγείας (ΠΟΥ) για το 2019, ο καρκίνος είναι η πρώτη ή δεύτερη κύρια αιτία θανάτου, πριν από την ηλικία των 70 ετών, σε 112 από τις 183 χώρες του κόσμου και κατατάσσεται στην τρίτη ή τέταρτη θέση σε άλλες 23 χώρες (Εικόνα 2) [1]. Η αυξανόμενη προβολή του καρκίνου ως κύρια αιτία θανάτου αντανακλάται, εν μέρει, και από τη σημαντική μείωση των ποσοστών θνησιμότητας λόγω εγκεφαλικών επεισοδίων και στεφανιαίας νόσου συγκριτικά με τον καρκίνο σε πολλές χώρες.

Η παγκόσμια επιβάρυνση από τον καρκίνο εκτιμάται ότι αυξήθηκε σε 18,1 εκατομμύρια νέες περιπτώσεις και 9,6 εκατομμύρια νέους θανάτους το 2018 [2]. Το GLOBOCAN (Παγκόσμιο

Παρατηρητήριο Καρκίνου του Διεθνούς Οργανισμού Ερευνών για τον Καρκίνο του Παγκόσμιου Οργανισμού Υγείας) προβλέπει ότι θα υπάρχουν 27,5 εκατομμύρια νέα περιστατικά καρκίνου παγκοσμίως κάθε χρόνο έως το 2040, δηλαδή θα σημειωθεί μία αύξηση κατά 61,7% [1]. Μεγάλο μέρος αυτής της προβλεπόμενης αύξησης οφείλεται σε δημογραφικούς παράγοντες, όπως η αύξηση του παγκόσμιου πληθυσμού, η γήρανση του πληθυσμού, αλλά και η αυξανόμενη επικράτηση ορισμένων τύπων καρκίνων που συνδέονται με την κοινωνική και οικονομική ανάπτυξη. Αναμένεται να υπάρξει μετατόπιση από τους καρκίνους που σχετίζονται με λοιμώξεις σε καρκίνους που σχετίζονται με τις επιλογές του δυτικού τρόπου ζωής. Επί του παρόντος, ο συχνότερα διαγνωσμένος καρκίνος και η κύρια αιτία καρκίνου θανάτου ποικίλλουν μεταξύ των χωρών και εντός της κάθε χώρας ανάλογα με το βαθμό της οικονομικής ανάπτυξης και τους συναφείς κοινωνικούς δείκτες και παράγοντες του τρόπου ζωής. Εξετάζοντας τα στοιχεία για την επίπτωση και τη θνησιμότητα του καρκίνου σε σχέση με το επίπεδο εισοδήματος των διαφορετικών χωρών καταδεικνύει περαιτέρω αυτή την τάση. Οι χώρες με υψηλό εισόδημα έχουν τα υψηλότερα ποσοστά επίπτωσης για όλους τους καρκίνους (συμπεριλαμβανομένου του καρκίνου του πνεύμονα, του παχέος εντέρου, του μαστού και του προστάτη). Ωστόσο, αξίζει να τονιστεί ότι στις χώρες αυτές τα ποσοστά θνησιμότητας μειώνονται. Οι χώρες χαμηλού και μεσαίου εισοδήματος έχουν υψηλότερα ποσοστά καρκίνου του στομάχου, του ήπατος, του οισοφάγου και του τραχήλου της μήτρας, αντιπροσωπεύοντας δυσανάλογο βάρος καρκίνων που σχετίζονται με λοιμώξεις, και τα ποσοστά θνησιμότητας παραμένουν υψηλά εξαιτίας της έλλειψης πρόσβασης σε αποτελεσματικό έλεγχο και θεραπεία.



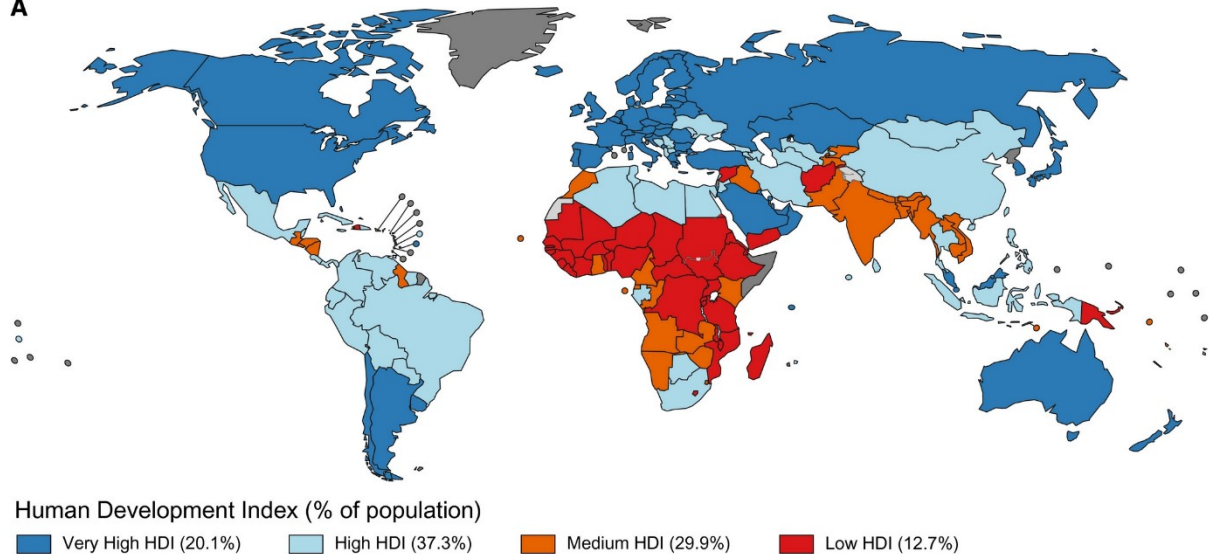
Εικόνα 2: Εθνική κατάταξη του καρκίνου ως αιτία θανάτου σε ηλικίες < 70 ετών για το έτος 2019.

(Πηγή: <https://gco.iarc.fr/today/online-analysis>)

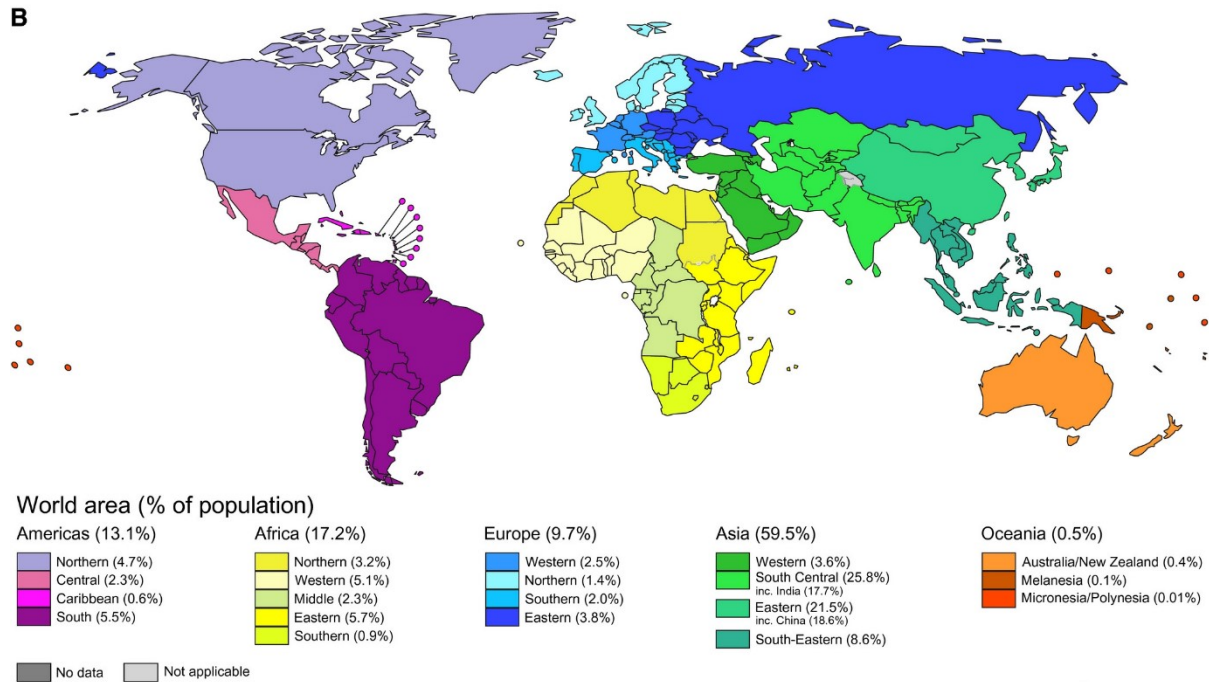
Είναι προφανές λοιπόν πως η επιβάρυνση από τη συχνότητα εμφάνισης και τη θνησιμότητα του καρκίνου αυξάνεται με ταχείς ρυθμούς παγκοσμίως. Το γεγονός αυτό αντανakλά τόσο τη γήρανση και την αύξηση του παγκόσμιου πληθυσμού όσο και τις αλλαγές στον επιπολασμό και την κατανομή των κύριων παραγόντων κινδύνου για τον καρκίνο, αρκετοί από τους οποίους συνδέονται με την κοινωνικοοικονομική ανάπτυξη. Ο βαθμός στον οποίο η θέση του καρκίνου ως αιτία πρόωρου θανάτου αντανakλά τα εθνικά επίπεδα κοινωνικής και οικονομικής ανάπτυξης μπορεί να γίνει αντιληπτός συγκρίνοντας τους χάρτες που παρουσιάζονται στη συνέχεια (Εικόνα 2 και Εικόνα 3). Η τελευταία

εικόνα απεικονίζει τον τετραβάθμιο Δείκτη Ανθρώπινης Ανάπτυξης (Human Development Index – HDI) με βάση την Έκθεση Ανθρώπινης Ανάπτυξης των Ηνωμένων Εθνών για το 2019.

A



B



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: UNDP
Map production: IARC
World Health Organization

World Health Organization
© WHO 2020. All rights reserved

Εικόνα 3: (Α) Ο 4-βάθμιος δείκτης ανθρώπινης ανάπτυξης (HDI) και (Β) 20 περιοχές του κόσμου με τα αντίστοιχα ποσοστά πληθυσμού.

(Πηγή: <https://gco.iarc.fr/today/online-analysis>)

Αίτια και Παράγοντες Κινδύνου

Ο καρκίνος είναι μια **γενετική ασθένεια**, δηλαδή προκαλείται από αλλαγές στα γονίδια που ελέγχουν τον τρόπο λειτουργίας των κυττάρων (Εικόνα 4), ιδίως τον τρόπο με τον οποίο αναπτύσσονται και διαιρούνται.

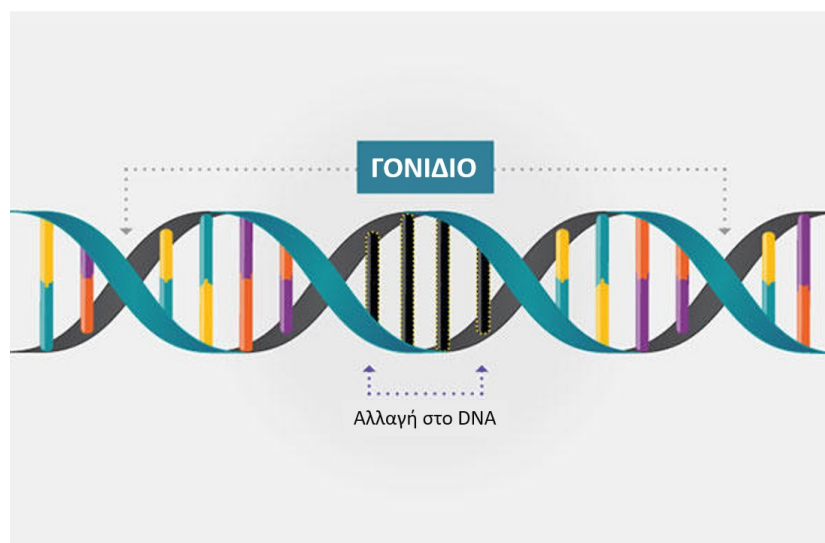
Οι γενετικές αλλαγές που προκαλούν τον καρκίνο μπορεί να συμβούν από:

- λάθη που είναι πιθανό να ανακύψουν καθώς τα κύτταρα διαιρούνται.

- βλάβες στο DNA οι οποίες προκαλούνται από επιβλαβείς ουσίες στο εξωτερικό περιβάλλον.
- κληρονομημένες μεταλλάξεις από τους γονείς.

Έτσι, ο καρκίνος προκύπτει από τον εκφυλισμό φυσιολογικών κυττάρων σε καρκινικά κύτταρα σε μια **διαδικασία πολλαπλών σταδίων** που συνήθως εξελίσσεται από μια προκαρκινική βλάβη σε έναν κακοήγη όγκο. Οι αλλοιώσεις αυτές είναι συχνά το αποτέλεσμα της αλληλεπίδρασης μεταξύ των γενετικών παραγόντων ενός ατόμου και τριών κατηγοριών εξωτερικών παραγόντων, οι οποίοι περιλαμβάνουν τις:

- **φυσικές** καρκινογόνες ουσίες, όπως η υπεριώδης και η ιονίζουσα ακτινοβολία.
- **χημικές** καρκινογόνες ουσίες, όπως ο αμιάντος, τα συστατικά του καπνού του τσιγάρου, το αλκοόλ, η αφλατοξίνη (μολυσματικός παράγοντας των τροφίμων) και το αρσενικό (μολυσματικός παράγοντας του πόσιμου νερού).
- **βιολογικές** καρκινογόνες ουσίες, όπως οι λοιμώξεις από ορισμένους ιούς, βακτήρια ή παράσιτα.



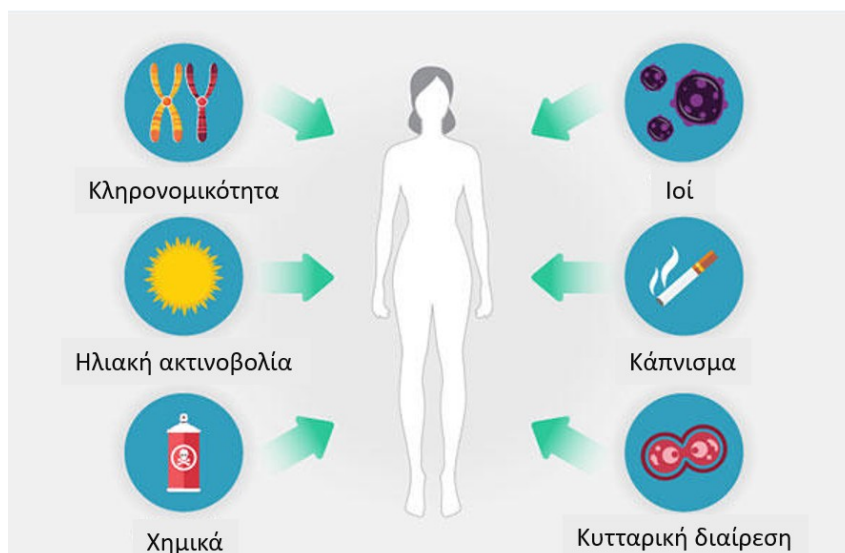
Εικόνα 4: Ο καρκίνος προκαλείται από αλλαγές στο DNA. Οι περισσότερες αλλαγές στο DNA που προκαλούν καρκίνο συμβαίνουν σε τμήματα του DNA που ονομάζονται γονίδια. Οι αλλαγές αυτές ονομάζονται επίσης γενετικές αλλαγές.

(Πηγή: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>)

Ο καρκίνος κάθε ατόμου έχει έναν μοναδικό συνδυασμό γενετικών αλλαγών. Καθώς ο καρκίνος συνεχίζει να αναπτύσσεται, θα προκύψουν πρόσθετες αλλαγές. Ακόμη και μέσα στον ίδιο όγκο, διαφορετικά κύτταρα μπορεί να έχουν διαφορετικές γενετικές αλλαγές. Οι καρκινικοί όγκοι δηλαδή μπορεί να παρουσιάζουν σημαντική **ετερογένεια**.

Η συχνότητα εμφάνισης καρκίνου **αυξάνεται σημαντικά με την ηλικία**, πιθανότατα λόγω της συσσώρευσης παραγόντων κινδύνου για συγκεκριμένους τύπους καρκίνου οι οποίοι αυξάνονται και αυτοί με την ηλικία. Η συνολική συσσώρευση των παραγόντων κινδύνου συνδυάζεται με την τάση της **μείωσης της αποτελεσματικότητας των μηχανισμών κυτταρικής επιδιόρθωσης** καθώς το άτομο μεγαλώνει. Ο οργανισμός δηλαδή χάνει την αποτελεσματική του ιδιότητα να εξαλείφει τα κύτταρα με κατεστραμμένο DNA πριν αυτά γίνουν καρκινικά.

Μερικοί από τους κυριότερους τροποποιήσιμους παράγοντες κινδύνου για την ανάπτυξη καρκίνου, καθώς και άλλων μη μεταδοτικών ασθενειών (π.χ. καρδιαγγειακών νοσημάτων), αποτελούν η χρήση καπνού, η κατανάλωση αλκοόλ, η ανθυγιεινή διατροφή, η σωματική αδράνεια και η έκθεση σε περιβάλλοντα με αυξημένη ατμοσφαιρική ρύπανση (Εικόνα 5).



Εικόνα 5: Οι γενετικές αλλαγές που προκαλούν καρκίνο μπορεί να είναι κληρονομικές ή να οφείλονται σε ορισμένους περιβαλλοντικούς παράγοντες. Οι γενετικές αλλαγές μπορούν επίσης να συμβούν λόγω σφαλμάτων που συμβαίνουν κατά τη διαίρεση των κυττάρων.

(Πηγή: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>)

Επίσης, ορισμένες χρόνιες λοιμώξεις αποτελούν και αυτές παράγοντες κινδύνου για την ανάπτυξη καρκίνου. Η συμβολή των παραγόντων αυτών αποτελεί ένα σημαντικό πρόβλημα σε χώρες χαμηλού και μεσαίου εισοδήματος. Περίπου το 13% των καρκίνων που διαγνώστηκαν το 2018 σε παγκόσμια κλίμακα αποδίδονταν σε καρκινογόνες λοιμώξεις, συμπεριλαμβανομένων του ελικοβακτηριδίου του πυλωρού, του ιού των ανθρώπινων θηλωμάτων (Human Papilloma Virus – HPV), του ιού της ηπατίτιδας Β, του ιού της ηπατίτιδας C και του ιού Epstein-Barr [3]. Αναλυτικότερα, οι ιοί της ηπατίτιδας Β και C και ορισμένοι τύποι HPV αυξάνουν τον κίνδυνο για καρκίνο του ήπατος και του τραχήλου της μήτρας, αντίστοιχα. Η μόλυνση από τον ιό της ανθρώπινης ανοσοανεπάρκειας (Human Immunodeficiency Virus – HIV) αυξάνει σημαντικά τον κίνδυνο εμφάνισης καρκίνου του τραχήλου της μήτρας, κατά περίπου έξι φορές, καθώς και τον κίνδυνο εμφάνισης άλλων μορφών καρκίνου, όπως το σάρκωμα Kaposi.

Όσον αφορά τους τροποποιήσιμους παράγοντες κινδύνου, εκτιμάται ότι το 45% των 607.000 θανάτων από καρκίνο που καταγράφηκαν στις ΗΠΑ το 2019 σχετιζόνταν με τροποποιήσιμους παράγοντες κινδύνου [4]. Αυτό καθιστά αυτούς τους τροποποιήσιμους παράγοντες κινδύνου μερικά από τα πιο σημαντικά κριτήρια επίπτωσης και τάσεων που μπορεί κανείς να παρακολουθεί όταν εξετάζει τρόπους για τη μείωση της επιβάρυνσης του καρκίνου στον πληθυσμό. Οι καρκίνοι του πνεύμονα, του μαστού, του προστάτη, του παχέος εντέρου και του ορθού αντιστοιχούν σε περισσότερο από το ήμισυ της συνολικής επιβάρυνσης από καρκίνο στις πιο ανεπτυγμένες περιοχές του κόσμου, όπως όπως η Βόρεια Αμερική. Ο αθροιστικός δια βίου κίνδυνος ενός ατόμου για την ανάπτυξη οποιοδήποτε από αυτούς τους τέσσερις καρκίνους ήταν πάνω από 3% στις πιο ανεπτυγμένες από ό,τι στις λιγότερο ανεπτυγμένες περιοχές [5]. Είναι λογικό να υποθέσει κανείς βέβαια ότι ένα μέρος αυτής της ανισότητας θα μπορούσε να αποδοθεί στην φυλετική/γενετική σύνθεση αυτών των πληθυσμών. Ωστόσο, η εξήγηση για το μεγαλύτερο μέρος αυτής της ανισότητας είναι πιθανότατα η συσχέτιση με τη διαφορετική περιβαλλοντική έκθεση των πληθυσμών. Οτιδήποτε στο οποίο εκτίθεται κάποιος κατά τη διάρκεια της ζωής του και με το οποίο δεν γεννιέται θεωρείται **τροποποιήσιμος παράγοντας κινδύνου** και αποτελεί μία δυνητική ευκαιρία για την πρόληψη του καρκίνου. Η μείωση των τροποποιήσιμων παραγόντων κινδύνου, που ενοχοποιούνται για την ανάπτυξη του καρκίνου, μαζί με τη βελτίωση του

προσυμπτωματικού ελέγχου απαιτεί βελτιωμένες κυβερνητικές πολιτικές, κοινωνικές προσπάθειες και αλλαγές σε ατομικές συμπεριφορές. Πολλές από αυτές τις στρατηγικές έχουν αποδειχθεί αποτελεσματικές, αλλά η εφαρμογή τους μέχρι στιγμής δεν είναι η βέλτιστη. Όπως εύστοχα τονίζουν η Susan Gapstur κ.α. σε πρόσφατη έρευνά τους: «Είναι ευθύνη των κυβερνήσεων, της βιομηχανίας και της δημόσιας υγείας, της ιατρικής και της επιστημονικής κοινότητας να συνεργαστούν με σκοπό να επενδύσουν σε ένα ολοκληρωμένο σχέδιο διαχείρισης και ελέγχου για τον καρκίνο και να το εφαρμόσουν σε εθνικό επίπεδο. Εάν αυτή η προσπάθεια αποτύχει, η ανθρωπότητα θα επιβραδύνει την πρόοδο στις εθνικές προσπάθειες για τη μείωση της επιβάρυνσης από τον καρκίνο» [6].

Όπως προαναφέρθηκε, το **κάπνισμα** αποτελεί έναν από τους βασικούς παράγοντες κινδύνου για την ανάπτυξη καρκίνου. Η εκτιμώμενη ετήσια απώλεια παραγωγικότητας στην εργασία που αποδίδεται στη χρήση καπνού στις ΗΠΑ είναι 156 δισεκατομμύρια δολάρια ετησίως, ενώ το άμεσο κόστος υγειονομικής περίθαλψης ανέρχεται στο ιλιγγιώδες ποσό των 170 δισεκατομμυρίων δολαρίων [7]. Υπάρχουν σημαντικές ενδείξεις ότι ο έλεγχος και περιορισμός του καπνίσματος μπορεί να αποτρέψει περισσότερους θανάτους από καρκίνο από οποιαδήποτε άλλη στρατηγική πρωτογενούς πρόληψης.

Μετά τη χρήση του καπνού, οι δεύτεροι πιο επιδραστικοί τροποποιήσιμοι παράγοντες κινδύνου για καρκίνο σχετίζονται με τη **διατροφή** και τον **τρόπο ζωής**. Μεταξύ του 1980 και του 2014, ο επιπολασμός του **υπέρβαρου ή παχύσαρκου πληθυσμού** στις ΗΠΑ αυξήθηκε κατά περισσότερο από 100% μεταξύ των παιδιών και των εφήβων και κατά 60% μεταξύ των ενηλίκων ηλικίας 20 έως 74 ετών [8]. Μια πιο εμπεριστατωμένη ανάλυση δείχνει ότι στην πραγματικότητα το ποσοστό των ανδρών και των γυναικών των ΗΠΑ που θεωρούνται υπέρβαροι είναι σταθερό, τουλάχιστον τα τελευταία χρόνια, στο 36,5% των ανδρών και στο 26,9% των γυναικών [9]. Ωστόσο, ο επιπολασμός της παχυσαρκίας έχει αυξηθεί δραστικά, με επτά πολιτείες να εμφανίζουν ποσοστά παχυσαρκίας τουλάχιστον 35% το 2017, από μηδέν πολιτείες το 2012. Τα ποσοστά αυτά διαφέρουν ανάλογα με την εθνικότητα, με τα υψηλότερα ποσοστά παχυσαρκίας να εμφανίζονται στους Αφροαμερικανούς (54,9%) και τα χαμηλότερα στους Ασιάτες (14,8%). Σε ενήλικες ηλικίας 30 ετών και άνω στις ΗΠΑ, το υπερβολικό σωματικό βάρος θα μπορούσε να ευθύνεται έως και για το 60% όλων των καρκίνων του ενδομητρίου, το 36% των καρκίνων της χοληδόχου κύστης, το 33% των καρκίνων των νεφρών, το 17% των καρκίνων του παγκρέατος και το 11% των πολλαπλών μυελωμάτων το 2014 [10, 11]. Υπάρχουν επίσης ολόένα και περισσότερα στοιχεία που υποστηρίζουν τη συσχέτιση μεταξύ της παιδικής ή εφηβικής παχυσαρκίας και του αυξημένου κινδύνου εμφάνισης καρκίνου του παχέος εντέρου, του ενδομητρίου και του παγκρέατος και πολλαπλού μυελώματος αργότερα στη ζωή [12]. Φαίνεται ότι η παχυσαρκία προκαλεί αυξημένο κίνδυνο καρκίνου στο οποιοδήποτε τελικό όργανο επηρεάζεται, κυρίως μέσω της φλεγμονής του λιπώδους ιστούς και των συναφών μεταβολών στο μικροπεριβάλλον, αλλά και συστηματικά μέσω των κυκλοφορούντων μεταβολικών και φλεγμονωδών μεσολαβητών που σχετίζονται με τη φλεγμονή του λίπους.

Το είδος της **διατροφής** που καταναλώνεται, και όχι μόνο η επίδρασή της στο σωματικό βάρος, φαίνεται επίσης να είναι σημαντική συσχέτιση με την καρκινογένεση. Μια μελέτη της γαλλικής κοόρτης NutriNet-Santé (2009-2017) κατέδειξε ότι η κατανάλωση ζαχαρούχων ποτών σχετίζεται σημαντικά με τον κίνδυνο εμφάνισης καρκίνου (λόγος κινδύνου για αύξηση 100 ml/ημέρα: 1,18 / 95% CI: 1,10-1,27 / τιμή-p < 0.0001) [13]. Η έλλειψη διαιτητικών ινών έχει επίσης συνδεθεί με υψηλότερη συχνότητα εμφάνισης καρκίνου του παχέος εντέρου τα τελευταία χρόνια. Η αυξημένη πρόσληψη φυτικών ινών μπορεί να οδηγήσει σε αραίωση των καρκινογόνων ουσιών στα κόπρανα, σε μειωμένο χρόνο διέλευσης των κοπράνων στο έντερο και σε βακτηριακή ζύμωση των φυτικών ινών σε λιπαρά οξέα βραχείας αλυσίδας με αντικαρκινικές ιδιότητες [14]. Σε μια μεγάλη προοπτική μελέτη, στην οποία όλα τα άτομα υποβλήθηκαν σε κοινό έλεγχο του παχέος εντέρου, η υψηλή πρόσληψη φυτικών ινών συσχετίστηκε με μειωμένο κίνδυνο εμφάνισης προκαρκινικών πολυπόδων αλλά και καρκίνου [15].

Σύμφωνα με το Κέντρο Ελέγχου και Πρόληψης Νοσημάτων – ΚΕΠΝ (Centers for Disease Control and Prevention – CDC), το ποσοστό των ενηλίκων στις Ηνωμένες Πολιτείες που υλοποιούν τις κατευθυντήριες γραμμές του CDC για την αερόβια **σωματική δραστηριότητα** είναι μόνο 53,3% [16]. Ακόμη, εκτιμάται ότι το 2,9% όλων των περιπτώσεων καρκίνου στις ΗΠΑ αποδίδεται σε χαμηλή σωματική δραστηριότητα, με τη συμβολή να είναι μεγαλύτερη μεταξύ των γυναικών (4,4%) από ό,τι μεταξύ των ανδρών (1,5%) [17]. Οι τύποι καρκίνου με τα υψηλότερα ποσοστά περιπτώσεων που σχετίζονται με τη χαμηλή σωματική δραστηριότητα ήταν ο καρκίνος της μήτρας (26,7%), ακολουθούμενος από τον καρκίνο του παχέος εντέρου (6,3%). Η συνήθης σωματική δραστηριότητα έχει βρεθεί ότι σχετίζεται με μειωμένη επίπτωση διαφόρων άλλων κοινών τύπων κακοήθειας, συμπεριλαμβανομένων του μαστού και του πνεύμονα. Η σωματική δραστηριότητα φαίνεται επίσης να μειώνει τη συνολική θνησιμότητα στον πληθυσμό αλλά και τη σχετική με τη νόσο θνησιμότητα στους ασθενείς με καρκίνο του μαστού και του παχέος εντέρου [18]. Φαίνεται επίσης ότι για ορισμένους τύπους καρκίνου η σωματική δραστηριότητα έχει ανεξάρτητη επίδραση στον κίνδυνο ανάπτυξης καρκίνου, όχι μόνο μέσω της ικανότητας της σωματικής δραστηριότητας να μειώνει τον δείκτη μάζας σώματος και την παχυσαρκία. Για τις έφηβες και τις νεαρές ενήλικες γυναίκες, που διαγνώστηκαν με καρκίνο του μαστού σε ηλικία 15 έως 39 ετών, έχει φανεί ότι τα υψηλότερα επίπεδα σωματικής δραστηριότητας, η χαμηλότερη πρόσληψη κόκκινου κρέατος και η υψηλότερη πρόσληψη φυτικών ινών φαίνεται να μειώνουν τον κίνδυνο εμφάνισης καρκίνου του μαστού [19]. Ωστόσο, η συσχέτιση μεταξύ της παχυσαρκίας και του κινδύνου καρκίνου του μαστού δεν ήταν εμφανής σε αυτόν τον πληθυσμό. Η συμβολή των τροποποιήσιμων παραγόντων κινδύνου στην επίπτωση του καρκίνου στον Καναδά είναι παρόμοια με εκείνη που παρατηρείται στις ΗΠΑ [20]. Σχεδόν ένας στους δύο Καναδούς αναμένεται να διαγνωστεί με καρκίνο κατά τη διάρκεια της ζωής του. Ωστόσο, υπάρχουν τρόποι για να μειωθεί ο αντίκτυπος των τροποποιήσιμων παραγόντων κινδύνου για τον καρκίνο μέσω καλά ενημερωμένων παρεμβάσεων και πολιτικών υγείας. Μια μελέτη του 2015 στον πληθυσμό εκτιμά ότι ένα ποσοστό μεταξύ 33% και 37% (έως 70.000 περιπτώσεις) των περιστατικών καρκίνου μεταξύ ενηλίκων, ηλικίας μεγαλύτερης των 30 ετών, αποδίδονταν σε παράγοντες κινδύνου που μπορούν να προληφθούν, με τη χρήση καπνού και την έλλειψη σωματικής δραστηριότητας να ευθύνονται για τα υψηλότερα ποσοστά των περιπτώσεων καρκίνου [21]. Οι καρκίνοι με τον υψηλότερο αριθμό περιπτώσεων που μπορούν να προληφθούν σε αυτή την καναδική κοόρτη ήταν ο καρκίνος του πνεύμονα, ο καρκίνος του παχέος εντέρου και ο καρκίνος του μαστού. Εάν οι τρέχουσες τάσεις στην επικράτηση των παραγόντων κινδύνου, που μπορούν να προληφθούν, συνεχιστούν και στο μέλλον, οι ερευνητές προβλέπουν ότι μέχρι το 2042 θα εκδηλώνονται περίπου 102.000 περιστατικά καρκίνου ετησίως τα οποία θα αποδίδονται σε τροποποιήσιμους παράγοντες κινδύνου, αντιπροσωπεύοντας το ίδιο ένα τρίτο όλων των περιστατικών καρκίνου που παρατηρούνται στις ΗΠΑ. Μέσω των διαφόρων παρεμβάσεων μείωσης του κινδύνου, των πολιτικών υγείας καθώς και των εκστρατειών δημόσιας υγείας, εκτιμάται ότι 10.600 έως 39.700 περιστατικά καρκίνου ετησίως στους Καναδούς πολίτες θα μπορούσαν να αποφευχθούν έως το 2042.

Αντιμετώπιση

Σε μία πρόσφατη έρευνα [4], διαπιστώνει ότι το 44,4% όλων των θανάτων από καρκίνο καθώς και το 42% των χαμένων υγιών ετών για το έτος 2019 θα μπορούσαν να αποδοθούν σε παράγοντες κινδύνου που μπορούσαν να προληφθούν. Βάσει του ΠΟΥ, μεταξύ 30 και 50% των καρκίνων μπορούν, επί του παρόντος, να προληφθούν τροποποιώντας τους παράγοντες κινδύνου με αλλαγή του τρόπου ζωής και εφαρμόζοντας τις υπάρχουσες στρατηγικές πρόληψης. Οι συνέπειες του καρκίνου μπορούν επίσης να αμβλυνθούν μέσω της **έγκαιρης ανίχνευσης και διάγνωσης** αλλά και της κατάλληλης θεραπείας και

φροντίδας των ασθενών που αναπτύσσουν καρκίνο. Πολλοί καρκίνοι έχουν μεγάλες πιθανότητες **ίασης** εάν **διαγνωστούν έγκαιρα και αντιμετωπιστούν κατάλληλα**.

Η θνησιμότητα από καρκίνο μειώνεται όταν τα περιστατικά εντοπίζονται και αντιμετωπίζονται έγκαιρα. Υπάρχουν δύο βασικοί τρόποι έγκαιρης ανίχνευσης: η **έγκαιρη διάγνωση** και ο **προσυμπτωματικός έλεγχος** (screening).

Όταν ο καρκίνος εντοπίζεται έγκαιρα είναι πιο πιθανό να ανταποκριθεί στη θεραπεία και ο ασθενής να εμφανίσει μεγαλύτερη πιθανότητα επιβίωσης με λιγότερη νοσηρότητα, καθώς και να απαιτηθεί μια λιγότερο δαπανηρή θεραπεία. Επίσης, σημαντικές είναι οι βελτιώσεις που μπορούν να γίνουν στη ζωή των καρκινοπαθών με την έγκαιρη ανίχνευση της ασθένειας και την αποφυγή καθυστερήσεων στη φροντίδα.

Η έγκαιρη διάγνωση μπορεί να επιτευχθεί μέσω της προώθησης τριών βασικών αξόνων:

- εκπαίδευση των ασθενών σχετικά με την επίγνωση των συμπτμάτων των διαφόρων μορφών καρκίνου και τη σημασία της αναζήτησης ιατρικής συμβουλής όταν παρατηρούνται μη φυσιολογικά ευρήματα.
- εξασφάλιση της πρόσβασης των ασθενών σε υπηρεσίες κλινικής αξιολόγησης και διάγνωσης.
- έγκαιρη παραπομπή των πασχόντων σε υπηρεσίες θεραπείας.

Η έγκαιρη διάγνωση των συμπτωματικών καρκίνων είναι σημαντική στην πλειοψηφία των καρκίνων. Τα προγράμματα για τον καρκίνο θα πρέπει να σχεδιάζονται για να μειώνουν τις καθυστερήσεις και τα εμπόδια στη διάγνωση, τη θεραπεία και την υποστηρικτική φροντίδα.

Ο προσυμπτωματικός έλεγχος στοχεύει στον εντοπισμό ατόμων με ευρήματα που υποδηλώνουν συγκεκριμένο καρκίνο ή προ-καρκινικό στάδιο πριν εμφανίσουν κάποια συμπτώματα. Όταν εντοπίζονται ανωμαλίες κατά τη διάρκεια του προ-συμπτωματικού ελέγχου, θα πρέπει να ακολουθήσουν περαιτέρω εξετάσεις για τον προσδιορισμό της τελικής διάγνωσης, καθώς και παραπομπή για θεραπεία, εάν αποδειχθεί ότι υπάρχει καρκίνος.

Τα προγράμματα προσυμπτωματικού ελέγχου είναι αποτελεσματικά για ορισμένους αλλά όχι για όλους τους τύπους καρκίνου. Επίσης, μπορεί κάποιες φορές να είναι πολύ περίπλοκα και να δεσμεύουν πόρους από τον τομέα της έγκαιρης διάγνωσης, καθώς απαιτούν ειδικό εξοπλισμό και ειδικευμένο προσωπικό. Ακόμη και όταν καθιερώνονται προγράμματα προσυμπτωματικού ελέγχου, εξακολουθούν να είναι απαραίτητα τα προγράμματα έγκαιρης διάγνωσης για τον εντοπισμό των περιπτώσεων καρκίνου που εμφανίζονται σε άτομα που δεν πληρούν τα κριτήρια της ηλικίας ή των παραγόντων κινδύνου για την ένταξη στα προγράμματα προσυμπτωματικού ελέγχου.

Η επιλογή ασθενών για τα προγράμματα προσυμπτωματικού ελέγχου βασίζεται στην ηλικία των υποψηφίων και τους παράγοντες κινδύνου που αυτοί εμφανίζουν για την αποφυγή υπερβολικών ψευδώς θετικών αποτελεσμάτων. Παραδείγματα προγραμμάτων προσυμπτωματικού ελέγχου αποτελούν:

- Το τεστ HPV (συμπεριλαμβανομένης της δοκιμής HPV DNA και mRNA), ως προτιμώμενη μέθοδος για τον προσυμπτωματικό έλεγχο του καρκίνου του τραχήλου της μήτρας.
- Ο μαστογραφικός έλεγχος για τον καρκίνο του μαστού για γυναίκες ηλικίας 50 – 69 ετών που διαμένουν σε περιβάλλοντα με καλά οργανωμένα συστήματα υγείας.

Είναι σαφές πως απαιτείται διασφάλιση ποιότητας τόσο για τα προγράμματα προσυμπτωματικού ελέγχου όσο και για τα προγράμματα έγκαιρης διάγνωσης.

Σημεία – Κλειδιά για τον καρκίνο

- Ο καρκίνος είναι μία από τις **κύριες αιτίες θανάτου** παγκοσμίως, αντιπροσωπεύοντας σχεδόν 10 εκατομμύρια θανάτους το 2020 ή σχεδόν έναν στους έξι θανάτους.
- Οι πιο συχνοί καρκίνοι είναι ο καρκίνος του **μαστού**, του **πνεύμονα**, του **παχέος εντέρου** και του **ορθού** και του **προστάτη**.
- Περίπου το **ένα τρίτο των θανάτων** από καρκίνο οφείλεται στη χρήση καπνικών προϊόντων, στον υψηλό δείκτη μάζας σώματος, στην κατανάλωση αλκοόλ, στη χαμηλή πρόσληψη φρούτων και λαχανικών και στην έλλειψη σωματικής δραστηριότητας.
- Οι λοιμώξεις που προκαλούν καρκίνο, όπως ο ιός των ανθρώπινων θηλωμάτων (HPV) και η ηπατίτιδα, ευθύνονται για περίπου το 30% των περιπτώσεων καρκίνου σε χώρες **χαμηλού και χαμηλού – μεσαίου εισοδήματος**.
- Πολλοί καρκίνοι **μπορούν να θεραπευτούν** εάν εντοπιστούν **έγκαιρα** και αντιμετωπιστούν **αποτελεσματικά**.

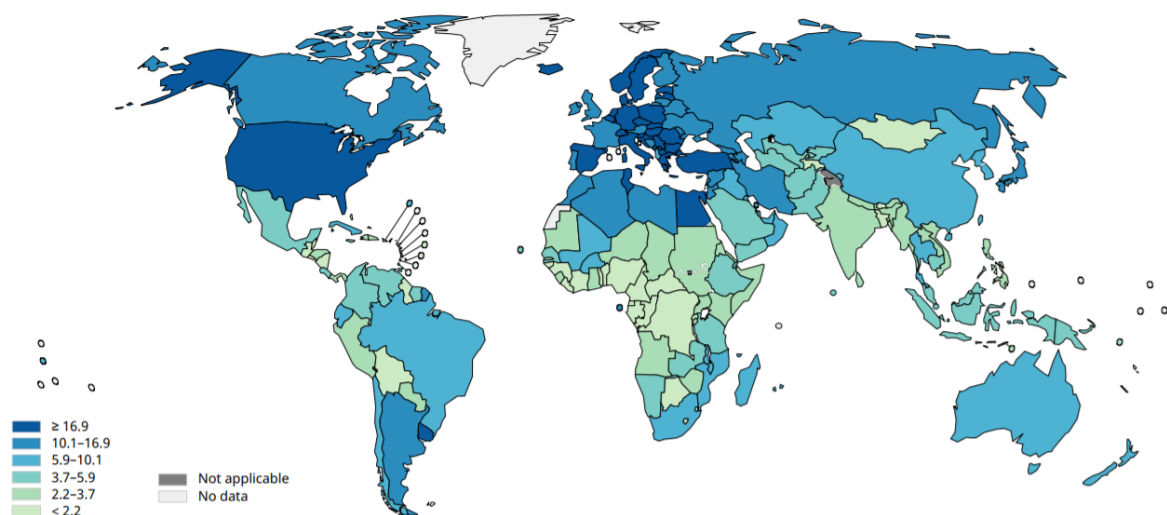
1.2 Καρκίνος ουροδόχου κύστης

Ορισμός

Ο καρκίνος της ουροδόχου κύστης – **KOK (Bladder Cancer – BCa | C67)** είναι η νόσος κατά την οποία **καρκινικά κύτταρα σχηματίζονται στα τοιχώματα της ουροδόχου κύστης**. Ο ΚΟΚ αποτελεί μία ιδιαίτερα περίπλοκη και ετερογενή ασθένεια η οποία απαιτεί εντατική επιτήρηση. Επίσης, εξαιτίας της υψηλής παγκόσμιας επίπτωσής του, του υψηλού ρυθμού επανεμφάνισής του, καθώς και της δυσμενούς πρόγνωσης στην περίπτωση της διηθητικής νόσου [22, 23], ο ΚΟΚ αποτελεί μία σημαντική επιβάρυνση για τα συστήματα υγείας και ένα σοβαρό κοινωνικοοικονομικό πρόβλημα.

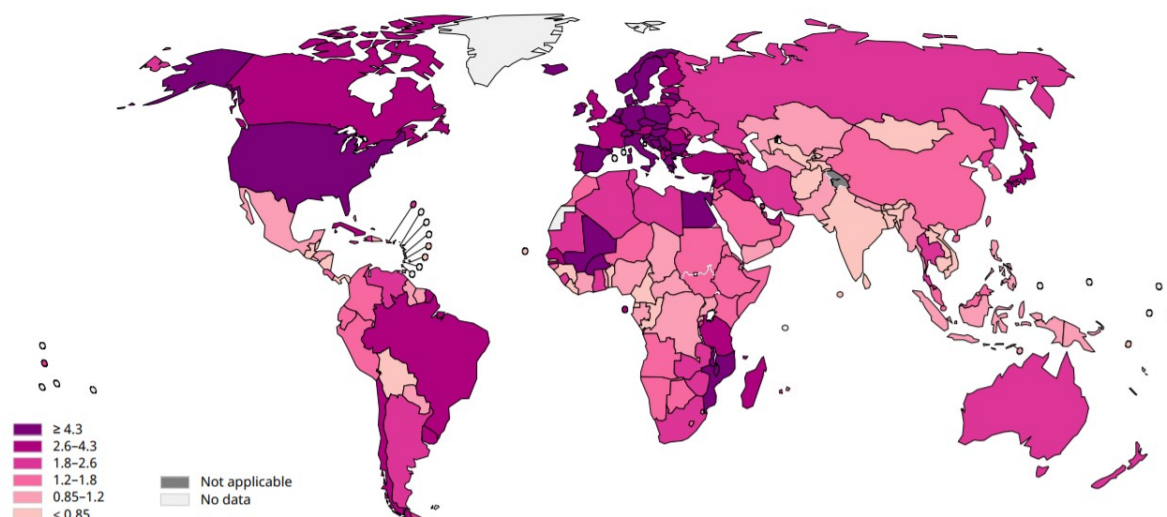
Επιδημιολογία και Παράγοντες Κινδύνου

Ο ΚΟΚ αποτελεί τη συχνότερη κακοήθεια του ουροποιητικού συστήματος. Παγκοσμίως, ήταν η **10^η πιο συχνή νεοπλασματική νόσος** και η **13^η πιο συχνή αιτία θανάτου από καρκίνο** παγκοσμίως για το 2020 [24]. Σύμφωνα με τον ΠΟΥ, το 2020 καταγράφηκαν 573.278 νέα περιστατικά ΚΟΚ και 212.536 θάνατοι λόγω της συγκεκριμένης ασθένειας. Επίσης, η Αμερικανική Αντικαρκινική Εταιρεία εκτιμά πως ο ΚΟΚ θα αποτελέσει τον **4^ο πιο συχνά εμφανιζόμενο τύπο καρκίνου** στις νέες διαγνώσεις, και τον **8^ο πιο συχνά εμφανιζόμενο τύπο στις αιτίες θανάτου** από καρκίνο για τους άνδρες στις ΗΠΑ για το φετινό έτος (2022) [25]. Πιο συγκεκριμένα και σύμφωνα με την τελευταία ετήσια έκθεση της Αμερικανικής Αντικαρκινικής Εταιρείας, υπολογίζεται ότι 81.180 ασθενείς θα διαγνωστούν με ΚΟΚ στις ΗΠΑ το 2022 και 17.100 άνθρωποι θα πεθάνουν από αυτή την ασθένεια την ίδια χρονιά [25]. Η προτυποποιημένη κατά ηλικία επίπτωση του ΚΟΚ στους άντρες και στις γυναίκες παρουσιάζεται στην Εικόνα 6 και Εικόνα 7, αντίστοιχα. Όπως φαίνεται καθαρά στις προηγούμενες εικόνες, οι άντρες πλήττονται συχνότερα σε σχέση με τις γυναίκες, με αναλογία που υπολογίζεται τουλάχιστον στο 3 προς 1 [26, 27].



Εικόνα 6: Προτυποποιημένη κατά ηλικία επίπτωση του ΚΟΚ στους άντρες παγκοσμίως.

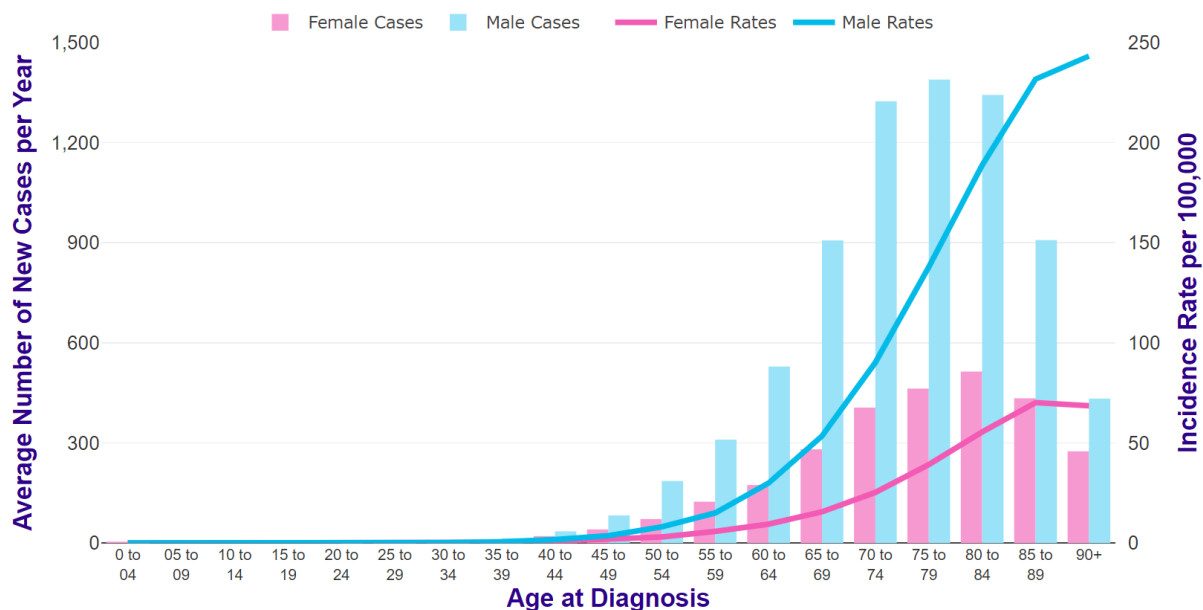
(Πηγή: <https://gco.iarc.fr/today/online-analysis>)



Εικόνα 7: Προτυποποιημένη κατά ηλικία επίπτωση του ΚΟΚ στις γυναίκες παγκοσμίως.

(Πηγή: <https://gco.iarc.fr/today/online-analysis>)

Όπως και με την πλειοψηφία των διαφόρων τύπων καρκίνου, η επίπτωση του ΚΟΚ αυξάνεται με την ηλικία. Στην Εικόνα 8 παρουσιάζεται ο μέσος αριθμός νέων διαγνώσεων ανά έτος και το ποσοστό επίπτωσης (ανά 100.000) για όλες τις ηλικιακές ομάδες [28]. Τα δεδομένα αφορούν το Ηνωμένο Βασίλειο και αναφέρονται στα έτη 2016 έως 2018. Όπως είναι εμφανές, η επίπτωση της νόσου παρουσιάζει ραγδαία αύξηση μετά την ηλικία των 65 ετών.



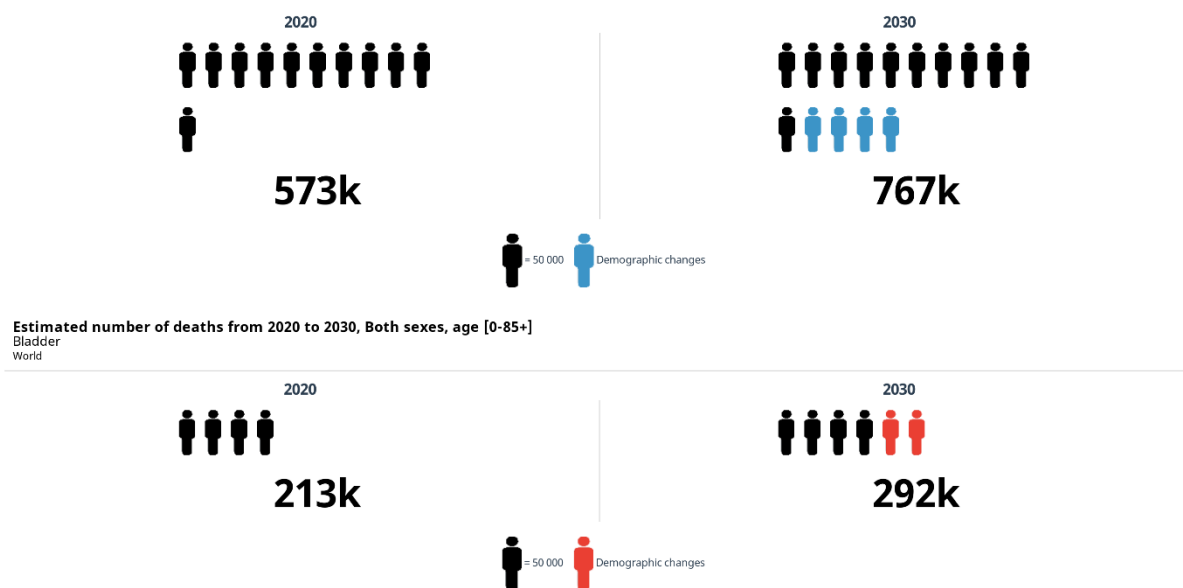
Εικόνα 8: Μέσος αριθμός νέων περιπτώσεων και επίπτωση του ΚΟΚ ανά ηλικιακή ομάδα.

(Πηγή: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/incidence>)

Πιο αναλυτικά, η επίπτωση του ΚΟΚ παρουσιάζει σημαντική διακύμανση μεταξύ διαφόρων γεωγραφικών περιοχών, με τέτοιο τρόπο ώστε η προτυποποιημένη κατά ηλικία επίπτωση (age-standardized incidence – ASI) να είναι σχεδόν τρεις φορές μεγαλύτερη στις πιο ανεπτυγμένες χώρες (ASI: 9,5/100.000) από ό,τι στις λιγότερο ανεπτυγμένες χώρες (ASI: 3,3/100.000). Όπως αναφέρθηκε σε προηγούμενη ενότητα, η τάση αυτή επαληθεύεται γενικά για τις περισσότερες μορφές καρκίνου. Το υψηλότερο ποσοστό ASI παρατηρείται στις ευρωπαϊκές χώρες (και συγκεκριμένα στην Ισπανία, την Ιταλία, τη Δανία και την Ελβετία), τη Βόρεια Αμερική, ορισμένες χώρες της βόρειας Αφρικής (για παράδειγμα, την Αίγυπτο) και τη δυτική Ασία (Τουρκία και Ισραήλ), και ιδιαίτερα στους άνδρες. Αντίθετα, τα χαμηλότερα ποσοστά σημειώνονται στην Κεντρική και Νότια Αμερική, την Υποσαχάρια Αφρική και τη νοτιοανατολική Ασία. Ο ουροθηλιακός ΚΟΚ παραμένει ο πιο συνηθισμένος τύπος ΚΟΚ στη Βόρεια Αμερική και την Ευρώπη [29].

Παρόλο που τα ποσοστά του ΚΟΚ είναι υψηλότερα στους λευκούς πληθυσμούς από ό,τι σε άλλες εθνότητες, τα ποσοστά επιβίωσης είναι χειρότερα για τους έγχρωμους ανθρώπους, σύμφωνα με έρευνες στις ΗΠΑ και το Σάο Πάολο της Βραζιλίας [30]. Δυστυχώς, οι πληροφορίες σχετικά με την επίπτωση του ΚΟΚ στους έγχρωμους πληθυσμούς σε όλο τον κόσμο είναι ελλιπείς. Τα χαμηλά ποσοστά στην Αφρική μπορούν να αποδοθούν στην ελλιπή πρόσβαση σε υπηρεσίες υγειονομικής περίθαλψης και σε ανταγωνιστικούς παράγοντες θνησιμότητας. Επιπλέον, η διακύμανση των ποσοστών θνησιμότητας από τον ΚΟΚ σε παγκόσμια βάση αποκλίνει λιγότερο συγκριτικά με την επίπτωση της νόσου. Αυτή η μειωμένη διακύμανση μπορεί να οφείλεται στον παρόμοιο τρόπο και χρόνο με τους οποίους οι ασθενείς με ΚΟΚ προχωρημένου σταδίου διαγιγνώσκονται και καταγράφονται στα συστήματα υγειονομικής περίθαλψης. Ωστόσο, οι διαφορές των ποσοστών θνησιμότητας σε τοπικό επίπεδο αντικατοπτρίζουν, τουλάχιστον εν μέρει, την άνιση πρόσβαση στα σύγχρονα συστήματα υγειονομικής περίθαλψης, τις ειδικές διαγνωστικές εγκαταστάσεις και τα πιο σύγχρονα πρωτόκολλα θεραπείας. Ευτυχώς, σε παγκόσμιο επίπεδο, τα προτυποποιημένα κατά ηλικία ποσοστά θνησιμότητας από ΚΟΚ έχουν μειωθεί τα τελευταία χρόνια, ειδικά στις πιο ανεπτυγμένες χώρες [31]. Ορισμένες ωστόσο εξαιρέσεις περιλαμβάνουν την κεντρική και νότια Αμερική (όπως για παράδειγμα, Βραζιλία και Κούβα), ορισμένες χώρες της κεντρικής, νότιας και ανατολικής Ευρώπης (όπως για παράδειγμα, Βουλγαρία, Κροατία,

Ουγγαρία και Ρουμανία), καθώς και τις χώρες της Βαλτικής (όπως για παράδειγμα, Λετονία), οι οποίες αντιμετώπισαν σημαντική οικονομική μετάβαση τα τελευταία 15 χρόνια. Ωστόσο, σύμφωνα με τον ΠΟΥ, ο αριθμός των νέων περιπτώσεων ΚΟΚ καθώς και των θανάτων αναμένεται να διπλασιαστεί στο εγγύς μέλλον (Εικόνα 9) [32]. Αυτό το φαινόμενο εξηγείται από την αύξηση του προσδόκιμου ζωής με την πάροδο του χρόνου, καθώς, στην πλειοψηφία των περιπτώσεων, ο ΚΟΚ διαγιγνώσκεται σε ηλικία > 65 ετών. Το μέσο προσδόκιμο ζωής μετά τη διάγνωση στους ασθενείς με ΚΟΚ έχει αυξηθεί κατά 3% από το 1950, και το χάσμα του προσδόκιμου ζωής μεταξύ των πιο ανεπτυγμένων και των λιγότερο ανεπτυγμένων χωρών έχει μειωθεί κατά 4,9 χρόνια από το 2000.



Εικόνα 9: Αναμενόμενος αριθμός νέων διαγνώσεων και θανάτων ΚΟΚ για το έτος 2030 συγκριτικά με το 2020.

(Πηγή: <https://gco.iarc.fr/tomorrow/en/dataviz/isotype>)

Η προχωρημένη ηλικία, το αρσενικό φύλο, το κάπνισμα και η έκθεση σε χημικές ουσίες στο χώρο εργασίας αποτελούν τους κυριότερους παράγοντες κινδύνου που συμβάλλουν στην εμφάνιση του ΚΟΚ [23].

Το **κάπνισμα** είναι ο πιο κοινός παράγοντας κινδύνου για τον ΚΟΚ. Πράγματι, οι εκτιμήσεις που έχουν αναφερθεί δείχνουν ότι ο καπνός ευθύνεται για τις μισές από όλες τις περιπτώσεις. Ωστόσο, μια καθυστέρηση περίπου 20 – 30 ετών παρατηρείται μεταξύ της έκθεσης στο τσιγάρο και της διάγνωσης. Η τρέχουσα επίπτωση του ΚΟΚ είναι υψηλότερη, αν και όχι ομοιόμορφα, σε περιοχές που είχαν υψηλά ποσοστά καπνίσματος τη δεκαετία του 1980. Ειδικότερα, στην Ισπανία και στην Ιταλία, τα προτυποποιημένα κατά ηλικία ποσοστά καπνίσματος το 1980 ανέρχονταν σε ποσοστό 44,4% και 44,3%, αντίστοιχα. Το υψηλότερο ποσοστό επίπτωσης του ΚΟΚ στους άνδρες σε αυτές τις χώρες ήταν 36,7 ανά 100.000 το 2003 και 33,2 ανά 100.000 το 2007, αντίστοιχα [33, 34]. Έκτοτε, ο επιπολασμός του καπνίσματος έχει μειωθεί σημαντικά στις χώρες υψηλού εισοδήματος και, ακολούθως, η επίπτωση και η θνησιμότητα τείνουν να αντικατοπτρίζουν αυτή την τάση. Ως εκ τούτου, το κάπνισμα μπορεί να εξηγήσει εν μέρει τις γεωγραφικές διαφοροποιήσεις στην επιδημιολογία του ΚΟΚ. Δυστυχώς, ο ΠΟΥ έχει αναφέρει αυξήσεις στην κατανάλωση καπνού σε μεγάλες περιοχές του λιγότερο ανεπτυγμένου κόσμου, συμπεριλαμβανομένης της Αφρικής, της Μέσης Ανατολής, της ανατολικής Ευρώπης, των χωρών της πρώην Σοβιετικής Ένωσης και της Ασίας, όπου ο κυβερνητικός έλεγχος στην αγορά τσιγάρων (συμπεριλαμβανομένης της διαφήμισης και προώθησης καπνικών προϊόντων) είναι

μικρότερος και η ευαισθητοποίηση του κοινού για τις βλαβερές συνέπειες του καπνίσματος είναι ελλιπής.

Εκτός από το κάπνισμα, έχουν διερευνηθεί εκτενώς και οι συσχετίσεις μεταξύ διαφόρων **περιβαλλοντικών παραγόντων** και του ΚΟΚ [35]. Η **δίαιτα** με χαμηλή περιεκτικότητα σε φρούτα και λαχανικά και η αστική διαβίωση συνδέονται - αν και όχι πάντα - με αυξημένο κίνδυνο εμφάνισης ΚΟΚ. Επιπλέον, ορισμένα στοιχεία δείχνουν ότι η πρόσληψη **αλκοόλ** αυξάνει ελαφρώς τον κίνδυνο, αλλά τα επιδημιολογικά δεδομένα συγχέονται με άλλους παράγοντες κινδύνου. Το **μεταβολικό σύνδρομο** στους άνδρες αναφέρθηκε επίσης ότι επηρεάζει τον κίνδυνο ανάπτυξης ΚΟΚ, αλλά δεν έχει αποδειχθεί τεκμηριωμένα καμία άμεση ή έμμεση συσχέτιση είτε με τον κίνδυνο ανάπτυξης είτε με την πρόγνωση. Η κατανάλωση νερού ή τροφίμων που έχουν μολυνθεί με **αρσενικό** μπορεί να εξηγήσει ορισμένα τοπικά περιστατικά ΚΟΚ. Για παράδειγμα, η ρύπανση από αρσενικό συσχετίστηκε με τον κίνδυνο ανάπτυξης ΚΟΚ στην Αργεντινή, τη Χιλή και το Μπαγκλαντές [36]. Η **ατμοσφαιρική ρύπανση** προτείνεται επίσης ως ένας παράγοντας που επηρεάζει τον κίνδυνο ανάπτυξης ΚΟΚ στις λιγότερο ανεπτυγμένες χώρες. Οι χημικές ουσίες **άμεσης καύσης** (κυρίως από τα καυσαέρια των κινητήρων ντίζελ, βενζίνης και τους σταθμούς παραγωγής ηλεκτρικής ενέργειας) είναι οι κύριες πηγές ατμοσφαιρικής ρύπανσης.

Η επαγγελματική έκθεση σε **καρκινογόνους παράγοντες** έχει από καιρό συσχετιστεί με τον κίνδυνο ανάπτυξης ΚΟΚ. Σύμφωνα με αναλύσεις, ο μεγαλύτερος κίνδυνος ανάπτυξης ΚΟΚ εμφανίζεται σε βιομηχανικές περιοχές στις οποίες γίνεται επεξεργασία χρωμάτων, καουτσούκ, προϊόντων πετρελαίου και βαφών, ενώ ο μεγαλύτερος κίνδυνος για θνησιμότητα από ΚΟΚ εμφανίζεται στους εργαζόμενους σε ηλεκτρικές και χημικές διεργασίες [37]. Μαζί με την εκτεταμένη αστικοποίηση, πολλές μεταποιητικές διαδικασίες μεταφέρονται από τις πιο ανεπτυγμένες στις λιγότερο ανεπτυγμένες χώρες, γεγονός που δυνητικά δημιουργεί αυξημένους επαγγελματικούς κινδύνους για τους εργαζόμενους στις τοπικές κοινωνίες. Παρόλα αυτά, δεν θεωρείται ότι πάνω από το 8% των περιπτώσεων ΚΟΚ αποδίδονται σε τέτοιες εκθέσεις και ο παγκόσμιος αντίκτυπος των επαγγελματικών κινδύνων στην επίπτωση του ΚΟΚ και στην ερμηνεία της γεωγραφικής του ποικιλομορφίας μένει να εξακριβωθεί πλήρως. Επιπλέον, η ανεργία, ο αριθμός των σωματικά ανθυγιεινών ημερών εργασίας, ο αριθμός των ημερών έκθεσης σε ατμοσφαιρική ρύπανση του όζοντος, το ποσοστό των σπιτιών με νερό που προέρχεται από πηγάδια, η απασχόληση στη βιομηχανία εξόρυξης, η αστική διαβίωση και η εθνικότητα συνδέονται με την θνητότητα από ΚΟΚ.

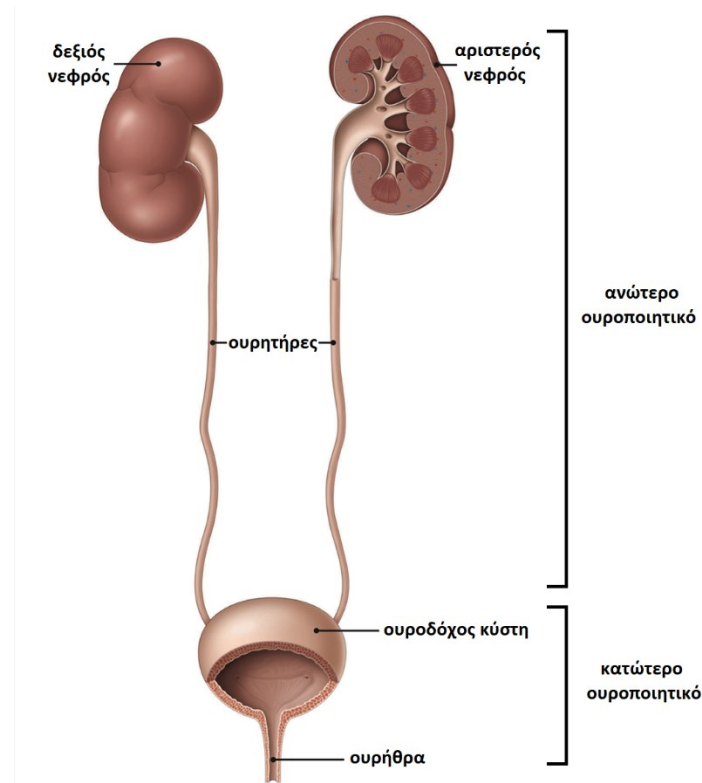
Ορισμένα στοιχεία υποστηρίζουν και μια **γενετική προδιάθεση** για την ανάπτυξη ΚΟΚ. Γονίδια που εμπλέκονται στο μεταβολισμό των καρκινογόνων ουσιών, όπως η N-ακετυλοτρανσφεράση (N-acetyltransferase – NAT) και ο πολυμορφισμός null genotype για το γονίδιο *GSTM1*, συνδέονται με αυξημένο κίνδυνο. Συγκεκριμένα, οι γυναίκες με ιστορικό καπνίσματος και οι οποίες φέρουν τον γονότυπο *GSTM1*-null είναι πιο επιρρεπείς στην ανάπτυξη ΚΟΚ από ό,τι οι αντίστοιχες μη καπνίστριες [38]. Μεγάλες μελέτες γονιδιωματικής συσχέτισης σε ολόκληρο το γονιδίωμα (Genome-wide association studies – GWAS) έχουν βρει παραλλαγές αλληλουχίας που μπορούν να αυξήσουν τον κίνδυνο για ΚΟΚ. Για παράδειγμα, μεταβολές στον μεταφορέα ουρίας, που κωδικοποιείται από το γονίδιο *SLC14A1*, σχετίζονται με τη συγκέντρωση των νεφρικών ούρων και μπορούν να επηρεάσουν την επαφή των καρκινογόνων ουσιών με τις ουροθηλιακές επιφάνειες [39]. Επιπλέον, βάσει μιας νέας, βιοπληροφορικής ανάλυσης που μέτρησε τις αλληλεπιδράσεις μεταξύ γονιδίων σε δύο GWAS, η ευαισθησία στον ΚΟΚ φέρεται να σχετίζεται με τα πρωτεϊνικά σύμπλοκα αποκαρβοξυλάσης (decarboxylase), τα οποία αποτελούν και δυνητικούς φαρμακευτικούς στόχους [40].

Ανατομία και Φυσιολογία της Ουροδόχου Κύστης

Το σώμα λαμβάνει θρεπτικά συστατικά από τις τροφές και τα μετατρέπει σε ενέργεια. Αφού ο οργανισμός λάβει τα συστατικά της τροφής που χρειάζεται, τα απόβλητα μένουν πίσω στο έντερο και στο αίμα. Η λειτουργία του **ουροποιητικού συστήματος** είναι να φιλτράρει το αίμα και να δημιουργεί ούρα ως άχρηστο υποπροϊόν. Τα όργανα του ουροποιητικού συστήματος περιλαμβάνουν τους νεφρούς, τη νεφρική πύελο, τους ουρητήρες, την ουροδόχο κύστη και την ουρήθρα. Η ανατομία του ουροποιητικού συστήματος παρουσιάζεται στην Εικόνα 10.

Τα νεφρά και το ουροποιητικό σύστημα βοηθούν το σώμα να αποβάλλει τα υγρά απόβλητα, που ονομάζονται ουρία, και να διατηρεί σε ισορροπία τις χημικές ουσίες, όπως το κάλιο, το νάτριο, και το νερό. Η ουρία παράγεται όταν οι τροφές που περιέχουν πρωτεΐνες, όπως το κρέας, τα πουλερικά και ορισμένα λαχανικά, διασπώνται στο σώμα. Η ουρία μεταφέρεται με την κυκλοφορία του αίματος στα νεφρά, όπου απομακρύνεται μαζί με το νερό και άλλα απόβλητα με τη μορφή ούρων.

Άλλες σημαντικές λειτουργίες των νεφρών περιλαμβάνουν τη ρύθμιση της αρτηριακής πίεσης και την παραγωγή ερυθροποιητίνης, η οποία ελέγχει την παραγωγή ερυθρών αιμοσφαιρίων στο μυελό των οστών. Οι νεφροί ρυθμίζουν επίσης την οξεοβασική ισορροπία και διατηρούν τα υγρά στο σώμα.



Εικόνα 10: Το ουροποιητικό σύστημα του ανθρώπου.

(Πηγή: <https://kidneystonemelbourne.com.au/kidney-stones/anatomy-urinary-system>)

Αναλυτικότερα, οι **νεφροί** στον ανθρώπινο οργανισμό λειτουργούν με σκοπό να απομακρύνουν τα απόβλητα και τις φαρμακευτικές ουσίες από το σώμα, να εξισορροπούν τα υγρά του σώματος, να απελευθερώνουν ορμόνες για τη ρύθμιση της αρτηριακής πίεσης και να ελέγχουν την παραγωγή ερυθρών αιμοσφαιρίων. Οι νεφροί απομακρύνουν την ουρία από το αίμα μέσω μικροσκοπικών μονάδων φιλτραρίσματος που ονομάζονται νεφρώνες. Κάθε νεφρώνας αποτελείται από μια σφαίρα που σχηματίζεται από μικρά τριχοειδή αγγεία του αίματος, που ονομάζονται αγγειώδες σπείραμα, και έναν μικρό σωλήνα που ονομάζεται νεφρικό σωληνάριο. Η ουρία, μαζί με το νερό και άλλες άχρηστες ουσίες,

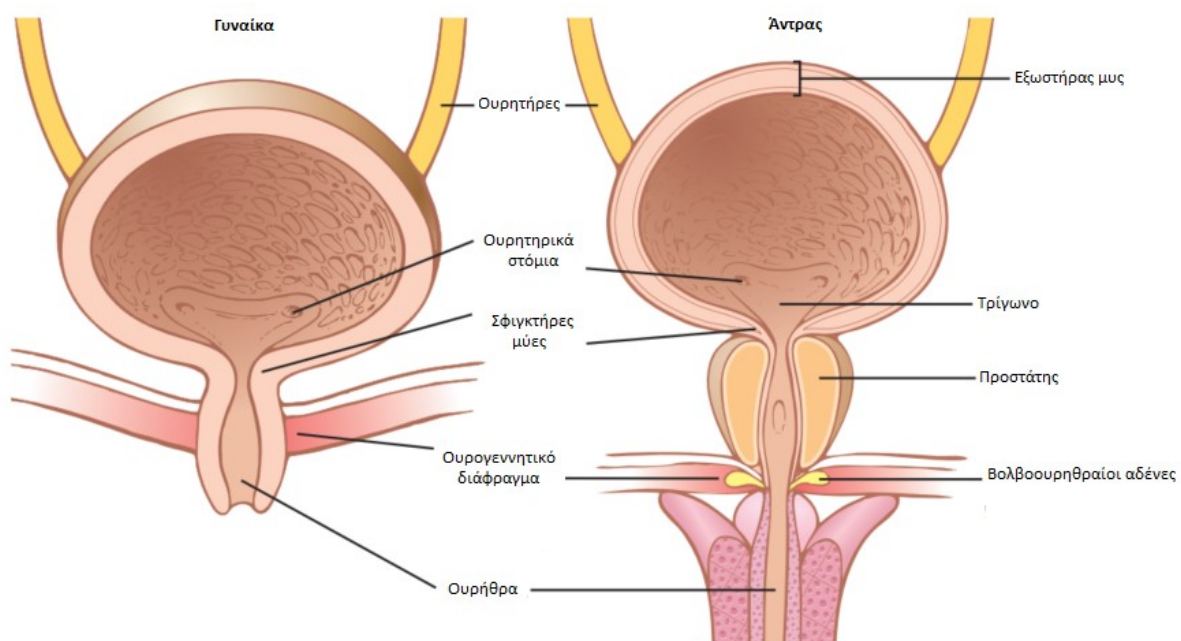
σχηματίζει τα ούρα καθώς περνούν μέσα από τους νεφρώνες και κάτω από τα νεφρικά σωληνάκια του νεφρού.

Οι στενοί σωλήνες που μεταφέρουν τα ούρα από τους νεφρούς στην ουροδόχο κύστη ονομάζονται **ουρητήρες**. Οι μύες στα τοιχώματα των ουρητήρων συσπώνται και χαλαρώνουν συνεχώς, ωθώντας τα ούρα προς τα κάτω, μακριά από τους νεφρούς. Εάν τα ούρα συσσωρευτούν ή αφεθούν να λιμνάσουν, μπορεί να αναπτυχθεί λοίμωξη των νεφρών. Περίπου κάθε 10 έως 15 δευτερόλεπτα, μικρές ποσότητες ούρων αδειάζουν στην ουροδόχο κύστη από τους ουρητήρες.

Η **ουροδόχος κύστη** αποτελεί το τριγωνικό, κοίλο όργανο που βρίσκεται στην κάτω κοιλιακή χώρα. Συγκρατείται στη θέση του από συνδέσμους που ενώνονται με άλλα όργανα και τα οστά της πυέλου. Τα τοιχώματα της ουροδόχου κύστης χαλαρώνουν και διαστέλλονται για την αποθήκευση των ούρων αλλά και συστέλλονται και ισιώνουν για την εκκένωση των ούρων μέσω της ουρήθρας. Η τυπική υγιής ενήλικη κύστη μπορεί να αποθηκεύσει έως και δύο φλιτζάνια ούρων για δύο με πέντε ώρες.

Οι θέσεις τυχόν ανωμαλιών μέσα στην κύστη περιγράφονται με συγκεκριμένους όρους ανάλογα με το που εντοπίζεται η ανωμαλία. Αυτοί οι όροι είναι το κυστικό τρίγωνο (trigone), το οποίο αποτελεί μια περιοχή σε σχήμα τριγώνου κοντά στη συμβολή της ουρήθρας με την ουροδόχο κύστη, το δεξί και αριστερό πλευρικό τοίχωμα, τα οποία αποτελούν τα τοιχώματα εκατέρωθεν της τριγωνικής περιοχής, το οπίσθιο τοίχωμα και ο θόλος της κύστης, ο οποίος αποτελεί και την «οροφή» της ουροδόχου κύστης (Εικόνα 11).

Επίσης, μέρος του ουροποιητικού συστήματος αποτελούν οι δύο **σφιγκτήρες μύες**, οι οποίοι βοηθούν ώστε να μην διαρρέουν τα ούρα από την κύστη, κλείνοντας σφιχτά σαν λάστιχο γύρω από το άνοιγμα της ουροδόχου κύστης.



Εικόνα 11: Ανατομία της ουροδόχου κύστης και της ουρήθρας σε άνδρες και γυναίκες. [41]

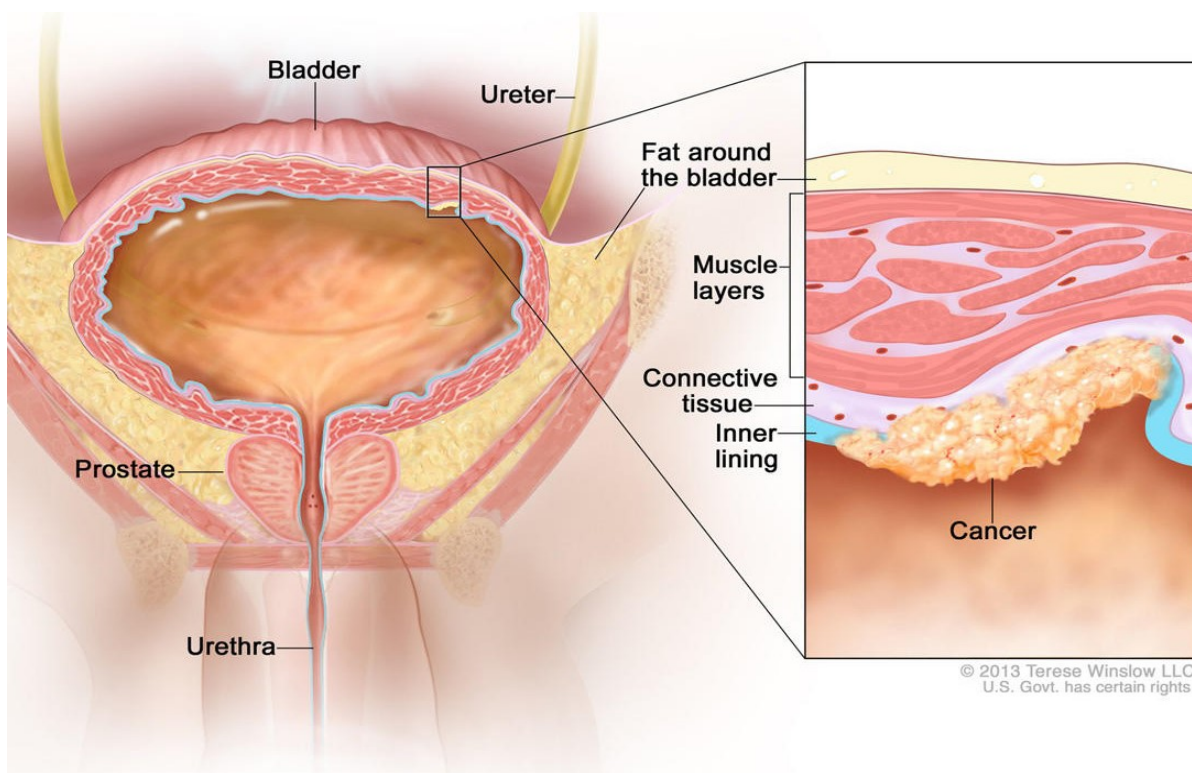
Τα **νεύρα της ουροδόχου κύστης** ειδοποιούν το άτομο όταν είναι ώρα να ουρήσει ή να αδειάσει την ουροδόχο κύστη.

Τέλος, η **ουρήθρα** αποτελεί τον σωλήνα ο οποίος επιτρέπει στα ούρα να περάσουν έξω από το σώμα. Ο εγκέφαλος δίνει σήμα στους μύες της ουροδόχου κύστης να συσπαστούν, γεγονός που εξωθεί τα ούρα από την ουροδόχο κύστη. Ταυτόχρονα, ο εγκέφαλος δίνει σήμα στους σφιγκτηριακούς μύες να

χαλαρώσουν για να αφήσουν τα ούρα να εξέλθουν από την ουροδόχο κύστη μέσω της ουρήθρας. Όταν όλα τα σήματα συμβαίνουν με τη σωστή σειρά, επέρχεται φυσιολογική ούρηση.

Μηχανισμοί και Παθοφυσιολογία του Καρκίνου της Ουροδόχου Κύστης

Ο ΚΟΚ συνήθως ξεκινά από τα κύτταρα του επιθηλίου (ουροθήλιο), το οποίο επενδύει τον εσωτερικό αυλό της ουροδόχου κύστης (Εικόνα 12). Ο συγκεκριμένος τύπος καρκίνου, γνωστός ως **καρκίνωμα ουροδόχου κύστης μεταβατικού επιθηλίου** ή, πιο απλά, **ουροθηλιακό καρκίνωμα** (urothelial carcinoma), αντιπροσωπεύει τον πιο συχνά απαντώμενο τύπο ΚΟΚ. Άλλοι τύποι ΚΟΚ που έχουν επίσης περιγραφεί και παρουσιάζουν ποικίλη ιστολογική εικόνα αποτελούν το πλακώδες καρκίνωμα, το μικροκυτταρικό καρκίνωμα και το αδενοκαρκίνωμα. Οι τύποι αυτοί καρκίνου αποτελούν περίπου το 10 – 15% των περιπτώσεων ΚΟΚ [42].



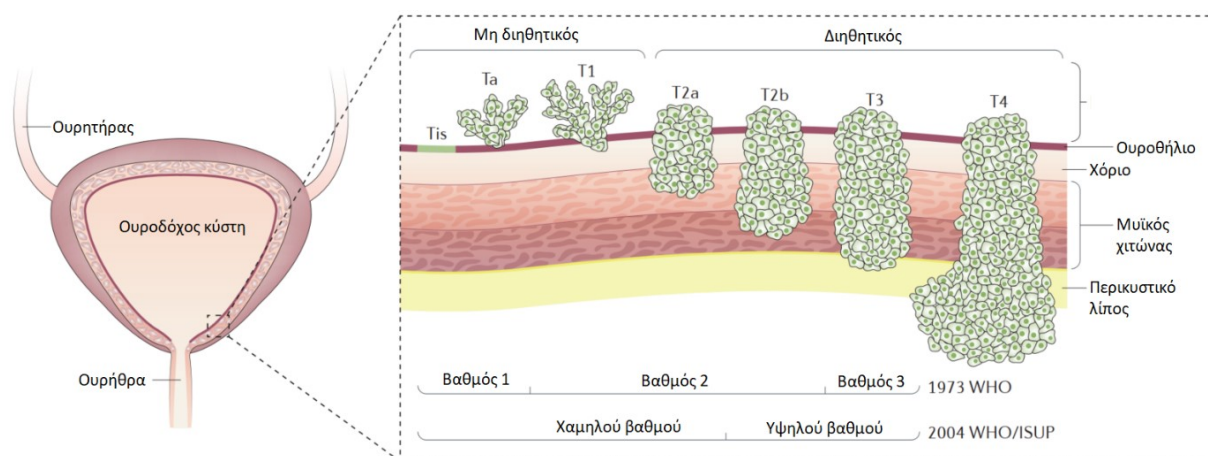
Εικόνα 12: Ο καρκίνος ουροδόχου κύστης (ΚΟΚ).

(Πηγή: <https://www.cancer.gov/types/bladder/patient/bladder-treatment-pdq>)

Τα υψηλού βαθμού (high grade) ουροθηλιακά καρκινώματα μπορεί να είναι μικροσκοπικά θηλώδη, σαρκωματοειδή, πλασματοκυτταρικά, φωλεακά και μικροκυστικά καρκινώματα, και έχουν την τάση για αποκλίνουσα διαφοροποίηση σε π.χ. πλακώδεις ή αδενικές ιστολογικές μορφές. Ο ΚΟΚ με ύπαρξη ποικίλων ιστολογικών μορφών έχει συσχετιστεί με τοπικά επιθετική νόσο, αυξημένη πιθανότητα μετάστασης και μικρή ανταπόκριση στα υπάρχοντα θεραπευτικά σχήματα. Ωστόσο, υπάρχει μια διχογνωμία στην επιστημονική κοινότητα σχετικά με την πραγματική επιρροή που έχει η ιστολογία του καρκίνου στην πρόγνωση και στα θεραπευτικά αποτελέσματα. Γενικά, θα πρέπει να εφαρμόζεται εξατομικευμένη διαχείριση σε κάθε ασθενή, υπό την παρούσα συνθήκη της περιορισμένης γνώσης και βιβλιογραφίας.

Ο ΚΟΚ χωρίζεται σε δύο βασικές κατηγορίες ανάλογα με το βαθμό διήθησης της νόσου: στον **μη μυοδιηθητικό** ή επιφανειακό ΚΟΚ – ΜΜΚΟΚ (non-muscle invasive BCa – NMIBC), ο οποίος αντιστοιχεί περίπου στο 70 με 75% των διαγνώσεων, και στον **μυοδιηθητικό ΚΟΚ** – ΜΚΟΚ (muscle

invasive BCa – MIBC), ο οποίος αναλογεί στο υπόλοιπο 25 με 30%. Οι δύο αυτές κατηγορίες διαφέρουν σημαντικά ως προς το γενετικό τους υπόβαθρο και εμφανίζουν διαφορετική πρόγνωση [43]. Έτσι, οι όγκοι οι οποίοι διηθούν τον εξωστήρα μυ έχουν περισσότερες πιθανότητες να οδηγήσουν σε μεταστάσεις στους λεμφαδένες και σε άλλα όργανα, εν αντιθέσει με τους καρκίνους οι οποίοι δεν έχουν φτάσει στο μυϊκό χιτώνα της ουροδόχου κύστης (Εικόνα 13). Το στάδιο του καρκίνου κατά τη στιγμή της διάγνωσης δεν έχει αλλάξει τα τελευταία τουλάχιστον 10 χρόνια, σύμφωνα με το μητρώο SEER [44], καθώς δεν υπάρχουν καθιερωμένα προγράμματα προληπτικού ελέγχου. Αντίστοιχα, και τα ποσοστά θνησιμότητας δεν έχουν αλλάξει, καθώς δεν υπάρχει θεραπεία για τη μεταστατική νόσο.



Εικόνα 13: Τύποι και στάδια του ΚΟΚ. [42]

Οι παθολογικές και κλινικές πληροφορίες από μοντέλα ποντικών και ανθρώπινα δείγματα επιβεβαιώνουν ότι το ουροθηλιακό καρκίνωμα αναπτύσσεται μέσω δύο διαφορετικών οδών: δημιουργώντας θηλώδη μη μυοδιηθητικά και μη θηλώδη (συμπαγή) μυοδιηθητικά καρκινώματα. Σε μοντέλα ποντικών, η χαμηλού επιπέδου έκφραση του μεταλλαγμένου γονιδίου *HRAS* οδηγεί σε επίπεδες ή θηλώδεις υπερπλαστικές αλλοιώσεις του ουροθηλίου, ενώ η υψηλού επιπέδου έκφρασή του οδηγεί σε μη μυοδιηθητικά καρκινώματα. Ομοίως, στον ανθρώπινο οργανισμό, οι συνήθεις πρόδρομοι του μη μυοδιηθητικού ΚΟΚ είναι οι επίπεδες ή θηλώδεις υπερπλαστικές αλλοιώσεις του ουροθηλίου. Δύο κοινές παραλλαγές του μη μυοδιηθητικού ΚΟΚ, η διαγραφή του χρωμοσώματος 9 και μια σημειακή μετάλλαξη στο γονίδιο *FGFR3* (που κωδικοποιεί τον υποδοχέα 3 του αυξητικού παράγοντα των ινοβλαστών), είναι επίσης εμφανείς σε αυτές τις πρόδρομες υπερπλαστικές αλλοιώσεις. Επίσης, κατά τη μελέτη των μη μυοδιηθητικών θηλωμάτων ΚΟΚ από τον ίδιο ασθενή, οι μεταβολές αυτές ήταν υπαρκτές. Το γεγονός αυτό υποδηλώνει μία κλωνική σχέση μεταξύ των πρόδρομων υπερπλασιών και των όγκων. Το ουροθηλιακό θήλωμα, το οποίο θεωρείται ένας καλοήθης όγκος, παρουσιάζει επίσης συχνές μεταλλάξεις του γονιδίου *FGFR3*.

Αντίθετα, η δημιουργία μυοδιηθητικού (συμπαγούς) ΚΟΚ σε μοντέλα ποντικών προϋποθέτει την απενεργοποίηση ενός ή περισσότερων από τα ογκοκατασταλτικά γονίδια. Τα γονίδια αυτά περιλαμβάνουν τα *TP53*, *RBI* και *PTEN*. Οι όγκοι σε αυτά τα μοντέλα ποντικών έπονται της ανάπτυξης μίας επίπεδης ουροθηλιακής δυσπλασίας και ενός καρκινώματος *in situ* (CIS). Ομοίως, στους ανθρώπους, ο υψηλός κίνδυνος ανάπτυξης μυοδιηθητικού ΚΟΚ σε ασθενείς με επίπεδη δυσπλασία ή CIS είναι καλά τεκμηριωμένος. Ακόμη, οι πρόδρομες αυτές βλάβες έχουν κοινά χαρακτηριστικά με τον υψηλού βαθμού και μυοδιηθητικό ΚΟΚ. Αυτά τα χαρακτηριστικά διευκολύνουν τον κυτταρικό πολλαπλασιασμό και περιλαμβάνουν συνήθως μεταλλάξεις στο γονίδιο *TP53* και σταθερή έκφραση του γονιδίου αυτού. Επιπλέον, η υπερ-έκφραση των γονιδίων *CK20* και *HER2* και η μειωμένη έκφραση του

γονιδίου *PTEN* με ταυτόχρονη υπερ-ρύθμιση του μονοπατιού της φωσφατιδυλνινοσιτόλης-3 κινάσης (*PI3K*) αποτελούν κοινά χαρακτηριστικά των μυοδιηθητικών όγκων.

Ο Γονιδιακός Άτλαντας του Καρκίνου (Cancer Genome Atlas – TCGA) έχει συμβάλει στην καλύτερη κατανόηση των παθοφυσιολογικών μηχανισμών του ΚΟΚ. Η μελέτη αυτή του γενετικού χαρακτηρισμού του μυοδιηθητικού ΚΟΚ παρείχε πληροφορίες, όχι μόνο για τους γενετικούς παράγοντες που θα μπορούσαν να χρησιμεύσουν ως θεραπευτικοί στόχοι, αλλά και για τους υπότυπους ή τις κατηγορίες της διηθητικής αυτής ασθένειας. Οι κατηγορίες αυτές μπορεί να σχετίζονται με προγνωστικούς παράγοντες και μοναδικές θεραπευτικές προσεγγίσεις και τελικά να συμβάλλουν καταλυτικά στην εξατομικευμένη προσέγγιση των ασθενών.

Διάγνωση, προληπτικός έλεγχος και πρόληψη

Το συνηθέστερο σύμπτωμα του ΚΟΚ είναι η μικροσκοπική ή μακροσκοπική (ορατή) αιματουρία, η οποία καταγράφεται σε ποσοστό 13,7% και 78,3% των ασθενών, αντίστοιχα [45]. Εναλλακτικά, η επίπτωση του ΚΟΚ είναι 10 - 20% σε ασθενείς με μακροσκοπική αιματουρία και 2 - 5% σε αναφερόμενους πληθυσμούς με μικροσκοπική αιματουρία. Συνήθως, η μικροσκοπική ανώδυνη αιματουρία αποτελεί ένα τυχαίο εύρημα σε γενικές εξετάσεις ούρων που διενεργούνται για την αξιολόγηση της υγείας ενός ατόμου. Επίσης, πολλοί ασθενείς με μικροσκοπική αιματουρία δεν αξιολογούνται επαρκώς, καθώς δεν υπάρχει κάποιος καθιερωμένος προσυμπτωματικός έλεγχος (screening test) για τον ΚΟΚ. Ωστόσο, στους ασθενείς με ΚΟΚ που παρατηρείται μακροσκοπική αιματουρία το στάδιο της νόσου είναι συνήθως προχωρημένο.

Ο θεράπων ιατρός μπορεί επίσης να υποπτευθεί την παρουσία του ΚΟΚ εάν ο ασθενής παρουσιάζει μη ειδικά συμπτώματα του κατώτερου ουροποιητικού συστήματος τα οποία σχετίζονται με διαταραχή της αποθήκευσης των ούρων, και πιο συγκεκριμένα, την αυξημένη ανάγκη για ούρηση, την αυξημένη συχνότητα ούρησης καθώς και τη δυσουρία. Τα συμπτώματα αυτά είναι συχνότερα σε ασθενείς με CIS παρά με θηλώδεις όγκους (παθολογικού σταδίου T_a) και θα πρέπει να προτρέπουν τους ασθενείς σε ουρολογική εκτίμηση από ειδικό. Η ανάλυση ούρων, με βυθομετρική ράβδο ή/και με μικροσκόπιο, θα πρέπει να υποδεικνύει την παρουσία αιματουρίας και να είναι σε θέση να αποκλείσει τη λοίμωξη του ουροποιητικού συστήματος, η οποία μπορεί να μιμείται ή/και να συνυπάρχει με τον ΚΟΚ. Κατά συνέπεια, μετά τη θεραπεία της ουρολοίμωξης, η απουσία αυτών των συμπτωμάτων και της αιματουρίας πρέπει να επιβεβαιώνεται.

Η αξιολόγηση των ασθενών για τους οποίους υπάρχουν υποψίες ότι πάσχουν από ΚΟΚ πραγματοποιείται με τη διενέργεια κυστεοσκόπησης, η οποία είναι μια ενδοσκοπική διαδικασία που εκτελείται στα εξωτερικά ιατρεία με εύκαμπτο ενδοσκόπιο υπό τοπική αναισθησία. Οποιοδήποτε παθολογικό εύρημα, όπως ερυθρές, επίπεδες, θηλώδεις ή συμπαγείς αλλοιώσεις, απαιτεί ιστολογική αξιολόγηση, διότι κάποιες καλοήθεις καταστάσεις, όπως οι φλεγμονώδεις παθήσεις, μπορεί να μιμούνται τον ΚΟΚ. Η λήψη ιστοτεμαχίου για ιστολογική εξέταση μπορεί να ληφθεί με διουρηθρική βιοψία ή με εκτομή ολόκληρης της περιοχής. Συχνά, η εξέταση των κυττάρων στα ούρα (κυτταρολογική εξέταση) πραγματοποιείται ως συμπληρωματικό μέτρο για την ανίχνευση του καρκίνου που έχει διαφύγει. Τα κύτταρα με κακοήγη εμφάνιση είναι άκρως ενδεικτικά της παρουσίας καρκινικών αλλοιώσεων στην ουροδόχο κύστη και ο εντοπισμός τους υποδεικνύει την ανάγκη διενέργειας νέας κυστεοσκόπησης και ιστολογικής διερεύνησης. Επί του παρόντος, **κανένας καρκινικός βιοδείκτης στα ούρα δεν έχει επιδείξει επαρκή ευαισθησία και ειδικότητα** ώστε να αντικαταστήσει την κυστεοσκόπηση στην ανίχνευση του ΚΟΚ.

Είναι σημαντικό ότι το CIS αποτελεί ένα είδος καρκίνου που είναι δύσκολο να διαγνωστεί κυστεοσκοπικά, επειδή οι βλάβες αυτές είναι δύσκολο να διακριθούν από τον φυσιολογικό ιστό της ουροδόχου κύστης. Αντ' αυτού, απαιτείται μικροσκοπική ανάλυση ούρων για τον εντοπισμό άτυπων

κυττάρων και η διάγνωση μπορεί να επιβεβαιώνεται μετά από ιστολογική αξιολόγηση δειγμάτων ιστού της ουροδόχου κύστης. Η κυστεοσκοπική ανίχνευση του CIS μπορεί να ενισχυθεί με κυστεοσκόπηση φθορισμού ή απεικόνιση στενής ζώνης (narrow-band imaging). Αυτές οι τεχνολογίες βελτιώνουν τη διαφοροποίηση των καρκινικών αλλοιώσεων από τον φυσιολογικό ιστό εκμεταλλεύόμενες την αυξημένη μεταβολική δραστηριότητα (μπλε φως) και την αρχιτεκτονική των αγγείων (στενή ζώνη) που εμφανίζεται στα καρκινικά κύτταρα και έχουν υψηλότερη ειδικότητα στον εντοπισμό του ΚΟΚ συγκριτικά με την παραδοσιακή κυστεοσκόπηση.

Η απεικόνιση του ανώτερου ουροποιητικού συστήματος, δηλαδή του νεφρικού συστήματος και των ουρητήρων, αποτελεί επίσης σημαντικό στοιχείο της αξιολόγησης ασθενών με αιματουρία, αλλά παίζει μικρό ρόλο στη διάγνωση του ΚΟΚ. Παρόλο που η αξονική και η μαγνητική ουρογραφία έχουν προταθεί για τη βελτίωση της διαγνωστικής ακρίβειας, η απεικόνιση με υπερήχους έχει φανεί ότι αρκεί. Σε ασθενείς με επιβεβαιωμένο μυοδιηθητικό ΚΟΚ, η αξονική τομογραφία είναι επιβεβλημένη για τη σταδιοποίηση της νόσου και την εκτίμηση της πιθανής απομακρυσμένης εξάπλωσης.

Η πρόγνωση (και η αντιμετώπιση) του ΚΟΚ εξαρτάται από τα ευρήματα της **ιστοπαθολογικής εξέτασης** (μη μυοδιηθητικός ή μυοδιηθητικός ΚΟΚ). Πράγματι, η ιστοπαθολογική εξέταση είναι ο μόνος αξιόπιστος καθοριστικός παράγοντας της μοριακής βιολογίας του όγκου, βάσει του οποίου θα γίνει η σταδιοποίηση (staging) της νόσου και θα ληφθεί η απόφαση για τη σωστή διαχείριση της νόσου. Ωστόσο, αν και η πρόγνωση δεν μπορεί να είναι ακριβής, το βάθος διήθησης του όγκου στο τοίχωμα της ουροδόχου κύστης μπορεί να παρέχει μια καλή διαστρωμάτωση του κινδύνου και της πρόγνωσης του ασθενούς. Για παράδειγμα, οι όγκοι που περιορίζονται στο εσωτερικό τοίχωμα της ουροδόχου κύστης (βλεννογόνο) και δεν διεισδύουν στη χοριακή στιβάδα (lamina propria) ταξινομούνται ως στάδιο Ta σύμφωνα με το σύστημα ταξινόμησης όγκου – λεμφαδένων – μετάστασης (Tumor Node Metastasis – TNM). Οι όγκοι που διεισδύουν στη χοριακή στιβάδα ταξινομούνται ως στάδιο T1. Οι όγκοι αυτοί χαρακτηρίζονται από δυσμενή βιολογία του όγκου παρόμοια με τους μυοδιηθητικούς όγκους. Οι όγκοι που διεισδύουν στον εξωστήρα μυ της ουροδόχου κύστης και πέραν αυτού είναι ιδιαίτερα επιθετικοί.

Επιπλέον, ο **βαθμός κακοήθειας** (δηλαδή ο βαθμός διαφοροποίησης των κυττάρων) είναι σημαντικός για την αξιολόγηση των όγκων σταδίου Ta, διότι οι καλά διαφοροποιημένοι (χαμηλού βαθμού κακοήθειας) όγκοι είναι λιγότερο επιθετικοί συγκριτικά με τους υψηλού βαθμού κακοήθειας. Το 2016 εκδόθηκε η τρίτη έκδοση της ταξινόμησης του ΠΟΥ για τα στάδια του ΚΟΚ, τονίζοντας τις αλλαγές που προτάθηκαν από το 2004 στην ταξινόμηση που ίσχυε από το 1973 [46]. Παρόλο που το όφελος του πιο πρόσφατου συστήματος ταξινόμησης δεν έχει ακόμη αξιολογηθεί συστηματικά, συγκριτικά με το σύστημα ταξινόμησης του 1973, η τρέχουσα έκδοση (2016) παρέχει ομοιόμορφες ορολογίες και σαφέστερους ορισμούς, κυρίως στην, ολόένα και περισσότερο αναγνωρισμένη, αποκλίνουσα διαφοροποίηση των κυττάρων στον μυοδιηθητικό ΚΟΚ.

Συνολικά, η σημασία της ιστοπαθολογικής εξέτασης των δειγμάτων της ουροδόχου κύστης στη διάγνωση, την πρόγνωση και τη θεραπεία του ΚΟΚ είναι αδιαμφισβήτητη. Ωστόσο, η μεταβλητότητα μεταξύ του ίδιου και των διαφορετικών παρατηρητών όσον αφορά τη σταδιοποίηση και την ταξινόμηση αποτελεί περιορισμό της τεχνικής. Για παράδειγμα, έχει αναφερθεί σημαντική μεταβλητότητα μεταξύ των παρατηρητών στην ταξινόμηση των όγκων σταδίου T1 έναντι των όγκων σταδίου Ta και του βαθμού της κακοήθειας με τη χρήση των κριτηρίων του 1997 και του 2004 του ΠΟΥ [47].

Ο προσυμπτωματικός έλεγχος (screening) για την έγκαιρη ανίχνευση του ΚΟΚ δεν έχει καθιερωθεί σε καμία χώρα λόγω των χαμηλών ποσοστών εμφάνισής του σε σύγκριση με άλλους πιο κοινούς καρκίνους, όπως ο καρκίνος του προστάτη και ο καρκίνος του μαστού. Ωστόσο, αρκετές μη τυχαιοποιημένες μελέτες έχουν καταδείξει τη δυνατότητα έγκαιρης ανίχνευσης του ΚΟΚ με τη διερεύνηση ανίχνευσης αίματος ή καρκινικών βιοδεικτών στα ούρα, όπως η κυτταρολογία ή το τεστ

πρωτεΐνης πυρηνικού σκελετού (NMP-22 test), το οποίο αντανακλά τη μιτωτική δραστηριότητα [48]. Παρόλο που η αξία, για παράδειγμα, της κυτταρολογίας των ούρων, είναι σημαντική στην παρακολούθηση των ασθενών με μη μυοδιηθητικό ΚΟΚ, τα χαμηλά ποσοστά ανίχνευσης της μεθόδου αυτής περιορίζουν την τρέχουσα εφαρμογή της στο πλαίσιο του προσυμπτωματικού ελέγχου. Επιπλέον, έχει προταθεί η διενέργεια προσυμπτωματικού ελέγχου σε άτομα υψηλού κινδύνου ανάπτυξης ΚΟΚ. Άτομα αυτού του πληθυσμού αποτελούν τα άτομα με σημαντική έκθεση σε παράγοντες κινδύνου, όπως οι καπνιστές. Σε αντίθεση με άλλες κοινές κακοήθειες, η γενετική προδιάθεση έχει μικρή χρησιμότητα στην πρόβλεψη του ποιος θα αναπτύξει ΚΟΚ, αλλά η γενετική σαφώς ασκεί επιρροή μέσω της ρύθμισης της ευαισθησίας σε περιβαλλοντικούς παράγοντες κινδύνου.

Έτσι, η πρόληψη του ΚΟΚ επικεντρώνεται στην αποφυγή της έκθεσης στον καπνό καθώς και σε περιβαλλοντικούς και επαγγελματικούς καρκινογόνους παράγοντες. Είναι προφανές πως η διακοπή του καπνίσματος αποτελεί πρόκληση και πολλοί ασθενείς δυσκολεύονται να διακόψουν το κάπνισμα. Ωστόσο, οι γιατροί έχουν την υποχρέωση να ενημερώνουν τους ασθενείς σχετικά με κινδύνους και να ενθαρρύνουν τη διακοπή του καπνίσματος.

Συμπερασματικά, δεν υπάρχει επί του παρόντος κάποια διαγνωστική εξέταση ρουτίνας η οποία να συστήνεται για το γενικό πληθυσμό ή τους ανθρώπους που εμφανίζουν μέτριο ρίσκο εμφάνισης της νόσου [49]. Τα βασικά διαγνωστικά εργαλεία για τους συμπτωματικούς ασθενείς ή τους ανθρώπους που βρίσκονται σε υψηλό ρίσκο αποτελούν η κυστεοσκόπηση - η οποία συνιστά το «χρυσό πρότυπο» (gold standard) για την αξιολόγηση της κατώτερης ουροποιητικής οδού - καθώς και τα τεστ ανάλυσης ούρων. Σε αυτά συμπεριλαμβάνονται η κυτταρολογική και η γενική εξέταση ούρων.

Διαχείριση – Αντιμετώπιση – Θεραπεία

Διουρηθρική εκτομή του όγκου της ουροδόχου κύστης

Αρχικά, όλοι οι νεοδιαγνωσθέντες όγκοι της ουροδόχου κύστης απαιτούν μια ενδοσκοπική εκτομή υπό γενική ή επισκληρίδιο αναισθησία. Αυτή η διαδικασία, η οποία ονομάζεται **διουρηθρική εκτομή του όγκου της ουροδόχου κύστης (TransUrethral Resection of Bladder Tumour – TURBT)**, εκτελείται για να καταστεί δυνατή μια ενδελεχής επισκόπηση της ουροδόχου κύστης και για την πλήρη εκτομή του όγκου συνοδευόμενου από ένα μέρος του μυός της κύστης, με σκοπό να επιτευχθεί μιας ακριβής σταδιοποίηση του όγκου. Διατίθενται διάφορες κατευθυντήριες οδηγίες για τη διαχείριση του ΚΟΚ, μεταξύ άλλων από την Ευρωπαϊκή Ένωση Ουρολογίας (European Association of Urology), την Αμερικανική Ουρολογική Εταιρεία (American Urological Association), την Εταιρεία Ουρολογικής Ογκολογίας (Society of Urologic Oncology) και το Εθνικό Δίκτυο Ογκολογίας των ΗΠΑ (National Comprehensive Cancer Network – NCCN). Αν και σε γενικές γραμμές οι κατευθυντήριες οδηγίες ταυτίζονται, υπάρχουν κάποιες σημαντικές διαφοροποιήσεις λόγω των διαφορετικών επιπέδων υποστήριξης τεκμηρίωσης. Σε γενικές γραμμές, οι μη μυοδιηθητικοί όγκοι αντιμετωπίζονται με ενδοσκοπική εκτομή και ενδοκυτικές θεραπείες εγχύσεων, ενώ οι μυοδιηθητικοί όγκοι αντιμετωπίζονται με πιο επιθετικές θεραπείες, όπως η κυστεκτομή (ολική αφαίρεση της ουροδόχου κύστης) με ή χωρίς χημειοθεραπεία.

Η TURBT έχει όχι μόνο διαγνωστικό αλλά και θεραπευτικό ρόλο και μπορεί να αποτελέσει επαρκή και δυνητικά ιάσιμη θεραπεία ανάλογα με τα παθολογικά χαρακτηριστικά του όγκου. Η TURBT πραγματοποιείται με τη διέλευση ενός ενδοσκοπικού οργάνου (ρεζεκτοσκόπιο) μέσω της ουρήθρας. Οι μικροί όγκοι μπορούν να αφαιρεθούν *en bloc*, με τον ηλεκτροφόρο συρμάτινο βρόχο του ρεζεκτοσκοπίου, ενώ οι μεγαλύτεροι όγκοι αφαιρούνται σε πολλαπλά τμήματα. Οι ασθενείς με μη μυοδιηθητικό ΚΟΚ μπορούν να διαστρωματωθούν σε τρεις ομάδες κινδύνου ανάλογα με τον αριθμό των όγκων, το μέγεθος του όγκου (ή των όγκων), το ποσοστό υποτροπής, το στάδιο της νόσου, την παρουσία CIS και τον βαθμό κακοήθειας του όγκου, με σκοπό τον περαιτέρω καθορισμό της θεραπείας

μετά την αρχική TURBT. Οι ασθενείς με νόσο χαμηλού κινδύνου συχνά αντιμετωπίζονται οριστικά με την αρχική TURBT, εάν όλος ο όγκος αφαιρεθεί ορατά, αλλά παραμένουν υπό στενή παρακολούθηση λόγω του αυξημένου κινδύνου υποτροπής. Παρόλο που έχουν συστηθεί πολλά πρωτόκολλα παρακολούθησης, η κυστεοσκόπηση πρέπει να διενεργείται τουλάχιστον σε διάστημα τριών μηνών μετά την εγχείρηση και με φθίνουσα συχνότητα για διάστημα έως και 5 ετών. Μελλοντικές υποτροπές χαμηλού βαθμού μπορούν να αντιμετωπιστούν με TURBT ή με καυτηριασμό στο ιατρείο (με χρήση διαθερμίας).

Η υποεκτίμηση της πραγματικής παθολογίας κατά την αρχική TURBT σε ασθενείς με νόσο υψηλού κινδύνου αποτελεί σημαντικό ρίσκο. Αξίζει να σημειωθεί πως υπάρχει πιθανότητα έως και 20% ένας ασθενής με νόσο T1 να αναβαθμιστεί σε T2 ακόμη και αν κάποια ποσότητα μυός ήταν παρών στον ιστό που αρχικά εκτομήθηκε, και πιθανότητα έως και 40% να αναβαθμιστεί αν δεν υπήρχε καθόλου μυς στο αφαιρεμένο δείγμα [50]. Επιπλέον, ακόμη και αν ο όγκος σταδιοποιηθεί με ακρίβεια κατά την αρχική TURBT, υπάρχει 50% πιθανότητα ατελούς εκτομής που προκύπτει από παράγοντες όπως η πολλαπλότητα, το μέγεθος και η θέση του όγκου. Η ακριβής σταδιοποίηση είναι εξαιρετικά σημαντική στον ΚΟΚ για τον καθορισμό της κατάλληλης θεραπείας μετά την αρχική TURBT. Έτσι, συνιστάται επαναληπτική TURBT εντός 2-6 εβδομάδων σε ασθενείς με γνωστό ατελώς αφαιρεμένο όγκο ή με όγκους που διεισδύουν στον εξωστήρα μυ (T1). Επίσης, η επαναληπτική TURBT θα πρέπει να εξετάζεται σε ασθενείς με υψηλού βαθμού κακοήθειας και μη διηθητική νόσο (εκτός από το CIS) για τη βελτίωση της ακρίβειας της σταδιοποίησης και την αύξηση του χρόνου επιβίωσης χωρίς υποτροπή. Μια τυχαιοποιημένη ελεγχόμενη μελέτη (randomized controlled trial – RCT) έδειξε ότι η επαναληπτική TURBT μετά από νεο-διαγνωσθέντα ΚΟΚ σταδίου T1 βελτιώνει την επιβίωση χωρίς υποτροπή και την επιβίωση χωρίς εξέλιξη της νόσου κατά 25% και 14% στα 5 έτη, αντίστοιχα [51].

Οι εξελίξεις στην τεχνική TURBT υιοθετούνται και αξιολογούνται με αργούς ρυθμούς. Για παράδειγμα, η κυστεοσκόπηση φθορισμού με μπλε φως χρησιμοποιεί εξαμινολεβουλινάτη (hexaminolevulinate hydrochloride) – μια φωτοευαισθητοποιητική ουσία αιμάτωσης που χορηγείται 1 ώρα πριν από την κυστεοσκόπηση – για την ανίχνευση της παθολογικής συσσώρευσης φθορίζοντων προϊόντων πορφυρίνης σε κύτταρα ΚΟΚ (Εικόνα 14). Η τεχνική φαίνεται να βελτιώνει τη διάγνωση και να μειώνει τη βραχυπρόθεσμη και μακροπρόθεσμη υποτροπή. Ωστόσο, ορισμένα ζητήματα με υψηλά ποσοστά ψευδώς θετικών αποτελεσμάτων, που σαφώς μειώνονται με την εμπειρία του ιατρού, αποτελούν σημαντικούς περιορισμούς της μεθόδου. Επίσης, ο ιατροτεχνολογικός εξοπλισμός που απαιτείται είναι δαπανηρός. Παρ' όλα αυτά, η τεχνική αυτή είναι πολλά υποσχόμενη όσον αφορά τη βελτίωση της διάγνωσης και της θεραπείας κυρίως του μη μυοδιηθητικού ΚΟΚ.



Εικόνα 14: Παραδείγματα εικόνων κυστεοσκόπησης που λαμβάνονται με τη χρήση λευκού (αριστερά) και μπλε φωτός (δεξιά) για την καλύτερη διάγνωση του ΚΟΚ.

(Πηγή: <https://www.medstarhealth.org/services/blue-light-cystoscopy>)

Η διπολική ηλεκτροκαυτηρίαση είναι μια άλλη νέα τεχνική που χρησιμοποιεί λιγότερη ενέργεια και τάση από την τυπική μονοπολική καυτηρίαση βρόχου της TURBT, καθώς το κύκλωμα δεν διέρχεται μέσω του ασθενούς. Επιπλέον, μπορούν να χρησιμοποιηθούν ισοτονικά υγρά που έχουν ελάχιστες επιπτώσεις στον ορό του ασθενούς σε σύγκριση με τη μονοπολική καυτηρίαση (αν και αυτό το ζήτημα είναι λιγότερο ανησυχητικό κατά τη διάρκεια της εκτομής της ουροδόχου κύστης συγκριτικά, για παράδειγμα, με όταν η τεχνική χρησιμοποιείται για την αφαίρεση ιστού του προστάτη). Άλλες πιθανές βελτιώσεις της TURBT που διερευνώνται περιλαμβάνουν την ελαχιστοποίηση του αντανάκλαστικού του οπισθοστερνικού μυός στην πύελο, τη μείωση των περιπτώσεων διάτρησης της ουροδόχου κύστης και τη μείωση των επιπτώσεων των τεχνουργημάτων καυτηρίασης στην παθολογική εικόνα του ιστού. Ωστόσο, απαιτούνται περισσότερες κλινικές δοκιμές πριν να εξαχθούν σχετικά συμπεράσματα.

Επικουρική ενδοκυστική θεραπεία

Σε ασθενείς με νόσο ενδιάμεσου και ιδιαίτερα υψηλού ρίσκου συνιστάται η διενέργεια επικουρικής ενδοκυστικής θεραπείας (adjuvant intravesical therapy). Το σκεύασμα **βακίλλου Calmette-Guérin** (Bacillus Calmette-Guérin – BCG), ένα εμβόλιο που χρησιμοποιείται κυρίως κατά της φυματίωσης, μπορεί τόσο να μειώσει την υποτροπή, όσο και να καθυστερήσει την εξέλιξη του ΚΟΚ και φέρεται να είναι αποτελεσματικότερο θεραπευτικό σχήμα από τη χημειοθεραπεία, σύμφωνα με πολλαπλές τυχαιοποιημένες ελεγχόμενες δοκιμές και μετα-αναλύσεις [52]. Η επικουρική θεραπεία πρέπει να περιλαμβάνει θεραπεία συντήρησης για ένα έτος σε νόσο ενδιάμεσου ρίσκου και για έως και τρία έτη (εάν είναι ανεκτή) σε νόσο υψηλού ρίσκου για την επίτευξη μέγιστης αποτελεσματικότητας. Όσον αφορά τον τρόπο χορήγησής της διατίθενται διάφορα πρωτόκολλα. Η συντήρηση με ενδοκυστική θεραπεία BCG μπορεί να μειώσει τον κίνδυνο εξέλιξης της νόσου κατά 37% σε σύγκριση με τη μη συντήρηση με BCG [53]. Οι ασθενείς που εμφανίζουν επίμονη ή επιδεινούμενη νόσο μετά από μια κατάλληλη θεραπευτική αγωγή με BCG, καθώς και εκείνοι που εμφανίζουν υποτροπή της νόσου κατά τη διάρκεια της θεραπείας συντήρησης, θεωρούνται αποτυχημένα σχήματα BCG. Η αποτυχία της θεραπείας με BCG μπορεί επίσης να συμβεί σε εκείνους που εμφανίζουν υποτροπή της νόσου λόγω της ανεπαρκούς θεραπευτικής πορείας, εξαιτίας μη ανεκτών ανεπιθύμητων ενεργειών. Για τους ασθενείς αυτούς, η πιο αποτελεσματική θεραπεία είναι η ριζική κυστεκτομή, αν και μπορεί να εξεταστούν στρατηγικές διατήρησης της ουροδόχου κύστης, συμπεριλαμβανομένης της ενδοκυστικής χημειοθεραπείας, της ενδοκυστικής θεραπείας με τη βοήθεια κάποιας συσκευής και των κλινικών δοκιμών.

Μετά από την TURBT, μια εφάπαξ δόση ενδοκυστικής θεραπείας με χημειοθεραπευτικούς παράγοντες (μιτομυκίνη ή δοξορουβικίνη στις ΗΠΑ, καθώς και επιρουβικίνη ή pirarubican στην Ευρώπη) εντός 24 ωρών μπορεί να μειώσει την υποτροπή κατά 40% σε 1 έτος και κατά 15% σε 5 έτη [54]. Η ενδοκυστική χημειοθεραπεία πρέπει να χορηγείται σε ασθενείς αμέσως μετά την TURBT με θηλώδεις αλλοιώσεις, εάν δεν υπάρχει κλινική ανησυχία για μυοδιηθητικό ΚΟΚ ή διάτρηση της ουροδόχου κύστης κατά τη διάρκεια της TURBT. Εάν υπάρχει υποψία για διάτρηση της ουροδόχου κύστης, δεν θα πρέπει να χορηγείται άμεση ενδοκυστική θεραπεία λόγω του αυξημένου κινδύνου επιπλοκών. Το ογκολογικό όφελος της επικουρικής ενδοκυστικής θεραπείας είναι ιδιαίτερα χρήσιμο σε ασθενείς με όγκους χαμηλού κινδύνου, καθώς οι ασθενείς με νόσο ενδιάμεσου και υψηλού κινδύνου βρίσκονται σε μεγαλύτερο κίνδυνο υποτροπής και είναι πιθανό να χρειαστούν περαιτέρω επικουρική θεραπεία.

Ριζική κυστεκτομή

Η ριζική κυστεκτομή είναι η θεραπεία «χρυσού προτύπου» για τους ασθενείς με μυοδιηθητικό ΚΟΚ καθώς και για εκείνους με μη μυοδιηθητικό ΚΟΚ που αποτυγχάνουν στην ενδοκυστική θεραπεία, όπως αυτή περιγράφηκε παραπάνω. Επιπλέον, ορισμένοι ασθενείς με μη μυοδιηθητικό ΚΟΚ σταδίου T1 (εισβολή στη χοριακή στιβάδα) μπορούν να θεωρηθούν κατάλληλοι για «πρώιμη» κυστεκτομή. Η πρώιμη κυστεκτομή μπορεί επίσης να θεωρηθεί κατάλληλη σε ασθενείς με χαρακτηριστικά υψηλού

κινδύνου κατά τη διενέργεια της TURBT, όπως για παράδειγμα, εκείνους με λεμφαγγειακή διασπορά, με ταυτόχρονη ύπαρξη CIS, παραλλαγή ιστολογίας (ιδίως μικροθηλωματώδους μορφολογίας), μεγάλου μεγέθους (>3 cm) και πολυεστιακούς όγκους και βαθιά εισβολή στη χοριακή στιβάδα. Η ριζική κυστεκτομή συνήθως περιλαμβάνει και προστατεκτομή στους άνδρες, και υστερεκτομή και μερική εκτομή του κόλπου και της ουρήθρας στις γυναίκες.

Αν και η ριζική κυστεκτομή εκτελείται παραδοσιακά με ανοικτή χειρουργική επέμβαση, υπάρχει αυξανόμενο ενδιαφέρον για τις ελάχιστα επεμβατικές προσεγγίσεις με τη χρήση του ρομποτικού συστήματος da Vinci (Intuitive Surgical, Sunnyvale, California, USA) [55]. Τα δυνητικά οφέλη αυτής της προσέγγισης περιλαμβάνουν τη μειωμένη απώλεια αίματος και άλλα πλεονεκτήματα που είναι συνυφασμένα με την ελάχιστα επεμβατική χειρουργική, με πολλαπλές μελέτες να καταδεικνύουν ισοδύναμα ογκολογικά και λειτουργικά αποτελέσματα. Μία τυχαιοποιημένη ελεγχόμενη μελέτη συνέκρινε τη ρομποτική με την ανοικτή κυστεκτομή με διφορούμενα ευρήματα [56]. Στη μελέτη αυτή, μόνο το μέρος της επέμβασης της ριζικής κυστεκτομής εκτελέστηκε ρομποτικά και το μέρος της εκτροπής των ούρων εκτελέστηκε μέσω μιας τυπικής ανοικτής τομής. Έτσι, τα αρχικά οφέλη της ρομποτικής προσέγγισης ενδεχομένως μετριάστηκαν. Μια άλλη τυχαιοποιημένη ελεγχόμενη μελέτη που συνέκρινε την ανοικτή και τη ριζική κυστεκτομή, η μελέτη RAZOR, βρίσκεται σε εξέλιξη [57]. Ωστόσο, και πάλι, η εκτροπή μπορεί να πραγματοποιηθεί μέσω ανοικτής επέμβασης. Συμπερασματικά, απαιτούνται περαιτέρω μελέτες για τη σύγκριση της αμιγούς ρομποτικής ριζικής κυστεκτομής σε συνδυασμό με την ενδοσωματική εκτροπή (intracorporeal urinary diversion – ICUD) και της ανοικτής ριζικής κυστεκτομής πριν εξαχθούν περαιτέρω συμπεράσματα.

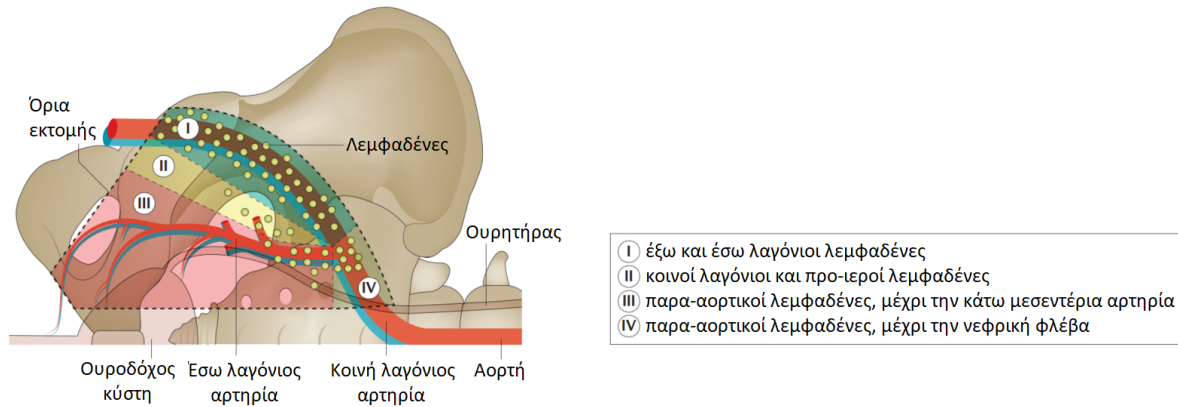
Λεμφαδενεκτομή

Ο ρόλος της λεμφαδενεκτομής (εκτομή λεμφαδένων) στην αντιμετώπιση του ΚΟΚ είναι ζωτικής σημασίας και παρόλο που η ενδεδειγμένη λεμφαδενεκτομή παρέχει αποδεδειγμένο πλεονέκτημα επιβίωσης, η συζήτηση στην επιστημονική κοινότητα παραμένει ανοικτή σχετικά με το πρότυπο που πρέπει να εφαρμόζεται και τον αριθμό των λεμφαδένων που πρέπει να αφαιρούνται. Η συνήθης διαδικασία εκτομής περιλαμβάνει την αφαίρεση των λεμφαδένων του πυελικού εδάφους κατά μήκος των έξω και έσω λαγόνιων αγγείων από την επιπολής περισπώμενη λαγόνια φλέβα μέχρι το σημείο διχασμού των λαγόνιων αγγείων και το μηριαίο γεννητικό νεύρο πλάγια προς τον ουρητήρα (**Εικόνα 15**). Στο εκτεταμένο πρότυπο, η διατομή διεξάγεται κατά μήκος των κοινών λαγόνιων αγγείων μέχρι το διχασμό της αορτής και την προϊερά περιοχή. Στο υπερ-εκτεταμένο πρότυπο, η διατομή φθάνει από την αορτή έως την κάτω μεσεντέριο αρτηρία. Δεδομένου ότι οι περισσότεροι χειρουργοί συμφωνούν ότι τουλάχιστον ένας τυπικός λεμφαδενικός καθαρισμός θα πρέπει να εκτελείται, μια τυχαιοποιημένη μελέτη (SWOG S1011) αξιολογεί τα πρότυπα του τυπικού έναντι του εκτενούς λεμφαδενικού καθαρισμού και τα αποτελέσματα αναμένονται [58].

Εκτροπή ούρων

Μετά από τη ριζική κυστεκτομή και την εκτομή των λεμφαδένων, απαιτείται η εκτροπή των ούρων για τη μόνιμη μετακίνηση της ροής των ούρων από την φυσιολογική τους έξοδο στην ουρήθρα προς άλλη κατεύθυνση. Όλες οι εκτροπές των ούρων χρησιμοποιούν ένα τμήμα του εντέρου στο οποίο αναστομώνονται οι ουρητήρες και το άλλο άκρο χρησιμοποιείται για την αποβολή των ούρων. Οι εκτροπές ούρων μπορεί να είναι είτε εγκρατείς, όπως οι εγκρατείς καθετηριαζόμενες νεοκύστες και οι ορθοτοπικές νεοκύστες, είτε μη εγκρατείς, όπως οι ειλεοστομίες [59]. Κατά τις ειλεοστομίες διενεργείται σύνδεση των ουρητήρων μέσω ενός τμήματος του εντέρου στο κοιλιακό δέρμα. Τα ούρα τρέχουν συνεχώς σε ένα κολλημένο σάκο στο στόμιο του δέρματος. Η ειλεοστομία αποτελεί μία ευρέως χρησιμοποιούμενη μέθοδο για την εκτροπή των ούρων. Οι εγκρατείς εκτροπές των ούρων μπορούν να ποικίλλουν ως προς την τεχνική. Πρακτικά, σε όλες, δημιουργείται μία νεοκύστη από διάφορα τμήματα του εντέρου και ένα κανάλι βγαίνει προς τα έξω στο επίπεδο του δέρματος, με τη χρήση κάποιου

μηχανισμού εγκράτειας ο οποίος απαιτεί καθετηριασμό για την εκκένωση. Σε μια ορθοτοπική νεοκύστη, πρακτικά διενεργείται ο σχηματισμός ενός νέου οργάνου αποθήκευσης ούρων, κυρίως από το λεπτό έντερο. Αυτός ο νέος «σάκος» αναστομώνεται στην αρχική ουρήθρα, με τους ασθενείς να αδειάζουν τα ούρα τους μέσω της μητρικής τους οδού. Αυτός ο τύπος εκτροπής απαιτεί επαρκή νεφρική λειτουργία και απουσία καρκίνου στο όριο της ουρήθρας.



Εικόνα 15: Ανατομική έκταση της τυπικής λεμφαδενεκτομής στον καρκίνο της ουροδόχου κύστης. [42]

Ανάρρωση

Μεγάλα βήματα έχουν γίνει όσον αφορά την αποκατάσταση και ανάρρωση των ασθενών μετά από ριζική κυστεκτομή, τα οποία οφείλονται στην εφαρμογή ειδικών πρωτοκόλλων ενισχυμένης αποκατάστασης μετά από χειρουργική επέμβαση (Enhanced Recovery after Surgery – ERAS) [60]. Προεγχειρητικά, τα εν λόγω πρωτόκολλα ενθαρρύνουν την ενυδάτωση με υγρά και τη χρήση φαρμάκων όπως η αλβιμοπάνη (alvimopan), η οποία είναι ένας ανταγωνιστής μ-υποδοχέων, που βοηθά στην αποκατάσταση μετά από χειρουργική επέμβαση στο έντερο. Αυτά τα σκευάσματα μετριάζουν τις ανεπιθύμητες ενέργειες σχετιζόμενες με το έντερο, οι οποίες αφορούν τα οπιοειδή φάρμακα, και αποτρέπουν την προετοιμασία ρουτίνας του εντέρου που μπορεί να οδηγήσει σε αφυδάτωση και την επακόλουθη ανάγκη για ενδοφλέβια χορήγηση υγρών που μπορεί με τη σειρά τους να οδηγήσουν σε οίδημα του εντέρου και ειλεό (απόφραξη). Κατά τη διάρκεια της εγχείρησης, τα πρωτόκολλα ERAS ελαχιστοποιούν την αναζωογόνηση με υγρά και συνδράμουν στην αφαίρεση των τυχόντων ρινογαστρικών σωλήνων πριν από την αποσωλήνωση. Μετεγχειρητικά, τα πρωτόκολλα ενθαρρύνουν την πρόωπη βάδιση, την πρόωπη σίτιση (συνήθως με κανονική διαίτα μέχρι τη δεύτερη μετεγχειρητική ημέρα) καθώς και την υποκατάσταση και ελαχιστοποίηση των οπιοειδών φαρμακευτικών σκευασμάτων με εναλλακτικές λύσεις όπως η κετορολάκη (ketorolac), παρακεταμόλη (paracetamol) και τραμαδόλη (tramadol). Τα πρωτόκολλα ERAS φαίνεται να συμβάλλουν σημαντικά στη μείωση της διάρκειας παραμονής στο νοσοκομείο και στην ταχύτερη ανάρρωση.

Επιβίωση

Μετά από τη ριζική κυστεκτομή, τα αποτελέσματα επιβίωσης εξαρτώνται σε μεγάλο βαθμό από την τελική παθολογική σταδιοποίηση του όγκου. Η δεκαετής επιβίωση χωρίς υποτροπή για ασθενείς στους οποίους δεν υπήρχε διήθηση των λεμφαδένων είναι 86% για παθολογοανατομικά επιβεβαιωμένους όγκους σταδίου T0, 76% για όγκους σταδίου T1-pT3a, 61% για όγκους σταδίου T3b και 45% για όγκους σταδίου T4, αλλά πέφτει στο 34% ανεξαρτήτως σταδίου όταν εμπλέκονται οι λεμφαδένες [61, 62]. Ο ΚΟΚ παραμένει δυστυχώς μία θανατηφόρα ασθένεια. Πράγματι, παρά τη ριζική κυστεκτομή με φαινομενικά καλά ογκολογικά αποτελέσματα, ακόμη και σε ασθενείς με φαινομενικά οργανικά περιορισμένη νόσο, αρνητικά χειρουργικά όρια και αρνητικούς λεμφαδένες, πολλοί από αυτούς τους

ασθενείς εξακολουθούν να εμφανίζουν υποτροπή, γεγονός που ενισχύει την ανάγκη χρήσης επικουρικών θεραπειών.

Νεοεπικουρική χημειοθεραπεία

Ο ρόλος της νεοεπικουρικής χημειοθεραπείας (neoadjuvant chemotherapy) πριν από τη ριζική κυστεκτομή, για τη βελτίωση της επιβίωσης, έχει διερευνηθεί σε αρκετές τυχαιοποιημένες ελεγχόμενες μελέτες και μετα-αναλύσεις. Μια μετα-ανάλυση 11 μελετών ανέφερε πλεονέκτημα επιβίωσης 5% στα πέντε έτη [63]. Αυτές οι μελέτες χρησιμοποίησαν χημειοθεραπεία με βάση την πλατίνα (platinum). Επί του παρόντος, δεν υπάρχουν επαρκή δεδομένα για την υποστήριξη θεραπειών με βάση διαφορετικές ουσίες στο πλαίσιο της νεοεπικουρικής θεραπείας. Ορισμένες μελέτες έχουν δείξει παρόμοια αποτελεσματικότητα με τον λιγότερο τοξικό συνδυασμό γεμισιταβίνης – σισπλατίνης (gemcitabine – cisplatin), ο οποίος έχει προταθεί για χρήση στο πλαίσιο νεοεπικουρικής θεραπείας [64]. Συνεπώς, σε ασθενείς με μυοδιηθητικό ΚΟΚ (και ιδίως σε εκείνους με κλινική νόσο σταδίου > T3), θα πρέπει να χορηγείται νεοεπικουρική χημειοθεραπεία, εάν είναι σε θέση να ανεχθούν το θεραπευτικό σχήμα με βάση την πλατίνα. Εάν η χημειοθεραπεία με βάση την πλατίνα δεν μπορεί να γίνει ανεκτή από τους ασθενείς, τότε αυτοί θα πρέπει να προχωρούν άμεσα σε ριζική κυστεκτομή.

Επικουρική θεραπεία

Ο ρόλος της επικουρικής θεραπείας (adjuvant therapy) μετά τη ριζική κυστεκτομή παραμένει ασαφής, επειδή οι μεμονωμένες μελέτες μέχρι σήμερα δεν έχουν επαρκή ισχύ. Μια πρόσφατη τυχαιοποιημένη ελεγχόμενη μελέτη δεν κατέδειξε σαφές όφελος στην αναβαλλόμενη έναντι της άμεσης χημειοθεραπείας σε ασθενείς με παθολογική νόσο σταδίου > T3 ή νόσο με συμμετοχή των λεμφαδένων [65]. Ωστόσο, μια μετα-ανάλυση που περιλάμβανε συνολικά 945 ασθενείς από 9 τυχαιοποιημένες ελεγχόμενες μελέτες έδειξε σχετική μείωση του κινδύνου θανάτου κατά 23% με χρήση επικουρικής θεραπείας γεμισιταβίνης – σισπλατίνης σε σύγκριση με εφαρμογή μονάχα της χειρουργικής επέμβασης [66]. Ομοίως, από επτά μελέτες που ανέφεραν αυτό το τελικό σημείο προέκυψε μείωση του κινδύνου υποτροπής της νόσου κατά 34%, η οποία ήταν υψηλότερη για τους ασθενείς με θετική συμμετοχή λεμφαδένων [67]. Τα πρώτα στοιχεία υποδηλώνουν ότι η επικουρική ακτινοθεραπεία μπορεί να έχει ρόλο σε ασθενείς με δυσμενή παθολογικά χαρακτηριστικά (παθολογικό στάδιο νόσου > T3, συμμετοχή λεμφαδένων ή θετικά χειρουργικά όρια) στη ριζική κυστεκτομή και διερευνάται περαιτέρω έως και σήμερα.

Πρωτόκολλα διατήρησης της ουροδόχου κύστης στον μυοδιηθητικό ΚΟΚ

Παρόλο που η ριζική κυστεκτομή παραμένει το «χρυσό πρότυπο» για την ογκολογική αποτελεσματικότητα σε ασθενείς με μυοδιηθητικό ΚΟΚ, η χειρουργική επέμβαση μπορεί να έχει σημαντικές επιπτώσεις στην ποιότητα ζωής των ασθενών. Επιπλέον, ορισμένοι ασθενείς δεν είναι κατάλληλοι για αυτή τη χειρουργική επέμβαση. Έτσι, έχουν διερευνηθεί διάφορες επιλογές διατήρησης της ουροδόχου κύστης που θα μπορούσαν, εάν εφαρμοστούν με ιδιαίτερη προσοχή, να ωφεληθούν επιλεγμένους ασθενείς. Η ριζική διουρηθρική εκτομή του όγκου (TURBT) περιλαμβάνει μια επιθετική διαδικασία ενδοσκοπικής εκτομής με σκοπό την εκτομή όλης της ορατής νόσου (προσέχοντας να μην προκληθεί διάτρηση στην κύστη). Στη διαδικασία αυτή χρησιμοποιείται επίσης καυτηρίαση (είτε μονοπολική είτε διπολική) για την εκτομή ιστού όσο το δυνατόν βαθύτερα και ευρύτερα, σε μια προσπάθεια να καταστραφεί όσο το δυνατόν περισσότερο μέρος του όγκου – ιδανικό ολόκληρο – με ασφάλεια. Η ριζική διουρηθρική εκτομή του όγκου ως αυτόνομη θεραπευτική προσέγγιση μπορεί να έχει μια διαρκή επίδραση στην επιβίωση, αλλά θα πρέπει να εξετάζεται μόνο για ασθενείς που δεν είναι επιλέξιμοι για ριζική κυστεκτομή ή χημειοακτινοθεραπεία. Σε μία πρόσφατη μελέτη, τα δεδομένα επιβίωσης 133 ασθενών αποκάλυψαν ικανοποιητική ειδική της νόσου επιβίωση (76,7%) και επιβίωση χωρίς εξέλιξη με διατήρηση της ουροδόχου κύστης (57,8%) στα 15 έτη [68]. Ωστόσο, η πλήρης εκτομή

του όγκου (που επιβεβαιώνεται με αρνητική βιοψία της βάσης του όγκου) είναι ζωτικής σημασίας για αυτόν τον τύπο διαχείρισης .

Η μερική κυστεκτομή μπορεί να εξεταστεί προσεκτικά για ασθενείς με όγκο μικρού μεγέθους (< 3 cm), συμπαγή όγκο στο θόλο της ουροδόχου κύστης, χωρίς σχετιζόμενο CIS ή όγκους σε εκκολπώματα. Σε μία μελέτη συγκρίθηκαν 86 ασθενείς που υποβλήθηκαν σε μερική κυστεκτομή με ασθενείς που υποβλήθηκαν σε ριζική κυστεκτομή, σε αντιστοιχία κατά ζεύγη (αναλογία 1:2), αλλά δεν διαπιστώθηκε σημαντική διαφορά στην δεκαετή επιβίωση ειδική της νόσου (58% έναντι 63%, τιμή-p = 0,63) ή στη συνολική επιβίωση (36% έναντι 39%, τιμή-p = 0,67) [69].

Η τρίπτυχη θεραπεία (trimodal therapy) συνδυάζει τη ριζική διουρηθρική εκτομή του όγκου, με ταυτόχρονη χημειο-ράδιο-θεραπεία και εξωτερική ακτινοθεραπεία. Αυτή η στρατηγική μπορεί να οδηγήσει σε πενταετή ποσοστά ειδική της νόσου επιβίωσης 50-82% [70]. Ωστόσο, το 25-30% των ασθενών αυτών χρειάζονται τελικά ριζική κυστεκτομή λόγω της αποτυχίας ανταπόκρισής τους στη θεραπεία. Είναι σημαντικό ότι η τρίπτυχη θεραπεία περιορίζει σημαντικά την πιθανότητα μελλοντικής εκτροπής των ούρων με εφαρμογή ορθοτοπικής νεοκύστης, σε περίπτωση που απαιτηθεί ριζική κυστεκτομή διάσωσης, γεγονός το οποίο αποτελεί έναν περιορισμό που πρέπει να συζητηθεί κατά τη συμβουλευτική των ασθενών.

Μεταστατική νόσος

Οι ασθενείς που διαγιγνώσκονται με μεταστατικό ΚΟΚ κατά τη διάρκεια της αρχικής εξέτασης ή μετά από ριζική κυστεκτομή αντιμετωπίζονται με συστηματική χημειοθεραπεία. Τα συνήθη σχήματα πρώτης γραμμής βασίζονται στη σισπλατίνη και περιλαμβάνουν τους ακόλουθους συνδυασμούς: μεθοτρεξάτη, βινβλαστίνη, δοξορουβικίνη και σισπλατίνη (methotrexate, vinblastine, doxorubicin, cisplatin – MVAC) – σισπλατίνη, μεθοτρεξάτη και βινβλαστίνη (cisplatin, methotrexate, vinblastine – CMV) – γεμισιταβίνη και σισπλατίνη (gemcitabine, cisplatin) – γεμισιταβίνη, σισπλατίνη και πακλιταξέλη (gemcitabine, cisplatin, paclitaxel) [71]. Οι ασθενείς που δεν είναι κατάλληλοι για θεραπεία με βάση τη σισπλατίνη μπορούν να λάβουν θεραπεία με καρβοπλατίνη (carboplatin) ή με ταξάνες (taxanes), οι οποίες όμως ενδέχεται να είναι κατώτερες σε αποτελεσματικότητα. Σε ασθενείς με εξέλιξη της νόσου μετά από χημειοθεραπεία με βάση την πλατίνη, υπάρχουν ορισμένες επιλογές δεύτερης γραμμής, με μικτά ωστόσο αποτελέσματα.

Ο ρόλος της συστηματικής ανοσοθεραπείας σε ασθενείς με μεταστατικό ΚΟΚ έχει αποδώσει ήδη κάποια ελπιδοφόρα αρχικά αποτελέσματα με τη χρήση των αναστολέων σημείων ελέγχου (checkpoint inhibitors) [72]. Ένα τέτοιο φάρμακο, η ατεζολιζουμάμπη (atezolizumab), το οποίο είναι ένα Fc-βιοτεχνολογικά παρασκευασμένο, εξανθρωποποιημένο μονοκλωνικό αντίσωμα G1 – IgG1 – που δεσμεύεται απευθείας στο PD-L1 και παρέχει διπλό αποκλεισμό των υποδοχέων PD-1 και B7.1, έχει εγκριθεί για θεραπεία δεύτερης γραμμής και τώρα διερευνάται και για θεραπεία πρώτης γραμμής [73]. Εν τω μεταξύ, βρίσκονται σε εξέλιξη μελέτες με χρήση ουσιών όπως η πεμπρολιζουμάμπη (pembrolizumab), η νιβολουμάμπη (nivolumab) και ιπιλιμουμάμπη (ipilimumab) σε ασθενείς με μεταστατικό ΚΟΚ.

Ποιότητα ζωής

Οι ασθενείς με ΚΟΚ αντιμετωπίζουν πολυάριθμες πιθανές προκλήσεις όσον αφορά την ποιότητα ζωής τους, αφενός λόγω της ίδιας της νόσου και αφετέρου λόγω των επιπλοκών των διαθέσιμων θεραπειών. Μετά από τις εκάστοτε θεραπευτικές παρεμβάσεις είναι πιθανό να επηρεαστούν η συνολική λειτουργία και ευεξία των ασθενών, καθώς και συγκεκριμένοι τομείς της σεξουαλικής, ουροποιητικής και εντερικής λειτουργίας τους [74]. Επιπλέον, τα υψηλά ποσοστά υποτροπής της συγκεκριμένης νόσου και η πιθανότητα εξέλιξής της διαμορφώνουν ένα από τα υψηλότερα κόστη ανά άτομο συγκριτικά με όλους τους τύπους καρκίνου. Οι ασθενείς με μη μυοδιηθητικό ΚΟΚ αντιμετωπίζουν υψηλό ποσοστό

υποτροπής, χαμηλά αλλά ποικίλα ποσοστά εξέλιξης και συχνή – ή ακόμα και δια βίου – επεμβατική παρακολούθηση με κυστεοσκόπηση καθώς και ενδοκυστικές θεραπείες. Οι ασθενείς με μυοδιηθητικό ΚΟΚ έρχονται αντιμέτωποι με τις δυνητικές επιλογές που αλλάζουν ολοκληρωτικά τη ζωή τους, όπως αυτή της ριζικής κυστεκτομής η οποία έχει ως αποτέλεσμα τη μόνιμη εκτροπή της ούρησης.

Ένας αυξανόμενος όγκος βιβλιογραφίας περιγράφει την ποιότητα ζωής των ασθενών με ΚΟΚ. Τα περισσότερα δεδομένα αφορούν τον πληθυσμό των ασθενών με μυοδιηθητικό ΚΟΚ, ειδικά σε θέματα σχετικά με τις επιπτώσεις της μεθόδου εκτροπής της ούρησης στην οποία υποβλήθηκαν. Για παράδειγμα, μια συστηματική ανασκόπηση σε μελέτες που δημοσιεύθηκαν από το 2012 και μετά υποδεικνύει ευνοϊκότερα αποτελέσματα για την εγκρατή ορθοτοπική εκτροπή από ό,τι για την ειλεοστομία [75]. Η υπάρχουσα βιβλιογραφία περιορίζεται από την ετερογένεια του σχεδιασμού των μελετών, από τα διαφορετικά ερωτηματολόγια, τη συχνά αναδρομική ή δια-τομεακή συλλογή δεδομένων και τα λίγα βασικά δεδομένα αναφοράς. Επί του παρόντος, έχουν περιγραφεί πέντε εργαλεία της μέτρησης των αποτελεσμάτων που αναφέρουν οι ασθενείς με ΚΟΚ (patient-reported outcomes – PROs) και τρία από αυτά έχουν επικυρωθεί [76]. Στόχος είναι η προώθηση της καταγραφής της εμπειρίας των ασθενών σχετικά με όλες τις διενεργούμενες διαδικασίες, παρεμβάσεις και θεραπείες για τον ΚΟΚ. Τα ερωτηματολόγια επικεντρώνονται στην εμπειρία των ατόμων με μη μυοδιηθητικό ή μυοδιηθητικό ΚΟΚ. Μια πρόσφατη ανασκόπηση περιγράφει λεπτομερώς τα μεθοδολογικά και τα σχετικά με τη νόσο πλεονεκτήματα αλλά και τους περιορισμούς αυτών των εργαλείων [77].

Πρόσφατα τέθηκαν σε εφαρμογή πρωτόκολλα γρήγορης αποκατάστασης των ασθενών (Enhanced Recovery After Surgery – ERAS) που εστιάζουν στην περιεγχειρητική φροντίδα, και συγκεκριμένα σε θέματα της μειωμένης προετοιμασίας του εντέρου, του τυποποιημένου προγράμματος σίτισης καθώς και του αναλγητικού σχήματος [78]. Στη μοναδική τυχαιοποιημένη κλινική μελέτη που έχει διεξαχθεί μέχρι σήμερα, το πρωτόκολλο ERAS μείωσε τη μετεγχειρητική νοσηρότητα όσον αφορά τις διαταραχές επούλωσης των τραυμάτων, τον πυρετό και τη θρόμβωση σε σύγκριση με ένα συντηρητικό σχήμα. Ο χρόνος που δαπανήθηκε στη μονάδα ενδιάμεσης φροντίδας μειώθηκε σημαντικά και η ποιότητα ζωής, η οποία προσδιορίστηκε με τη χρήση του ερωτηματολογίου EORTC QLQ-30, ήταν σημαντικά καλύτερη μετεγχειρητικά για τα άτομα που ακολουθούσαν το πρωτόκολλο ERAS [79]. Ωστόσο, χρειάζονται μεγαλύτερες μελέτες για να επιβεβαιωθούν τα παραπάνω αποτελέσματα. Η βελτιωμένη περιεγχειρητική διαχείριση φαίνεται πως μπορεί να μετριάσει πιθανώς μέρος της νοσηρότητας που σχετίζεται με τις διαδικασίες εκτροπής της ούρησης.

Ένας αυξανόμενος όγκος εμπειρίας υποστηρίζει την αξία της συλλογής των αποτελεσμάτων που αναφέρουν οι ασθενείς (PROs) στη συνήθη ιατρική περίθαλψη, ιδανικά ενσωματωμένων στον ηλεκτρονικό ιατρικό φάκελο (ePROs) [80]. Πολλαπλές προοπτικές μελέτες έχουν καταδείξει τη βελτίωση της ικανοποίησης των ασθενών, της διαχείρισης των συμπτωμάτων, της ποιότητας ζωής και της επικοινωνίας μεταξύ ασθενών και ιατρών με την ενσωμάτωση των ePROs στη συνήθη φροντίδα για τον καρκίνο. Καθώς οι προσπάθειες αυτές κερδίζουν έδαφος και τα δεδομένα των αποτελεσμάτων που αναφέρουν οι ασθενείς συλλέγονται πλέον συστηματικότερα, η φροντίδα του ΚΟΚ μπορεί να βελτιωθεί σε πολλαπλές διαστάσεις. Επίσης, η συνήθης συλλογή δεδομένων των αποτελεσμάτων που αναφέρουν οι ασθενείς θα μπορούσε να σωρευτεί με την πάροδο του χρόνου σε διάφορους πληθυσμούς ασθενών, ώστε να γίνει αποτελεσματικότερη η κατανόηση της συγκριτικής αποδοτικότητας των διαφόρων θεραπευτικών επιλογών για τον ΚΟΚ. Τελικά, η περαιτέρω ανάπτυξη σε αυτόν τον τομέα θα μπορούσε να συμβάλει στην καθιέρωση μέτρων ποιοτικής επίδοσης στον ΚΟΚ στηριζόμενοι στις αναφορές των ασθενών, αλλά και με βάση τα πρότυπα που έχουν θεσπιστεί από τους ρυθμιστικούς φορείς.

Προοπτικές

Παρόλο που η χρήση καπνικών προϊόντων μειώνεται σε ορισμένα μέρη του κόσμου, οι επιπτώσεις της έκθεσης στον καπνό μπορεί να διαρκέσουν για δεκαετίες. Αυτό έχει ως αποτέλεσμα η επίπτωση του ΚΟΚ να είναι αρκετά σταθερή και να δύναται να αυξηθεί καθώς ο πληθυσμός γηράσκει. Με απώτερο στόχο τη βελτίωση του χρόνου επιβίωσης των ασθενών με ΚΟΚ, είναι επιτακτική μια πιο έγκαιρη ανίχνευση και διάγνωση, ένας αποτελεσματικότερος τοπικός έλεγχος της νόσου ή/και ανάπτυξη νέων θεραπειών για τη μεταστατική νόσο.

Διάγνωση

Τα τελευταία χρόνια φαίνεται να αναδύονται δύο τεχνικές για την έγκαιρη ανίχνευση του ΚΟΚ. Η πρώτη τεχνική αφορά στον προσυμπτωματικό έλεγχο των ασθενών υψηλού κινδύνου, ο οποίος έχει αποδειχθεί δύσκολος λόγω της χαμηλής επίπτωσης της νόσου αλλά και της έλλειψης γενετικών δεικτών για τον προσδιορισμό της επιρρέπειας. Είναι γεγονός πως το κόστος του σχεδιασμού μιας τυχαιοποιημένης ελεγχόμενης μελέτης για την απόδειξη του οφέλους επιβίωσης μέσω του προσυμπτωματικού ελέγχου είναι εξαιρετικά υψηλό. Η δεύτερη τεχνική αφορά στη βελτίωση της ανίχνευσης μέσω της αύξησης της συμμόρφωσης ιατρών και ασθενών με τις τρέχουσες κατευθυντήριες γραμμές για την αξιολόγηση της αιματοουρίας. Οι τρέχουσες κατευθυντήριες γραμμές συχνά αγνοούνται, αλλά είναι πιθανό ότι η συμπερίληψη εργαλείων διαστρωμάτωσης κινδύνου με χρήση κλινικών πληροφοριών ή καρκινικών βιοδεικτών στα ούρα να μπορούν να βελτιώσουν την έγκαιρη ανίχνευση.

Χειρουργική επέμβαση

Η τεχνολογική πρόοδος στην απεικόνιση των όγκων με τη χρήση της φθορίζουσας κυστεοσκόπησης και της απεικόνισης στενής ζώνης είναι πιθανό να βελτιώσει την αποτελεσματικότερη ανίχνευση των όγκων και να μειώσει τις υποτροπές λόγω ανεπαρκούς αρχικής εκτομής. Επιπλέον, η ριζική χειρουργική επέμβαση κυστεκτομής και η έκταση της λεμφαδενεκτομής για το μυοδιηθητικό ΚΟΚ εξακολουθούν να εξελίσσονται. Πρόσφατες μελέτες έχουν δείξει ότι η έκταση της λεμφαδενεκτομής είναι πιο σημαντική από τον αριθμό των λεμφαδένων που θα αφαιρεθούν [81]. Πιο συγκεκριμένα, δύο πρόσφατες μετα-αναλύσεις αποκάλυψαν ότι η εκτεταμένη λεμφαδενεκτομή μπορεί να συμβάλει στην πενταετή ελεύθερης υποτροπής επιβίωση χωρίς να αυξάνει τα συνολικά ποσοστά των μετεγχειρητικών επιπλοκών, παρά το γεγονός ότι οι ασθενείς της εκτεταμένης λεμφαδενεκτομής είχαν χειρότερους προεγχειρητικούς δείκτες σε σύγκριση με την ομάδα της τυπικής λεμφαδενεκτομής [82, 83]. Οι τυχαιοποιημένες ελεγχόμενες μελέτες που μελετούν το ζήτημα των βέλτιστων χειρουργικών ανατομικών ορίων βρίσκονται ακόμα σε εξέλιξη.

Επίσης, από την εισαγωγή του χειρουργικού ρομπότ da Vinci, πριν από δύο και πλέον δεκαετίες, έχουν συγκεντρωθεί αρκετά στοιχεία σχετικά με τα αποτελέσματα της ρομποτικά υποβοηθούμενης ριζικής κυστεκτομής σε σύγκριση με τη συμβατική ανοικτή χειρουργική επέμβαση. Τρεις τυχαιοποιημένες ελεγχόμενες δοκιμές έδειξαν ότι μόνο η απώλεια αίματος κατά τη διάρκεια της επέμβασης πλεονεκτούσε στη ρομποτικά υποβοηθούμενη ριζική κυστεκτομή σε σύγκριση με την ανοικτή ριζική κυστεκτομή [84, 85]. Ωστόσο, οι μελέτες αυτές συμπεριλάμβαναν περιορισμένο αριθμό ασθενών. Επιπλέον, η μελέτη του Bochner κ.α. τερματίστηκε πρόωρα αφού η ανάλυση με πρόθεση θεραπείας δεν κατάφερε να καταδείξει μείωση κατά 20% των περιεγχειρητικών επιπλοκών σύμφωνα με την αρχική υπόθεση [56]. Παρά την έλλειψη σημαντικής υπεροχής της ρομποτικά υποβοηθούμενης χειρουργικής επέμβασης σε αυτές τις πρώιμες μελέτες, αναμένονται τα αποτελέσματα της δια-ιδρυματικής, τυχαιοποιημένης, προοπτικής, μη ήσσονος φάσης III μελέτης RAZOR, πριν από την τελική εξαγωγή συμπερασμάτων και την κρίση σχετικά με το μέλλον της ρομποτικά υποβοηθούμενης χειρουργικής, όσον αφορά τα περιεγχειρητικά και ογκολογικά αποτελέσματα στην αντιμετώπιση του μυοδιηθητικού ΚΟΚ [57].

Ενδοκυστικές θεραπείες

Οι όγκοι συχνά υποτροπιάζουν παρά την επαγωγική και συντηρητική θεραπεία με BCG και η απουσία αποτελεσματικών θεραπειών σε αυτή την περίπτωση έχει ως συνέπεια τη σύσταση της ολικής κυστεκτομής σε πολλούς ασθενείς πριν από την ανάπτυξη μυοδιηθητικής νόσου. Κατά συνέπεια, αξιολογούνται νέες θεραπείες, για χρήση σε ασθενείς με νόσο που δεν ανταποκρίνεται στη θεραπεία με BCG, ορισμένες από τις οποίες περιλαμβάνουν τη χρήση ιών για θεραπευτικούς σκοπούς. Για τον ΚΟΚ διερευνώνται ογκολυτικοί ιοί που στοχεύουν τα καρκινικά κύτταρα [86]. Επιπλέον, διερευνάται η χρήση τροποποιημένου αδενοϊού, που παράγει ιντερφερόνη α2b (IFNα2b), με σκοπό τη διέγερση του ανοσοποιητικού συστήματος του ξενιστή με ΚΟΚ [87]. Σε μια μελέτη φάσης Ι, η γονιδιακή θεραπεία με τη μεσολάβηση αδενοϊού IFNα2b χορηγήθηκε με ενστάλαξη στην ουροδόχο κύστη σε 17 ασθενείς οι οποίοι δεν ανταποκρίνονταν στη θεραπεία BCG [88]. Από τους 14 ασθενείς που έλαβαν τη θεραπεία με επαρκή δόση, έξι (43%) παρουσίασαν πλήρη ανταπόκριση με μέση διάρκεια 31 μηνών και δύο ήταν ελεύθεροι νόσου κατά την τελευταία παρακολούθηση. Σε μια άλλη μελέτη, πέντε από τους επτά ασθενείς που υποβλήθηκαν σε θεραπεία με αυξημένη δόση αυτής της γονιδιακής θεραπείας ήταν ελεύθεροι νόσου μετά από ελάχιστη παρακολούθηση 23,9 μηνών [89].

Συστηματικές θεραπείες

Όσον αφορά τις συστηματικές θεραπείες, εξετάζεται κατά πόσον οι νέοι μοριακοί υπότυποι που ανακαλύπτονται επηρεάζουν την ανταπόκριση στη χημειοθεραπεία. Μια μελέτη κατέδειξε ότι ένας σημαντικός αριθμός ασθενών με υψηλού βαθμού νόσο (67%) που εμφάνιζε τον αδενικό κυτταρικό τύπο ήταν ευαίσθητοι στη νεοεπικουρική χημειοθεραπεία MVAC, ενώ οι ασθενείς με τον υπότυπο *TP53* ήταν ανθεκτικοί στη συγκεκριμένη νεοεπικουρική χημειοθεραπεία [90]. Μία άλλη ερευνητική ομάδα ανέφερε ότι οι σωματικές μεταλλάξεις στο γονίδιο *ERCC2* συσχετίζονται με την ανταπόκριση στη θεραπεία με βάση τη σισπλατίνη στον μυοδιηθητικό ΚΟΚ [91]. Κατά συνέπεια, η πρόβλεψη της ανταπόκρισης στη νεοεπικουρική χημειοθεραπεία χρησιμοποιώντας το γενετικό υπόβαθρο του όγκου που έχει αφαιρεθεί αποτελεί ένα νέο ελπιδοφόρο και πολλά υποσχόμενο επιστημονικό πεδίο για την έρευνα του καρκίνου. Πράγματι, μια πρόσφατη κλινική δοκιμή νεοεπικουρικής θεραπείας έχει ξεκινήσει για τη σύγκριση του σχήματος γεμισταβίνης – σισπλατίνης με το σχήμα MVAC και τη μελέτη της ικανότητα ενός αλγορίθμου βασισμένου στο προφίλ της γονιδιακής έκφρασης να προβλέψει την πλήρη παθολογική ανταπόκριση [92].

Στην ανάλυση του Γονιδιακού Άτλαντα του Καρκίνου, το 69% των όγκων της ουροδόχου κύστης έφεραν μεταλλάξεις δυνητικά επιδεκτικές σε φαρμακευτική αντιμετώπιση, συμπεριλαμβανομένου ενός ποσοστού 42% με μεταβολές στο μονοπάτι *PI3K-AKT-mTOR* και ενός ποσοστού 45% με μεταβολές στα μονοπάτια που ρυθμίζονται από την κινάση ενεργοποιούμενη από μιτογόνα (mitogen-activated protein kinase – MAPK) (συμπεριλαμβανομένου του γονιδίου *HER2*) [93]. Οι ασθενείς με συγκεκριμένες στοχεύσιμες μεταλλάξεις μπορούν να έχουν καλές μακροχρόνιες ανταποκρίσεις σε στοχευμένες θεραπείες, οι οποίες είναι αναποτελεσματικές όταν χορηγούνται σε ασθενείς χωρίς τις μεταλλάξεις. Μελλοντικές κλινικές δοκιμές, όπως το NCI-MATCH, αναμένεται να ρίξουν φως στις αποτελεσματικές στρατηγικές για την εξατομικευμένη θεραπεία των ασθενών και για τον εντοπισμό παραγόντων που είναι αποτελεσματικοί σε επιλεγμένους ασθενείς [94].

Όπως αναφέρθηκε εν συντομία, οι αναστολείς σημείων ελέγχου διερευνώνται ως δυνητικά υποσχόμενες και αποτελεσματικές θεραπείες στον ΚΟΚ, με βάση τις προόδους και σε άλλους τύπους καρκίνου. Οι κυριότεροι στόχοι είναι το κυτταροτοξικό T-λεμφοκυτταρικό αντιγόνο 4 (*CTLA4*), η προγραμματισμένη πρωτεΐνη κυτταρικού θανάτου 1 (*PD1*) και ο προγραμματισμένος υποκαταστάτης θανάτου 1 (*PDL1*), πρωτεΐνες οι οποίες συμμετέχουν στην εξασθένιση των φλεγμονωδών ανοσολογικών αποκρίσεων με σκοπό την πρόληψη της ανεξέλεγκτης καταστροφικής φλεγμονής [95]. Ένας άλλος πολλά υποσχόμενος τομέας έρευνας είναι ο συνδυασμός της ακτινοθεραπείας με την

ανοσοθεραπεία. Η στρατηγική αυτή αποσκοπεί στη χρήση ανοσοθεραπευτικών παραγόντων, όπως οι αναστολείς σημείων ελέγχου, για την ενίσχυση της ανοσολογικής απόκρισης (η λεγόμενη καταιγίδα κυτταροκινών) που πυροδοτείται από τον θάνατο των καρκινικών κυττάρων που προκαλείται από την ακτινοθεραπεία. Μελέτες με χρήση πεμπρολιζουμάμπης (pembrolizumab) σε μη μυοδιηθητικό ΚΟΚ υψηλού κινδύνου και υποτροπιάζοντα μη μυοδιηθητικό ΚΟΚ βρίσκονται σε εξέλιξη [96], αλλά επί του παρόντος δεν υπάρχουν κλινικά δεδομένα σε συνδυασμό με ακτινοθεραπεία.

Η ατεζολιζουμάμπη (atezolizumab) είναι ένα μονοκλωνικό αντίσωμα έναντι της πρωτεΐνης *PDI* που εγκρίθηκε από τον Οργανισμό Φαρμάκων των ΗΠΑ τον Μάιο του 2016 για ασθενείς με μεταστατικό ΚΟΚ που δεν ανταποκρίνονται στη χημειοθεραπεία με βάση την πλατίνα [97]. Σε μια μελέτη φάσης II (n = 310 ασθενείς), προσδιορίστηκαν τέσσερις κατηγορίες ανοσολογικής διήθησης κυττάρων (immune cells): IC3 (όπου $\geq 10\%$ της περιοχής του όγκου διηθείται), IC2 ($\geq 5\%$ και $< 10\%$), IC1 ($\geq 1\%$ και $< 5\%$) και IC0 ($< 1\%$). Η ατεζολιζουμάμπη είχε ως αποτέλεσμα τη σημαντική βελτίωση της απόκρισης (με βάση τα κριτήρια του Response Evaluation Criteria in Solid Tumours – RECIST) για κάθε προκαθορισμένη ομάδα (IC2-3: 27% (95% CI: 19-37% / τιμή-p $< 0,0001$) και IC1-3: 18% (95% CI: 13-24% / τιμή-p = 0,0004)) και σε όλους τους ασθενείς (15% (95% CI: 11-20%, τιμή-p = 0,0058)) σε σύγκριση με ένα ποσοστό συνολικής ανταπόκρισης 10% σε ιστολογικό έλεγχο [98]. Τα κύτταρα του ανοσολογικού που διηθούν τον όγκο (tumor-infiltrating immune cells) και εκφράζουν σε υψηλά επίπεδα το γονίδιο *PDL1* συσχετίστηκαν με υψηλότερα ποσοστά ανταπόκρισης στην ατεζολιζουμάμπη. Επιπλέον, το 84% των ασθενών που ανταποκρίθηκαν στη θεραπεία παρουσίασαν συνεχιζόμενη ανταπόκριση μετά από διάμεση παρακολούθηση των 12 μηνών. Ωστόσο, παρά τα ανωτέρα ευρήματα, υπάρχουν ορισμένες επιφυλάξεις για τη χρήση του γονιδίου *PDL1* ως προγνωστικού βιοδείκτη και για τη χρήση των αναστολέων *PDI* ή *PDL1*. Οι επιφυλάξεις αυτές αποδίδονται στην πληθώρα των ειδικών για το γονίδιο *PDL1* αντισωμάτων, μικροσυστοιχιών, συστημάτων βαθμονόμησης και ορίων θετικότητας. Στο μέλλον, αυτές οι κατηγορίες φαρμάκων μαζί με στοχευμένους θεραπευτικούς παράγοντες μπορεί να προσφέρουν περισσότερες δυνατότητες για τη θεραπεία ασθενών με σε προχωρημένο στάδιο της νόσου.

Συνολικά, οι επιστημονικές έρευνες για την θεραπεία του ΚΟΚ είναι πολλά υποσχόμενες, με σημαντικές προόδους να σημειώνονται όσον αφορά την κατανόηση και τη διαχείριση του ΚΟΚ. Οι προοπτικές αυτές προσφέρουν σημαντικές ελπίδες για την αποτελεσματικότερη διαχείριση της ασθένειας και την βελτίωση της επιβίωσης και της ποιότητας ζωής των ασθενών.

1.3 Βιοδείκτες για τον καρκίνο ουροδόχου κύστης

Δεδομένου του ότι η κυστεοσκόπηση αποτελεί μία επεμβατική και σχετικά κοστοβόρα διαδικασία και η κυτταρολογική εξέταση περιορίζεται από χαμηλή ευαισθησία, κυρίως για τα πρώιμα στάδια της νόσου και για τους χαμηλού βαθμού όγκους, έχουν εφευρεθεί νέα τεστ ανάλυσης ούρων για την ανίχνευση καρκινικών βιοδεικτών. Τα τεστ αυτά ξεπερνούν κάποιους από τους περιορισμούς των προηγούμενων μεθόδων [99]. Πιο αναλυτικά, ανιχνεύουν συγκεκριμένες ουσίες στα ούρα, τα λεγόμενα εξωκυτταρικά κυστίδια, τα οποία συνιστούν μία νέα και ιδιαίτερα ελπιδοφόρα πηγή διαγνωστικών και προγνωστικών βιοδεικτών σε υγρές βιοψίες [100]. Ωστόσο, συχνά παρουσιάζουν χαμηλή ευαισθησία και ειδικότητα, κυρίως για τους χαμηλού βαθμού και πρώιμου σταδίου όγκους και για τις περιπτώσεις επανεμφάνισης, και επιστρέφουν πολλές φορές ψευδώς θετικά αποτελέσματα [101, 102]. Για το λόγο αυτό τα τεστ αυτά δεν έχουν αντικαταστήσει τα ισχύοντα διαγνωστικά εργαλεία τα οποία αποτελούν η κυστεοσκόπηση και η κυτταρολογική εξέταση των ούρων [103].

Οι αναδυόμενες κατηγορίες βιοδεικτών στα ούρα που χρησιμοποιούν τη μεθυλίωση του DNA και τα microRNAs (miRNAs) μελετώνται ενδελεχώς επί του παρόντος. Για παράδειγμα, μια πρόσφατη ανασκόπηση ανέφερε ότι οι δείκτες μεθυλίωσης στα ούρα αποκάλυψαν ευαισθησία της τάξης του 65-

100% και ειδικότητα της τάξης του 77-100% στο πλαίσιο της ανίχνευσης. Ομοίως, η ευαισθησία της αξιολόγησης των miRNAs στα ούρα που αποκολλήθηκαν από κύτταρα KOK ήταν 71-94%, με ειδικότητα στο εύρος 51-100%. Ένα άλλο μοντέλο που αξιολογεί την έκφραση των γονιδίων *IGFBP5*, *HOXA13*, *MDK*, *CDK1* και *CXCR2* σε ένα δείγμα ούρων (γονοτυπικά δεδομένα) συνδυαστικά με την ηλικία, το φύλο, τη συχνότητα της μακροσκοπικής αιματουρίας και το ιστορικό καπνίσματος (φαινοτυπικά δεδομένα) ήταν σε θέση να διαστρωματώσει σωστά το 80% των ασθενών με μικροσκοπική αιματουρία που δεν είχαν KOK, ώστε να αποφύγουν μία αχρείαστη πλήρη ουρολογική εξέταση. Η ενσωμάτωση βιοδεικτών στην ανίχνευση του KOK θα απαιτήσει εκτενή έλεγχο εγκυρότητας της κλινικής τους αξίας ώστε να μειωθεί ο κίνδυνος ψευδώς αρνητικών αποτελεσμάτων. Ωστόσο, ο εντοπισμός ασθενών με υψηλό κίνδυνο καρκίνου θα βελτιώσει σημαντικά την ανίχνευση εάν ο εν λόγω ασθενής δεν επρόκειτο να αξιολογηθεί εξ αρχής.

Ο ρόλος των στεροειδών φύλου (ορμόνες φύλου) και των ειδικών υποδοχέων τους είναι ένας αναδύμενος ερευνητικός τομέας στην εξέλιξη του KOK. Η ιδέα αυτή βασίζεται σε επιδημιολογικές μελέτες οι οποίες καταδεικνύουν ότι οι γυναίκες διαγιγνώσκονται με νόσο υψηλότερου σταδίου και έχουν χειρότερη έκβαση μετά τη θεραπεία συγκριτικά με τους άνδρες. Επιπλέον, η πρώιμη ηλικία κατά την εμμηνόπαυση (≤ 45 ετών έναντι ≥ 50 ετών) σχετίζεται με αυξημένο κίνδυνο εμφάνισης KOK, ενώ οι άγαμες γυναίκες, οι γυναίκες που ανέφεραν καθυστερημένη εμμηναρχή (≥ 15 ετών) και οι γυναίκες που έχουν χρησιμοποιήσει θεραπεία με οιστρογόνα και προγεστερόνη έχουν χαμηλότερο κίνδυνο εμφάνισης KOK από ό,τι οι γυναίκες που δεν έχουν λάβει θεραπεία για την εμμηνόπαυση. Δεδομένα από βασικές επιστημονικές μελέτες έχουν δείξει ότι ένας από τους πιο πιθανούς μηχανισμούς για την εξέλιξη του όγκου αφορά την γλυκουρονουλτρανσφεράση 1 (UDP glucuronosyltransferase 1 – UGT1A), ένα ένζυμο ζωτικής σημασίας για την αποτοξικοποίηση σημαντικών καρκινογόνων ουσιών, όπως οι αρωματικές αμίνες. Η UGT1A ρυθμίζεται διαφορετικά από τα οιστρογόνα και έχει προστατευτικό ρόλο στο φυσιολογικό ουροθήλιο, αλλά τα μειωμένα επίπεδά της σχετίζονται με υποτροπή και εξέλιξη του καρκίνου στο νεοπλασματικό ουροθήλιο. Η UGT1A έχει ταυτοποιηθεί ότι ρυθμίζεται από τον μεταγραφικό παράγοντα του υποδοχέα ανδρογόνων (τουλάχιστον στον ιστό του προστάτη) και ο ανδρογόνο-επαγόμενα σήματα προάγουν την καρκινογένεση της ουροδόχου κύστης μέσω της μείωσης της έκφρασης των UGTs. Το εύρημα αυτό υποστηρίχθηκε επίσης από μια κλινική μελέτη που αξιολογεί τη σχέση μεταξύ του KOK και της θεραπείας στέρησης ανδρογόνων. Από 162 ασθενείς που είχαν τόσο καρκίνο του προστάτη όσο και μη μυοδιηθητικό KOK, το 22% των ασθενών που έλαβαν θεραπεία στέρησης ανδρογόνων παρουσίασαν υποτροπή του KOK σε σύγκριση με το 50% στην ομάδα ελέγχου, μετά από μια διάμεση παρακολούθηση 62 μηνών και η πενταετής αναλογιστική επιβίωση χωρίς υποτροπή ήταν 76% έναντι 40%. Κατά συνέπεια, ο πιθανός ρόλος των στεροειδών του φύλου ή των ειδικών υποδοχέων τους μπορεί να ανοίξει το πεδίο για νέες στοχευμένες θεραπείες για τον KOK.

Είναι φανερό πως τις τελευταίες δεκαετίες, πολλές έρευνες έχουν εστιάσει στον εντοπισμό μοριακών δεικτών υψηλής αξίας για τον KOK. Δεν υπάρχει αμφιβολία ότι, παρόλο που οι ακριβείς μοριακοί μηχανισμοί που λαμβάνουν χώρα στην ανάπτυξη και εξέλιξη του KOK παραμένουν αδιευκρίνιστοι [104], έχουμε διευρύνει σημαντικά την κατανόησή μας όσον αφορά τη μοριακή παθοφυσιολογία του KOK, γεγονός που μας έχει επιτρέψει να καθιερώσουμε νέους προγνωστικούς και προβλεπτικούς βιοδείκτες χρησιμοποιώντας τις γονιδιακές τεχνολογίες υψηλής απόδοσης. Ωστόσο, δεν υπάρχει ακόμα κανένας εγκεκριμένος βιοδείκτης στην κλινική πρακτική και οι εξελίξεις στη διαχείριση και τη θεραπεία του KOK υστερούν συγκριτικά με άλλους τύπους καρκίνου [105]. Στο πλαίσιο αυτό, η ανάγκη για την ανάπτυξη αξιόπιστων και μη επεμβατικών μεθόδων οι οποίες θα μπορούν να ανιχνεύουν και να προβλέπουν την βιολογική συμπεριφορά του KOK είναι αδιαμφισβήτητη και συνεχίζεται εντατικά έως και σήμερα [106].

1.4 Γονιδιακή έκφραση

Η **γονιδιακή έκφραση** είναι η διαδικασία κατά την οποία οι πληροφορίες από ένα γονίδιο χρησιμοποιούνται στη σύνθεση ενός λειτουργικού γονιδιακού προϊόντος το οποίο επιτρέπει να παράγονται τελικά προϊόντα, πρωτεΐνες ή μη κωδικοποιημένα RNA, και, τελικά, επηρεάζει έναν φαινότυπο, ως τελικό αποτέλεσμα. Τα προϊόντα αυτά είναι στην πλειονότητα των περιπτώσεων πρωτεΐνες, αλλά για τα γονίδια που δεν κωδικοποιούν πρωτεΐνες, όπως το μεταφορικό RNA (tRNA) και το μικρό πυρηνικό RNA (snRNA), το προϊόν είναι ένα λειτουργικό μη κωδικοποιητικό RNA (ncRNA). Η γονιδιακή έκφραση συνοψίζεται στο **κεντρικό δόγμα της μοριακής βιολογίας** που διατυπώθηκε για πρώτη φορά από τον Francis Crick το 1957, αναπτύχθηκε περαιτέρω στο άρθρο του το 1970 και επεκτάθηκε με τις μεταγενέστερες ανακαλύψεις της αντίστροφης μεταγραφής και της αντιγραφής του RNA (Εικόνα 16) [107].

Το κεντρικό δόγμα της μοριακής βιολογίας είναι η εξήγηση της ροής της γενετικής πληροφορίας μέσα σε ένα βιολογικό σύστημα. Συχνά διατυπώνεται ως εξής: «Το DNA δημιουργεί το RNA και το RNA δημιουργεί την πρωτεΐνη», αν και αυτό δεν είναι το αρχικό του νόημα. Η αρχική διατύπωση του Francis Crick το 1957 και στη συνέχεια δημοσίευσή του το 1958, αναφέρουν τα εξής:

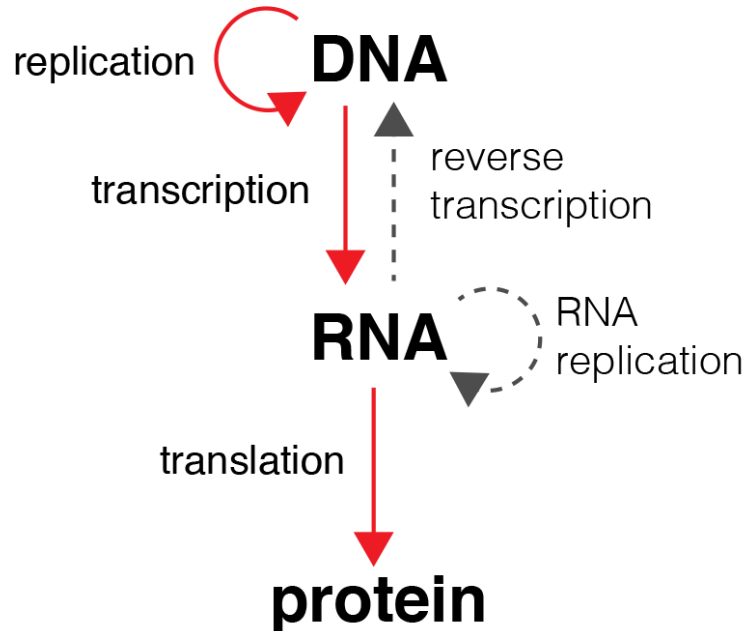
«Το κεντρικό δόγμα: Αυτό δηλώνει ότι από τη στιγμή που η πληροφορία έχει περάσει στην πρωτεΐνη δεν μπορεί να ξαναβγεί. Πιο αναλυτικά, η μεταφορά πληροφοριών από νουκλεϊκό οξύ σε νουκλεϊκό οξύ ή από νουκλεϊκό οξύ σε πρωτεΐνη μπορεί να είναι δυνατή, αλλά η μεταφορά από πρωτεΐνη σε πρωτεΐνη ή από πρωτεΐνη σε νουκλεϊκό οξύ είναι αδύνατη. Ο όρος πληροφορία εδώ σημαίνει τον ακριβή προσδιορισμό της αλληλουχίας, είτε των βάσεων στο νουκλεϊκό οξύ, είτε των αμινοξέων στην πρωτεΐνη».

Ο Crick το διατύπωσε εκ νέου σε ένα άρθρο του στο Nature που δημοσιεύθηκε το 1970, στο οποίο αναφέρει: «Το κεντρικό δόγμα της μοριακής βιολογίας ασχολείται με τη λεπτομερή μεταφορά της διαδοχικής πληροφορίας από νουκλεϊκό οξύ σε νουκλεϊκό οξύ. Δηλώνει ότι η πληροφορία αυτή δεν μπορεί να μεταφερθεί πίσω από την πρωτεΐνη στην πρωτεΐνη ή στο νουκλεϊκό οξύ».

Μια δεύτερη εκδοχή του κεντρικού δόγματος είναι δημοφιλής αλλά λανθασμένη. Πρόκειται για την απλουστευτική διαδρομή DNA → RNA → πρωτεΐνη που δημοσίευσε ο James Watson στην πρώτη έκδοση του βιβλίου «The Molecular Biology of the Gene» (1965). Η εκδοχή του Watson διαφέρει από εκείνη του Crick επειδή ο Watson περιγράφει μια διαδικασία δύο σταδίων (DNA → RNA και RNA → πρωτεΐνη) ως το κεντρικό δόγμα. Ενώ το δόγμα, όπως είχε αρχικά διατυπωθεί από τον Crick, εξακολουθεί να ισχύει σήμερα, η εκδοχή του Watson δεν ισχύει.

Σε γενικές γραμμές, το δόγμα είναι ένα πλαίσιο για την κατανόηση της μεταφοράς των πληροφοριών αλληλουχίας μεταξύ βιοπολυμερών που μεταφέρουν αυτές τις πληροφορίες στους ζωντανούς οργανισμούς. Υπάρχουν τρεις μεγάλες κατηγορίες τέτοιων βιοπολυμερών: το DNA, το RNA (που και τα δύο αποτελούν νουκλεϊκά οξέα) και οι πρωτεΐνες. Υπάρχουν $3 \times 3 = 9$ πιθανές άμεσες μεταφορές πληροφοριών που μπορούν να συμβούν μεταξύ αυτών. Το δόγμα τις ταξινομεί σε τρεις ομάδες των τριών: τρεις γενικές μεταφορές (πιστεύεται ότι συμβαίνουν κανονικά στα περισσότερα κύτταρα), δύο ειδικές μεταφορές (είναι γνωστό ότι συμβαίνουν, αλλά μόνο υπό συγκεκριμένες συνθήκες στην περίπτωση ορισμένων ιών ή σε ένα εργαστήριο) και τέσσερις άγνωστες μεταφορές (πιστεύεται ότι δεν συμβαίνουν ποτέ). Οι γενικές μεταφορές περιγράφουν τη φυσιολογική ροή της βιολογικής πληροφορίας: Το DNA μπορεί να αντιγραφεί σε DNA (αντιγραφή του DNA – **DNA replication**), οι πληροφορίες του DNA μπορούν να αντιγραφούν σε mRNA (μεταγραφή – **transcription**) και οι πρωτεΐνες μπορούν να συντεθούν χρησιμοποιώντας τις πληροφορίες του mRNA ως πρότυπο (μετάφραση – **translation**). Οι ειδικές μεταφορές περιγράφουν: το RNA μπορεί να αντιγραφεί από το RNA (αντιγραφή RNA – **RNA replication**), το DNA μπορεί να συντεθεί χρησιμοποιώντας ένα πρότυπο RNA (αντίστροφη μεταγραφή – **reverse transcription**). Οι άγνωστες μεταφορές περιγράφουν:

πρωτεΐνες που συντίθενται απευθείας από ένα πρότυπο DNA χωρίς τη χρήση mRNA, μια πρωτεΐνη που αντιγράφεται από μια πρωτεΐνη, σύνθεση RNA χρησιμοποιώντας την πρωταρχική δομή μιας πρωτεΐνης ως πρότυπο, καθώς και σύνθεση DNA χρησιμοποιώντας την πρωταρχική δομή μιας πρωτεΐνης ως πρότυπο, οι οποίες δεν πιστεύεται ότι συμβαίνουν στη φύση.



Εικόνα 16: Το κεντρικό δόγμα της μοριακής βιολογίας.

(Πηγή: <https://lms.su.edu.pk/lesson/184/week-1-introduction-to-molecular-biology>)

Η διαδικασία της γονιδιακής έκφρασης συναντάται σε **όλες τις γνωστές μορφές ζωής**, όπως τους ευκαρυωτικούς οργανισμούς (συμπεριλαμβανομένων των πολυκύτταρων οργανισμών), τους προκαρυωτικούς οργανισμούς (βακτήρια και αρχαία) και τους ιούς, για τη δημιουργία του μακρομοριακού μηχανισμού για τη ζωή.

Στη γενετική, η γονιδιακή έκφραση είναι το πιο θεμελιώδες επίπεδο στο οποίο ο γονότυπος δημιουργεί τον **φαινότυπο**, δηλαδή το παρατηρήσιμο χαρακτηριστικό. Η γενετική πληροφορία που είναι αποθηκευμένη στο DNA αντιπροσωπεύει τον **γονότυπο**, ενώ ο φαινότυπος προκύπτει από την «ερμηνεία» αυτής της πληροφορίας. Τέτοιοι φαινότυποι δύνανται να εμφανίζονται συχνά με τη σύνθεση πρωτεϊνών που ελέγχουν τη δομή και την ανάπτυξη του οργανισμού ή που δρουν ως ένζυμα που καταλύουν συγκεκριμένες μεταβολικές οδούς.

Όλα τα στάδια της διαδικασίας της γονιδιακής έκφρασης μπορούν να ρυθμιστούν. Αυτά συμπεριλαμβάνουν τη διαδικασία της μεταγραφής, της ωρίμανσης του RNA, της μετάφρασης και της μετα-μεταφραστικής τροποποίησης μιας πρωτεΐνης. Η ρύθμιση της γονιδιακής έκφρασης μπορεί να ελέγξει τον χρόνο, τη θέση και την ποσότητα ενός συγκεκριμένου γονιδιακού προϊόντος (πρωτεΐνη ή ncRNA) που υπάρχει σε ένα κύτταρο και μπορεί να έχει βαθιά επίδραση στην κυτταρική δομή και τη λειτουργία. Η ρύθμιση της γονιδιακής έκφρασης αποτελεί τη βάση για την κυτταρική διαφοροποίηση, την ανάπτυξη, τη μορφογένεση, την ευελιξία και την προσαρμοστικότητα του κάθε οργανισμού. Η γονιδιακή ρύθμιση μπορεί επομένως να χρησιμεύσει ως υπόστρωμα για τις εξελικτικές αλλαγές.

Ο πιο διαισθητικός τρόπος για τη διερεύνηση ενός συγκεκριμένου φαινοτύπου είναι η μέτρηση των επιπέδων έκφρασης των λειτουργικών πρωτεϊνών που υπάρχουν σε μια δεδομένη στιγμή στο κύτταρο. Ωστόσο, η μέτρηση της συγκέντρωσης των πρωτεϊνών μπορεί να είναι δύσκολη, λόγω των

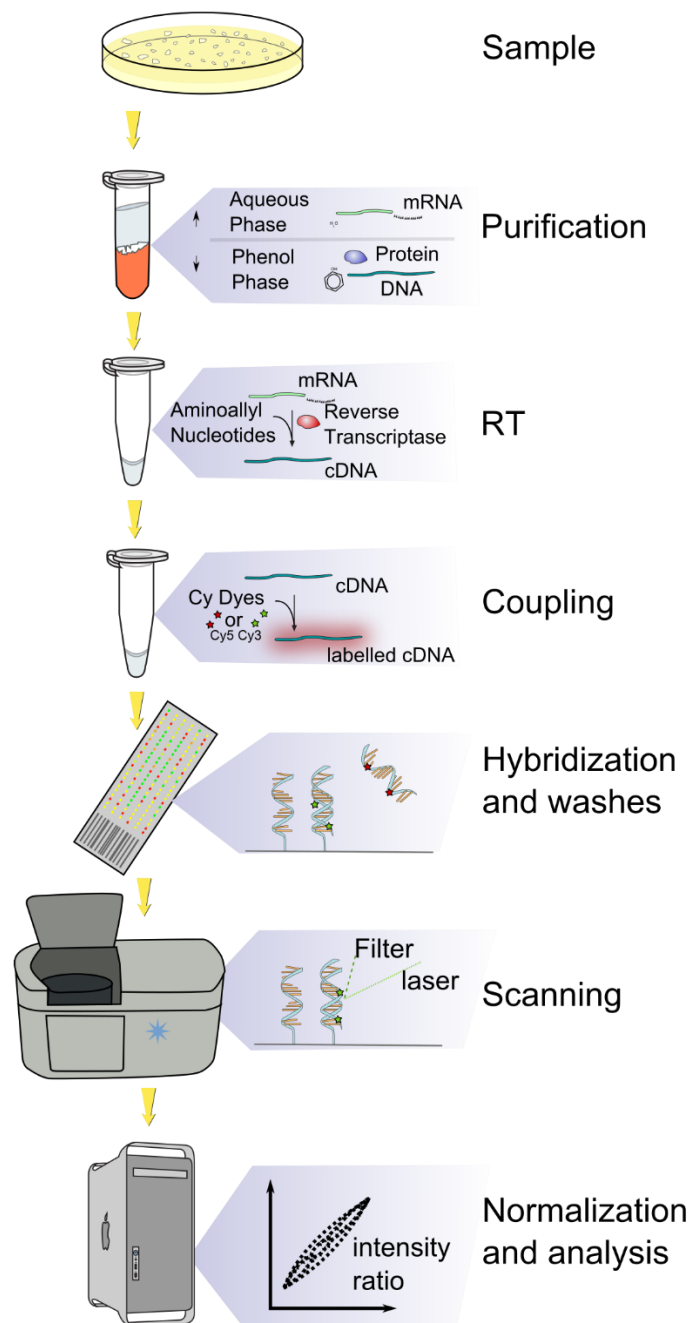
διαφορετικών θέσεων, τροποποιήσεων και πλαισίων στα οποία βρίσκονται, καθώς και λόγω της μη πληρότητας του πρωτεώματος. Τα επίπεδα έκφρασης mRNA, ωστόσο, είναι ευκολότερο να μετρηθούν και συχνά αποτελούν μια καλή προσέγγιση για την ποσοτικοποίηση της γονιδιακής έκφρασης. Με τη μέτρηση του mRNA, αναλύουμε τη ρύθμιση στο **επίπεδο της μεταγραφής**, χωρίς τις πρόσθετες επιπλοκές της μεταφραστικής ρύθμισης και της ενεργού αποικοδόμησης των πρωτεϊνών, γεγονός που απλοποιεί την ανάλυση με κόστος φυσικά την απώλεια κάποιων πληροφοριών. Στην επόμενη υποενότητα, θα εξετάσουμε την πρώτη και ευρέως διαδομένη τεχνική για τη δημιουργία δεδομένων γονιδιακής έκφρασης, τις μικροσυστοιχίες DNA, στην οποία βασίζονται και τα πρωτογενή δεδομένα της παρούσας έρευνας. Άλλες τεχνικές μέτρησης της γονιδιακής έκφρασης αποτελούν η αποτύπωση Northern (Northern blotting), η αλυσιδωτή αντίδραση πολυμεράσης σε πραγματικό χρόνο (Real-Time PCR) και, πιο πρόσφατα, η ανάλυση μεταγραφώματος (RNA-Seq).

1.5 Μικροσυστοιχίες

Ο όρος «συστοιχία» (array) σημαίνει μια συγκεκριμένη, διατεταγμένη απεικόνιση ενός ορισμένου αντικειμένου και χρησιμοποιείται για να περιγράψει μια αξιοσημείωτη πρόσφατη τεχνολογία, όπου αλληλουχίες DNA στερεώνονται σε αντικειμενοφόρες πλάκες μικροσκοπίου ή τσιπ, κατά συστοιχίες, σχηματίζοντας χημικούς δεσμούς. Ο σκοπός αυτών των «συστοιχιών DNA» είναι να προσδιοριστεί αν σε ένα συγκεκριμένο βιολογικό δείγμα υπάρχουν ή όχι νουκλεϊκά οξέα που έχουν επισημανθεί με ουσίες φθορισμού, καθώς επίσης και να ποσοτικοποιηθεί η αφθονία τους στο δείγμα αυτό. Αυτά τα νουκλεϊκά οξέα υβριδοποιούνται με τα μόρια DNA στις αντικειμενοφόρες πλάκες των συστοιχιών βάσει της **αρχής της συμπληρωματικότητας** των J. Watson και F. Crick, και, καθώς είναι επισημασμένα με ειδικές ουσίες, ανιχνεύονται εύκολα. Ως αποτέλεσμα, χιλιάδες μόρια DNA που είναι συνδεδεμένα στην αντικειμενοφόρο μικροσυστοιχία (microarray) υβριδοποιούνται και άρα ελέγχονται ταυτόχρονα. Η βιοπληροφορική, δηλαδή η χρήση της επιστήμης των υπολογιστών στη διαχείριση των βιολογικών δεδομένων, παίζει καθοριστικό ρόλο σε αυτή την τεχνολογία, καθώς παρέχει τις διαδικασίες για τη μετέπειτα ανάλυση των δεδομένων που προέρχονται από την υβριδοποίηση των μορίων DNA από το δείγμα – στόχο. Στην πραγματικότητα, δεν έχει νόημα να διεξαχθούν πειράματα μικροσυστοιχιών DNA χωρίς τη χρήση της βιοπληροφορικής, καθώς τα αποτελέσματα αυτά δε θα επιδέχονται κάποιας ερμηνείας.

Η πρώιμη ιστορία των συστοιχιών DNA ξεκίνησε με την ιδέα της μεθόδου του υβριδισμού αποικιών που προτάθηκε από τους Hogness και Grunstein το 1975. Αυτοί οι επιστήμονες κλωνοποίησαν τυχαία DNA σε πλασμίδια *E. coli* και χρησιμοποίησαν φίλτρο νιτροκυτταρίνης για να καλύψουν τα τρυβλία Πέτρι (Petri dish) από άγαρ, τα οποία επιστρώθηκαν με μετασχηματισμένα κύτταρα. Το DNA των αποικιών μετουσιώθηκε και σταθεροποιήθηκε στα φίλτρα, μετά από λύση, δημιουργώντας συλλογή κλωνοποιημένων τμημάτων DNA. Κατά συνέπεια, αυτά υβριδοποιήθηκαν με ραδιοσημασμένο ανιχνευτή ενδιαφέροντος που τους επέτρεψε να ελέγξουν χιλιάδες των αποικιών οι οποίες περιείχαν το κλωνοποιημένο DNA. Με βάση αυτή την προσέγγιση, το 1979, ένας άλλος επιστήμονας, ο J. Gergen, κατάφερε να κατασκευάσει μια μηχανική συσκευή που αναπαρήγαγε πολυάριθμες πλάκες μικροτιτλοδότησης σε άγαρ δημιουργώντας τελικά συστοιχίες αποικιών σε μια περιοχή μονάχα 26 × 38 cm, το 1728. Το πρωτόκολλό του περιλάμβανε επίσης τη μεταφορά αποικιών σε διηθητικό χαρτί και στη συνέχεια την ανάλυση των κυττάρων, τη μετουσίωση και τη στερέωση του DNA στο φίλτρο, γεγονός που επέτρεπε τη δημιουργία συστοιχιών DNA σε φίλτρα που επαναχρησιμοποιούνταν πολλές φορές. Αυτό ήταν ένα μεγάλο πλεονέκτημα, καθώς εξοικονομούσε πολύ χρόνο και προσπάθεια. Αργότερα, τα πρωτόκολλα που ανέπτυξε ο Gergen αυτοματοποιήθηκαν με τη χρήση ρομποτικών συστημάτων από τον Patrick Brown και την ερευνητική του ομάδα το 1995. Κατάφεραν έτσι να μετρήσουν τη γονιδιακή έκφραση 48 γονιδίων του φυτού *Arabidopsis thaliana* χρησιμοποιώντας

συμπληρωματικό DNA (complementary DNA – cDNA) που προήλθε από προϊόντα PCR και εκτυπώνοντας τη συστοιχία με τη χρήση ενός κατασκευαστικού ρομπότ. Μέχρι την επόμενη δεκαετία, η ανίχνευση φθορισμού έγινε μέρος της τεχνολογίας των συστοιχιών DNA στην οποία έδωσε ένα σημαντικό πλεονέκτημα, καθώς η απεικόνιση του υβριδοποιημένου DNA ήταν εύκολη και διαβάζονταν γρήγορα από το σύστημα. Επιπλέον, πραγματοποιήθηκε περαιτέρω έρευνα σε αυτή την αξιολογή τεχνολογία και σταδιακά βελτιώθηκαν τα πρωτόκολλά της. Ωστόσο, η χρήση πολύ μεγάλων αλληλουχιών δεν απέφερε αρχικά μεγάλη επιτυχία. Έτσι, η χρήση μικρότερων ολιγονουκλεοτιδίων, των 25 – 60 βάσεων, οδήγησε σε καλύτερη ειδικότητα, καθώς αυτά τα ολιγονουκλεοτίδια σχεδιάστηκαν έτσι ώστε να προσδένονται σε θέσεις – στόχους στο γονίδιο που διέφεραν σημαντικά από άλλες θέσεις ή γονίδια.



Εικόνα 17: Τα κύρια στάδια ενός πειράματος με τη χρήση μικροσυστοιχιών.
(Πηγή: https://en.wikipedia.org/wiki/Microarray_analysis_techniques)

Όπως αναφέρθηκε προηγουμένως, η τεχνική αυτή βασίζεται στην αντιστοίχιση άγνωστων και γνωστών δειγμάτων DNA, μέσω της αρχής της συμπληρωματικότητας των βάσεων των Watson και Crick. Τα γνωστά δείγματα DNA, που ονομάζονται «ανιχνευτές» (probes), εναποτίθενται και στερεώνονται σε μικροσκοπικά τζαμάκια ή τσιπ πυριτίου κατά χιλιάδες. Αυτά τα δείγματα μπορεί να είναι ολιγονουκλεοτίδια (oligonucleotides), συμπληρωματικό DNA (cDNA) ή ακόμη και απλώς DNA. Από την άλλη πλευρά, τα άγνωστα δείγματα DNA είναι αυτά που τίθενται προς ανάλυση, ως προς για παράδειγμα το επίπεδο γονιδιακής έκφρασής τους, και επισημαίνονται με τη χρήση μορίων – μαρτύρων, όπως τα φθοροφόρα, τα οποία αντικατέστησαν τα ραδιενεργά μόρια λόγω των πιθανών κινδύνων για την υγεία των αναλυτών. Εάν βρεθούν οι αλληλουχίες των ανιχνευτών στα γονίδια των αγνώστων δειγμάτων, θα σχηματίσουν συμπληρωματικούς δεσμούς με τους ανιχνευτές και θα εκπέμψουν, κατά την ακτινοβόλησή τους, σήμα φθορισμού το οποίο μπορεί να διαβαστεί από εξειδικευμένες κάμερες και υπολογιστικά συστήματα. Στη συνέχεια, αναλύονται με περισσότερη λεπτομέρεια τα κύρια βήματα για τη μέτρηση των επιπέδων της γονιδιακής έκφρασης σε ένα βιολογικό δείγμα με τη χρήση της τεχνολογίας μικροσυστοιχιών DNA. Τα βήματα αυτά περιλαμβάνουν την προετοιμασία και την επισήμανση του δείγματος (coupling), τον υβριδισμό (hybridization), το πλύσιμο (washing), την απόκτηση εικόνας (scanning) και την κανονικοποίηση (normalization) (Εικόνα 17).

Προετοιμασία δείγματος και σήμανση

Αρχικά, το mRNA πρέπει να εξαχθεί από το βιολογικό δείγμα ενδιαφέροντος και να καθαριστεί (purification). Στο πείραμα πρέπει επίσης να συμπεριληφθεί ένας παθολογικός έλεγχος (π.χ. να αποσαφηνιστεί εάν πρόκειται για ασθενή ή υγιή ιστό). Στη συνέχεια, η σήμανση (tagging) περιλαμβάνει την εκτέλεση αντίδρασης αντίστροφης μεταγραφής (reverse transcription polymerase chain reaction – RT-PCR) για τη σύνθεση της συμπληρωματικής αλυσίδας DNA (cDNA). Σε αυτή τη μέθοδο, ο εκκινητής poly T προσδένεται στο mRNA για να ξεκινήσει η διαδικασία της αντίστροφης μεταγραφής από το σήμα πολυαδενυλίωσης στην 3' αμετάφραστη περιοχή (untranslated region – UTR) του mRNA. Ένα μέρος των νουκλεϊκών οξέων, 5' - τριφωσφορικών δεοξυνουκλεοζιτών (dATP, dGTP, dCTP και dTTP), που προστίθενται σε αυτή τη φάση ενσωματώνονται με μια φωσφορίζουσα χρωστική (π.χ. μόνο dCTP επισημασμένο με Cy) μέσω ομοιοπολικού δεσμού (coupling). Εάν πρόκειται για μικροσυστοιχία δύο καναλιών, τα ασθενή και τα υγιή δείγματα μπορούν να επισημανθούν με διαφορετικές χρωστικές ουσίες, όπως π.χ. η Cy3 (διεγείρεται από πράσινο λέιζερ) και η Cy5 (διεγείρεται από κόκκινο λέιζερ), για τη διάκρισή τους και εναποτίθενται στην ίδια μικροσυστοιχία. Έτσι, οι ανιχνευτές cDNA που είναι συμπληρωματικοί ως προς τα σεσημασμένα μετάγραφα θα υβριδοποιηθούν (hybridization) και τελικά θα οπτικοποιηθούν ως έγχρωμες κηλίδες στην τελική εικόνα (Εικόνα 18). Ένας λόγος για τον οποίο η αντίστροφη μεταγραφή ξεκινάει στο σημείο 3' UTR είναι διότι θεωρείται ως η πιο μεταβλητή περιοχή στα γονίδια, γεγονός που δίνει το πλεονέκτημα της καλύτερης ειδικότητας κατά το σχεδιασμό των ανιχνευτών.

Επιπλέον, ένας άλλος τρόπος με τον οποίο το cDNA μπορεί να επισημανθεί με τη χρήση της αντίστροφης μεταγραφάσης είναι η χρήση του θραύσματος Klenow της DNA πολυμεράσης I και της τυχαίας εκκίνησης. Σε αυτή την αντίδραση, το cDNA που έχει προετοιμαστεί με τυχαίους εκκινητές επεκτείνεται χρησιμοποιώντας θραύσμα Klenow μαζί με την παρουσία επισημασμένων τριφωσφορικών δεοξυκυτιδινών (dCTPs). Το προϊόν της αντίδρασης θα περιλαμβάνει έτσι επισημασμένα μετάγραφα που είναι συμπληρωματικά και στις δύο αλυσίδες του γονιδίου. Αυτό βοηθά τους ερευνητές να ελέγχουν για διασταυρούμενο υβριδισμό στις συστοιχίες.

Υβριδισμός

Σε αυτό το βήμα, οι ανιχνευτές DNA που βρίσκονται στην αντικειμενοφόρο μικρο-πλάκα και τα cDNA – στόχοι με την ετικέτα στο υπό εξέταση δείγμα θα «ζευγαρώσουν» σύμφωνα με την αρχή της συμπληρωματικότητας των βάσεων των Watson και Crick. Η διαδικασία αυτή μπορεί να επιτευχθεί είτε χειροκίνητα είτε με τη χρήση κάποιου ρομποτικού συστήματος. Στην πρώτη προσέγγιση, η συστοιχία τοποθετείται σε ειδικό θάλαμο, όπου ο ερευνητής εγχύνει το διάλυμα που περιέχει τα cDNA – στόχους πάνω στη συστοιχία, υπό αποστειρωμένες συνθήκες, και την επωάζει σε συγκεκριμένη θερμοκρασία για 12 έως 24 ώρες. Κατά τη δεύτερη προσέγγιση, όλα εκτελούνται από ένα προγραμματισμένο ρομπότ το οποίο εξοικονομεί σαφώς χρόνο και προσπάθεια, εκτελεί το πρωτόκολλο σε εξειδικευμένο σταθμό και παρέχει καλύτερο έλεγχο της θερμοκρασίας που συνήθως κυμαίνεται μεταξύ 45 και 65°C. Είναι επίσης σημαντικό να αναφερθεί ότι ο υβριδισμός επηρεάζεται από πολλές περιβαλλοντικές συνθήκες, όπως η συγκέντρωση αλάτων, η θερμοκρασία, η συγκέντρωση φορμαμιδίου, η υγρασία και η ποσότητα του διαλύματος στόχου. Για παράδειγμα, η υψηλότερη θερμοκρασία και η χαμηλότερη συγκέντρωση αλάτων θα αυξήσουν την αυστηρότητα (stringency), πράγμα που σημαίνει ότι μόνο συγκεκριμένες αλυσίδες θα υβριδοποιηθούν. Προκειμένου να περιοριστεί ή να προληφθεί ο διασταυρούμενος υβριδισμός, μπορεί να προστεθεί μια επαναλαμβανόμενη αλληλουχία DNA και poly T ή poly A για να καλυφθεί η γονιδιωματική επαναλαμβανόμενη αλληλουχία και οι θέσεις πολυαδενυλίωσης στο cDNA, αντίστοιχα.

Πλύσιμο

Η εξάλειψη της περίσσειας του διαλύματος υβριδισμού από τη μικροσυστοιχία είναι ένας από τους λόγους για τους οποίους αυτό το βήμα είναι ζωτικής σημασίας, καθώς διασφαλίζει ότι μόνο τα cDNA – στόχοι με ετικέτα που είναι επιθυμητό να μετρηθούν θα δεσμευτούν στη μικροσυστοιχία. Επιπλέον, η πλύση αυξάνει την αυστηρότητα περιορίζοντας τη διασταυρούμενη υβριδοποίηση. Οι ερευνητές μπορούν να χρησιμοποιήσουν διαλύματα χαμηλής συγκέντρωσης αλάτων που περιέχουν λαουρυλοθειικό νάτριο (Sodium Dodecyl Sulfate – SDS) και αλατούχο διάλυμα κιτρικού νατρίου (Standard Saline Citrate – SSC). Πολλοί αυτοματοποιημένοι σταθμοί υβριδισμού μπορούν να περιλαμβάνουν έναν κύκλο πλύσης ως μέρος της όλης διαδικασίας.

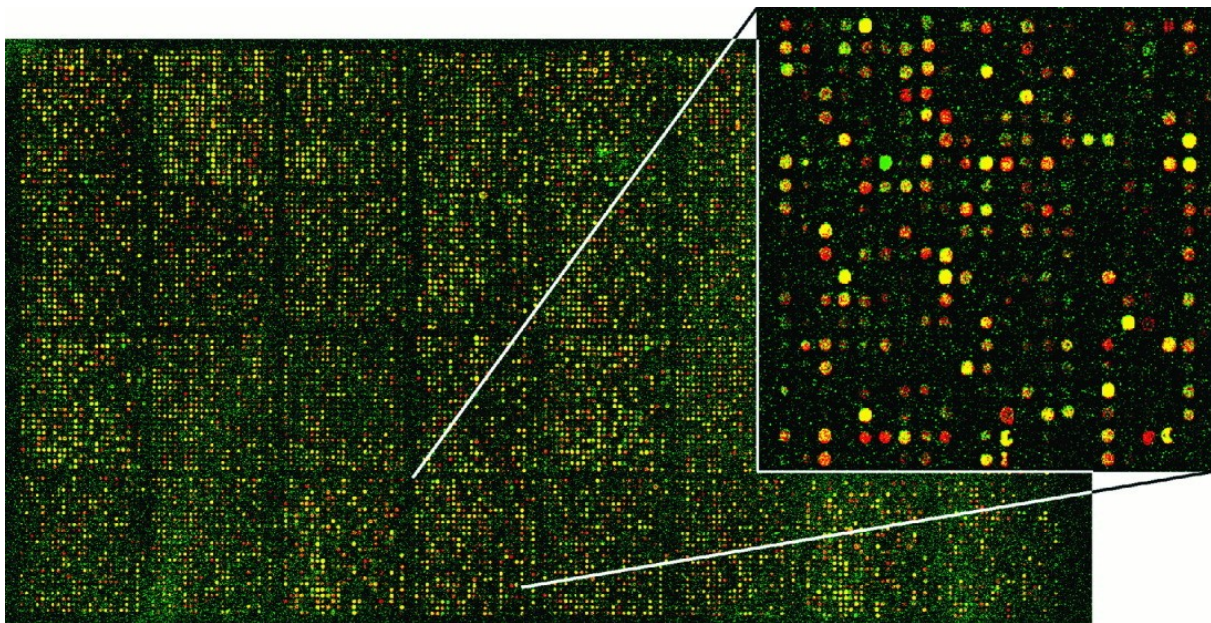
Λήψη εικόνας

Αυτό το στάδιο μπορεί να θεωρηθεί ως το τελικό βήμα της πειραματικής διαδικασίας, κατά το οποίο λαμβάνεται μια εικόνα των αποτελεσμάτων (Εικόνα 18). Επειδή το cDNA που συνδέεται με τον στόχο είναι σεσημασμένο με φθορίζουσες χρωστικές, αυτές μπορούν να διεγερθούν από φως κατάλληλου μήκους κύματος δίνοντας συγκεκριμένα χρώματα. Έτσι, οι μικροσυστοιχίες τοποθετούνται κάτω από το σαρωτή που διαθέτει δύο λέιζερ (π.χ. για τη διέγερση δύο διαφορετικών χρωστικών για ασθενείς και υγιείς ιστούς) για να ακτινοβοληθούν, να φθορίσουν και να διαβαστούν. Για μεγαλύτερη ακρίβεια, ο σαρωτής «διατρέχει» ολόκληρη την αντικειμενοφόρο πλάκα για να διαβαστεί κάθε σημείο της μικροσυστοιχίας. Επίσης, η ρύθμιση του μεγέθους του εικονοστοιχείου (αντιπροσωπεύει το μέγεθος του φυσικού χώρου) είναι το ίδιο με το πάχος της δέσμης του λέιζερ για να διασφαλιστεί ότι το φως που διαβάζεται δεν προέρχεται από γειτονικά σημεία της μικροσυστοιχίας αλλά από έναν συγκεκριμένο ανιχνευτή.

Κανονικοποίηση

Λαμβάνοντας υπόψη ότι μπορεί να έχουν προκύψει κάποια σφάλματα από τη σάρωση και απόκτηση των εικόνων, θα μπορούσε κανείς να εφαρμόσει μια διαδικασία διόρθωσης των «μεροληψιών» στις

εικόνες των πειραμάτων μικροσυστοιχιών πριν από την τελική ανάλυση. Αυτή η διαδικασία είναι γνωστή ως κανονικοποίηση (normalization) και χρησιμεύει ως «βαθμονόμηση» για την εξάλειψη των συστηματικών διακυμάνσεων μεταξύ των δειγμάτων. Ορισμένες από τις διάφορες «ασυνέπειες» και σφάλματα που μπορεί να προκύψουν περιλαμβάνουν τις διαφορετικές ρυθμίσεις του σαρωτή, τις ιδιότητες του υβριδισμού, την αποδοτικότητα των χρωστικών ουσιών, καθώς και τις περιβαλλοντικές συνθήκες. Αυτές έχουν μεγάλη επίδραση στα πειραματικά αποτελέσματα και συνεπώς μπορεί να οδηγήσουν σε παραπλανητικά συμπεράσματα σχετικά με την ανάλυση του υβριδισμού του DNA. Αν και υπάρχουν διάφορες μέθοδοι κανονικοποίησης, συνήθως χρησιμοποιείται η μέθοδος συνταιριασμού καμπύλης (LOcally WEighted Scatterplot Smoothing – LOWESS), η οποία ανιχνεύει συστηματικές διαφοροποιήσεις με χρήση γραμμικής παλινδρόμησης, ως συνάρτηση του $\log_{10}$ (στην περίπτωση της έντασης), και εξισορροπεί την παρατηρούμενη αναλογία με την καλύτερη προσαρμογή του μέσου $\log_2$ (στην περίπτωση του λόγου των εντάσεων). Μετά την κανονικοποίηση, ο λόγος της έκφρασης, ο οποίος είναι η κανονικοποιημένη τιμή του εκφραζόμενου γονιδίου σε σχέση με εκείνη των δειγμάτων ελέγχου, μπορεί να υπολογιστεί χρησιμοποιώντας τον ακόλουθο τύπο: $T_i = R_i/G_i$, όπου το i αντιπροσωπεύει το γονίδιο και τα R (Red) και G (Green) αντιπροσωπεύουν το δείγμα – στόχο και το δείγμα – έλεγχο, αντίστοιχα. Εάν αντί αυτού χρησιμοποιηθεί ο λόγος $T_i = \log_2(R_i/G_i)$, αυτό θα προεκτείνει το δυναμικό εύρος των σημάτων του επιπέδου έκφρασης των γονιδίων.



Εικόνα 18: Ενδεικτική εικόνα μικροσυστοιχίας γονιδίων. [108]

1.6 Επεξεργασία και Ανάλυση δεδομένων μικροσυστοιχιών

Οι μικροσυστοιχίες μπορούν να χρησιμοποιηθούν σε πολλούς τύπους πειραμάτων, όπως ο γονότυπος, η επιγενετική, η ανάλυση του μεταφραστικού προφίλ και η ανάλυση του προφίλ γονιδιακής έκφρασης. Η σκιαγράφηση του **προφίλ της γονιδιακής έκφρασης (gene expression profiling)** είναι μακράν η πιο κοινή χρήση της τεχνολογίας μικροσυστοιχιών. Για αυτό το είδος πειραμάτων μπορούν να χρησιμοποιηθούν τόσο μονοχρωματικές όσο και δίχρωμες μικροσυστοιχίες. Η διαδικασία **ανάλυσης των δεδομένων γονιδιακής έκφρασης** είναι παρόμοια και για τους δύο τύπους μικροσυστοιχιών και περιλαμβάνει (Εικόνα 24):

- εξαγωγή χαρακτηριστικών,
- ποιοτικό έλεγχο,

- κανονικοποίηση,
- ανάλυση διαφορικής έκφρασης,
- βιολογική ερμηνεία των αποτελεσμάτων, και
- υποβολή δεδομένων σε δημόσια βάση δεδομένων.

Αναλυτικότερα, η **εξαγωγή χαρακτηριστικών** είναι η διαδικασία μετατροπής της σαρωμένης εικόνας της μικροσυστοιχίας σε μετρήσιμες τιμές (ποσοτικοποίηση) και η αντιστοίχιση των ανιχνευτών με τα αναγνωριστικά των γονιδίων, η συμπλήρωση των ονομάτων των δειγμάτων και άλλες χρήσιμες πληροφορίες, όπως τα φαινοτυπικά χαρακτηριστικά τους και τα κλινικά χαρακτηριστικά των ασθενών από τους οποίους προέρχονται. Η διαδικασία αυτή εκτελείται συνήθως με τη χρήση του λογισμικού που παρέχεται από τον κατασκευαστή της μικροσυστοιχίας. Η έξοδος αυτής της διαδικασίας είναι ακατέργαστα (δηλαδή μη επεξεργασμένα) αρχεία δεδομένων που μπορεί να είναι σε δυαδική μορφή ή σε μορφή αριθμών και κειμένου (Πίνακας 1).

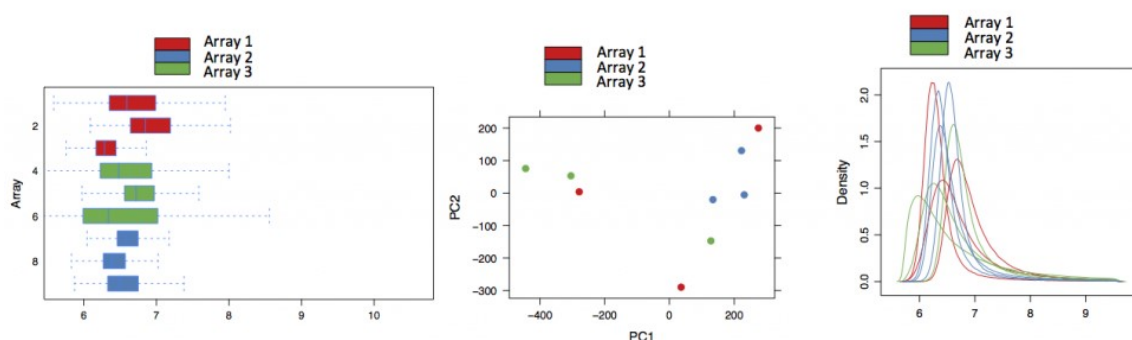
Πίνακας 1: Κοινοί τύποι αρχείων πρωτογενών δεδομένων μικροσυστοιχιών.

Κατασκευαστής	Τυπικό μορφή πρωτογενούς αρχείου	Λογισμικό / πακέτο ανάλυσης δεδομένων
Affymetrix	.cel αρχείο (δυαδικό αρχείο)	πακέτα της R (affy, limma, oligo...)
Agilent	αρχείο εξαγωγής χαρακτηριστικών (αρχείο κειμένου με διαχωρισμένες καρτέλες ανά υβριδισμό)	πακέτα της R (π.χ. limma) Λογισμικό λογιστικών φύλλων (Excel, OpenOffice, κ.τ.λ.)
Illumina	.idat (δυαδικό αρχείο)	πακέτα της R (π.χ. illuminaio)
	.txt αρχείο (πίνακας κειμένου με διαχωρισμένες καρτέλες για όλα τα δείγματα)	πακέτα της R (π.χ. lumi) Λογισμικό λογιστικών φύλλων (Excel, OpenOffice, κ.τ.λ.)

Μετά τη διαδικασία εξαγωγής των χαρακτηριστικών, τα δεδομένα μπορούν να αναλυθούν. Οι κατασκευαστές μικροσυστοιχιών συχνά παρέχουν λογισμικό για το άνοιγμα και την ανάλυση των πρωτογενών αρχείων δεδομένων τους. Τα προγράμματα αυτά μπορεί να μην είναι πάντα διαθέσιμα, να έχουν ξεπεραστεί μετά από μερικά χρόνια ή να μην είναι αρκετά ευέλικτα για τις ανάγκες των επιστημόνων και ερευνητών. Υπάρχουν διάφορα δωρεάν εργαλεία λογισμικού που είναι κατάλληλα για την περαιτέρω επεξεργασία αρχείων με δεδομένα μικροσυστοιχιών. Μερικά παραδείγματα αποτελούν η πλατφόρμα Galaxy, το GenePattern, το GeneSpring (απαιτείται άδεια χρήσης) και η γλώσσα προγραμματισμού R. Ιδιαίτερα ευρέως χρησιμοποιούμενα πακέτα της R αποτελούν τα πακέτα *oligo*, *limma* και *lumi* για την ανάλυση δεδομένων μικροσυστοιχιών των εταιρειών Affymetrix, Agilent και Illumina, αντίστοιχα.

Στη συνέχεια, ο **ποιοτικός έλεγχος** των δεδομένων λαμβάνει χώρα με σκοπό να προσδιοριστεί εάν η όλη διαδικασία έχει λειτουργήσει αρκετά καλά ώστε τα δεδομένα να μπορούν να θεωρηθούν αξιόπιστα. Δεν υπάρχουν τυποποιημένες μέθοδοι για τον έλεγχο ποιότητας των δεδομένων μικροσυστοιχιών, αν και υπάρχουν έργα όπως το MicroArray Gene Expression (MAGE) που προσπαθούν να αναπτύξουν πρότυπα ποιότητας. Μία από τις σημαντικότερες μελέτες σε αυτόν τον τομέα αποτελεί η μελέτη The MicroArray Quality Control (MAQC) η οποία που βασίστηκε σε εκατοντάδες μικροσυστοιχίες και

αφορά στην ανασκόπηση αυτού του προβλήματος. Οι περισσότεροι ποιοτικοί έλεγχοι βασίζονται σε εικόνες και διαγράμματα, αν και, στην περίπτωση της των δεδομένων από μικροσυστοιχίες της εταιρείας Affymetrix, έχουν επίσης αναπτυχθεί αριθμητικές αναφορές των οποίων η χρήση είναι πολύ εκτεταμένη. Ένα ευρέως χρησιμοποιούμενο πακέτο της R που μπορεί να εφαρμοστεί σε όλες τις τεχνολογίες μικροσυστοιχιών (Affymetrix, Agilent και Illumina) αποτελεί το *arrayQualityMetrics*. Ο ποιοτικός έλεγχος των δεδομένων μικροσυστοιχιών αρχίζει με την οπτική επιθεώρηση των σαρωμένων εικόνων των μικροσυστοιχιών που αναλύονται για να εξακριβωθεί ότι δεν υπάρχουν εμφανείς κηλίδες, γρατσουνιές ή κενές περιοχές. Μετά την εξαγωγή των χαρακτηριστικών, τα πακέτα λογισμικού ανάλυσης δεδομένων μπορούν να χρησιμοποιηθούν για τη δημιουργία διαγνωστικών διαγραμμάτων (για παράδειγμα του σήματος υποβάθρου, των μέσων τιμών έντασης και του ποσοστού των γονιδίων πάνω από το υπόβαθρο) για να βοηθήσουν στον εντοπισμό προβληματικών μικροσυστοιχιών ή/και δειγμάτων (Εικόνα 19).



Εικόνα 19: Παραδείγματα γραφικών παραστάσεων ελέγχου ποιότητας κατά την ανάλυση δεδομένων διαφορικής έκφρασης. Από αριστερά προς τα δεξιά: Κατανομές έντασης συστοιχιών, διάγραμμα PCA, εκτιμήσεις πυκνότητας.

(Πηγή: <https://www.ebi.ac.uk/training/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/microarrays/analysis-of-microarray-data/quality-control/>)

Η **κανονικοποίηση** των δεδομένων μικροσυστοιχιών χρησιμοποιείται για τον έλεγχο της τεχνικής διακύμανσης μεταξύ των δοκιμών/πειραμάτων, διατηρώντας παράλληλα τη βιολογική διακύμανση η οποία είναι επιθυμητό να μελετηθεί. Σε γενικές γραμμές, οι μέθοδοι κανονικοποίησης βασίζονται στην ακόλουθη γενική αρχή: «τα περισσότερα γονίδια στη μικροσυστοιχία είτε δεν εκφράζονται είτε εκφράζονται εξίσου σε οποιαδήποτε κατάσταση. Μόνο μια μικρή ποσότητα γονιδίων παρουσιάζει αλλαγές στην έκφραση μεταξύ των μελετούμενων συνθηκών».

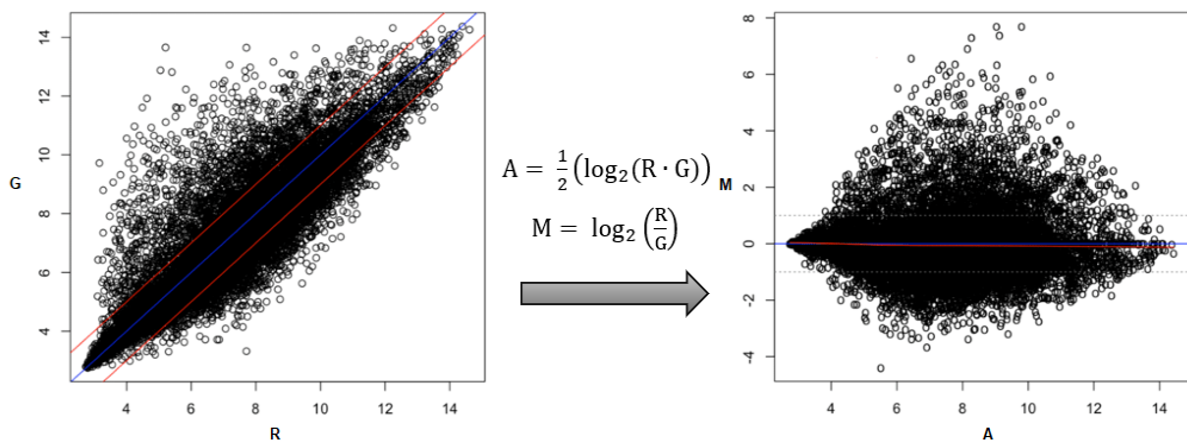
Αυτό υποδεικνύει και το πως θα πρέπει να μοιάζει ένα διάγραμμα των εντάσεων. Για παράδειγμα, εάν δεν υπήρχαν τεχνικά σφάλματα, σε μια μικροσυστοιχία δύο καναλιών, ένα διάγραμμα διασποράς των εντάσεων του κόκκινου έναντι του πράσινου θα πρέπει να αφήνει τα περισσότερα σημεία γύρω από μια διαγώνιο γραμμή. Οποιαδήποτε απόκλιση από αυτή την κατάσταση θα πρέπει να αποδίδεται σε τεχνικά σφάλματα, δηλαδή μη βιολογικούς λόγους, και κατά συνέπεια θα πρέπει να αφαιρείται. Αυτό έχει οδηγήσει σε μια πολύ δημοφιλή μέθοδο κανονικοποίησης που συνίσταται στην εκτίμηση του μετασχηματισμού που πρέπει να εφαρμοστεί ως συνάρτηση των εντάσεων, χρησιμοποιώντας τη μέθοδο Locally Weighted Scatterplot Smoothing (LOWESS) σε μία μετασχηματισμένη αναπαράσταση του διαγράμματος διασποράς, γνωστό και ως διάγραμμα MA.

Στην Εικόνα 20 απεικονίζεται ένα διάγραμμα διασποράς (scatter plot) του κόκκινου έναντι του πράσινου καναλιού για ένα πείραμα μικροσυστοιχιών. Το γεγονός ότι τα δεδομένα δεν είναι κεντραρισμένα γύρω από το διαγώνιο υποδηλώνει την ανάγκη κανονικοποίησης. Μια πολύ δημοφιλής

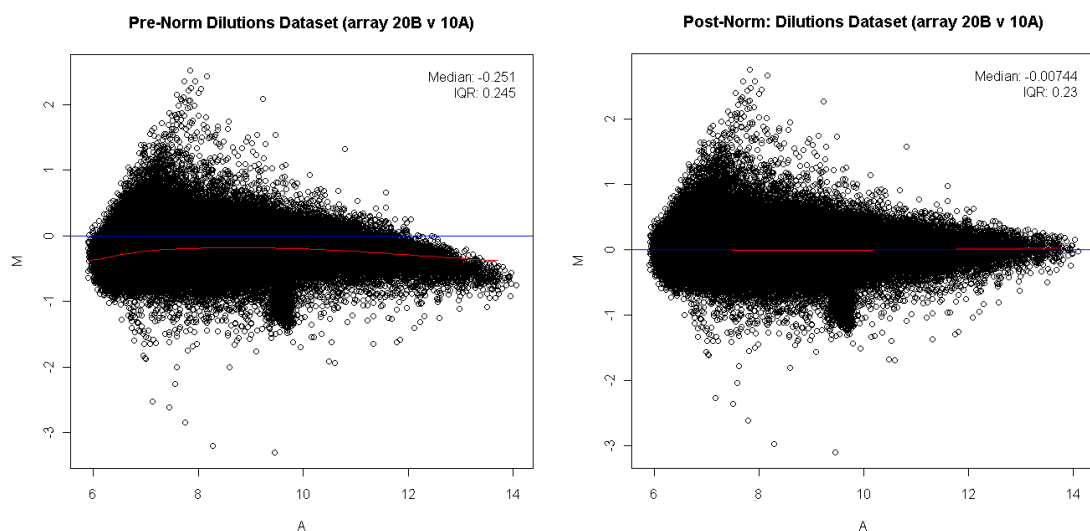
αναπαράσταση, η οποία βοηθά στην καλύτερη απεικόνιση αυτής της ασυμμετρίας είναι το διάγραμμα MA (Εικόνα 21). Γεωμετρικά αντιπροσωπεύουν ουσιαστικά μια περιστροφή του διαγράμματος διασποράς, όπου η έννοια του νέων αξόνων είναι:

- $A = \frac{1}{2}(\log_2(R \cdot G))$: η μέση λογαριθμική ένταση των δύο καναλιών
- $M = \log_2\left(\frac{R}{G}\right)$: Ο λογάριθμος της σχετικής έκφρασης μεταξύ των δύο καναλιών (συνήθως γνωστή ως «λογαριθμικός λόγος»)

Η Εικόνα 21 δείχνει το αποτέλεσμα της κανονικοποίησης των δεδομένων με τη μέθοδο LOWESS. Μετά την εφαρμογή της LOWESS στα δεδομένα, αφαιρείται πρακτικά μια διαφορετική ποσότητα από κάθε σημείο, ανάλογα με την έντασή του (τιμή A). Κατά συνέπεια, τα μετασχηματισμένα δεδομένα δεν είναι μόνο κεντραρισμένα αλλά και συμμετρικά γύρω από το μηδέν.



Εικόνα 20: Διάγραμμα διασποράς κόκκινου (R) έναντι πράσινου (G) καναλιού (αριστερά)
Διάγραμμα MA: μέση ένταση (A) έναντι λογαριθμικού λόγου (M) (δεξιά). [109]



Εικόνα 21: Διαγράμματα MA στις αρχικές / πρωτογενείς (αριστερά) και στις κανονικοποιημένες τιμές (δεξιά).
(Πηγές: https://commons.wikimedia.org/wiki/File:Pre-normalization_MAplot.png και
https://commons.wikimedia.org/wiki/File:Post-normalization_MAplot.png)

Η μέθοδος LOWESS κανονικοποιεί τις τιμές έκφρασης ώστε οι εντάσεις να είναι συνεπείς σε κάθε μικροσυστοιχία. Αυτή η τεχνική αποτελεί μια προσέγγιση κανονικοποίησης εντός ενός δείγματος

(within slide normalization). Σε πολλές περιπτώσεις, είναι επίσης απαραίτητο να επιτευχθεί συνοχή μεταξύ των μικροσυστοιχιών και μπορούν να εφαρμοστούν μέθοδοι όπως η προτυποποίηση Z (Z-score transformation), η κλιμακοποίηση (scaling), και η κανονικοποίηση ποσοστημορίων (quantile normalization). Η ιδέα της κλιμακοποίησης είναι απλώς να κλιμακωθούν οι λογάριθμοι έτσι ώστε να έχουν την ίδια διάμεση απόλυτη απόκλιση (median-absolute-deviation – MAD) σε όλες τις μικροσυστοιχίες. Η κανονικοποίηση ποσοστημορίων, η οποία μπορεί να χρησιμοποιηθεί τόσο σε μικροσυστοιχίες ενός όσο και σε μικροσυστοιχίες δύο χρωμάτων, εξασφαλίζει ότι οι εντάσεις έχουν την ίδια εμπειρική κατανομή σε όλες τις μικροσυστοιχίες και σε όλα τα κανάλια.

Οι μικροσυστοιχίες ενός καναλιού, παρουσιάζουν κάποια διαφορετικά τεχνικά σφάλματα που απαιτούν διαφορετικές μεθόδους κανονικοποίησης. Η πιο συχνά χρησιμοποιούμενη μέθοδος για αυτού του τύπου τις μικροσυστοιχίες είναι ο Robust Multi-array Average (RMA). Ο αλγόριθμος αυτό αποτελείται από τρία βήματα: αρχικά, γίνεται προσαρμογή του υποβάθρου με βάση ένα μοντέλο σε επίπεδο ανιχνευτών, στη συνέχεια λαμβάνει χώρα η κανονικοποίηση ποσοστημορίων και, τέλος, γίνεται η σύνοψη που πρακτικά ενσωματώνει τις τιμές όλων των ανιχνευτών που αντιστοιχούν σε ένα γονίδιο. Η μέθοδος RMA είναι πολύ δημοφιλής μεταξύ ερευνητών και βιο-στατιστικολόγων επειδή βασίζεται σε εύληπτα μαθηματικά μοντέλα που επιτρέπουν την κατανόηση της λογικής που διέπει τη μέθοδο. Μια εννοιολογικά απλούστερη προσέγγιση είναι αυτή που προτείνεται από τους κατασκευαστές των μικροσυστοιχιών: ο αλγόριθμος MAS5. Ορισμένες μελέτες που συγκρίνουν τις δύο (και άλλες) μεθόδους κανονικοποίησης καταλήγουν στην υπεροχή του RMA.

Γενικότερα, έχουν προταθεί πολλοί τρόποι για την κανονικοποίηση των δεδομένων μικροσυστοιχιών. Η επιλογή της μεθόδου που θα χρησιμοποιηθεί εξαρτάται από:

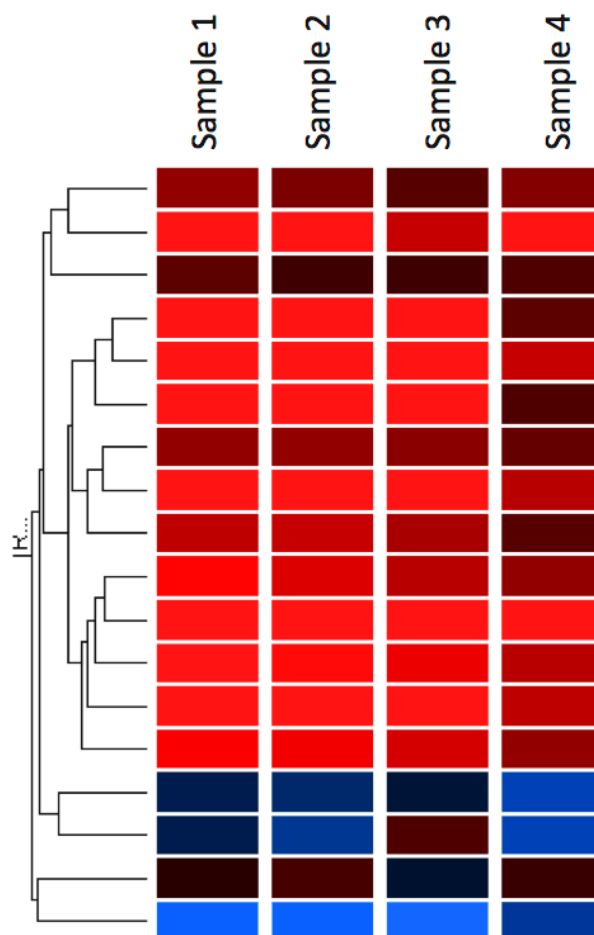
- τον τύπο της μικροσυστοιχίας,
- το σχεδιασμό του πειράματος,
- τις παραδοχές που γίνονται σχετικά με τα δεδομένα (π.χ. «η πλειοψηφία των γονιδίων που αντιπροσωπεύονται στη μικροσυστοιχία δεν αναμένεται να εκφράζονται διαφορετικά στην ομάδα δοκιμής σε σχέση με την ομάδα ελέγχου»), και
- το πακέτο που χρησιμοποιείται για την ανάλυση των δεδομένων.

Όσον αφορά τις μικροσυστοιχίες των μεγαλύτερων κατασκευαστούν, τα δεδομένα μικροσυστοιχιών της Affymetrix κανονικοποιούνται ως επί το πλείστον με τη μέθοδο RMA, η οποία υλοποιείται στα πακέτα της R *affy* και *oligo*. Τα δεδομένα μικροσυστοιχιών της Agilent κανονικοποιούνται με τη χρήση του πακέτου της R *limma*, και συγκεκριμένα με χρήση της συνάρτησης *quantile normalisation* για δεδομένα μικροσυστοιχιών ενός χρώματος. Τέλος, το ίδιο πακέτο μπορεί να εφαρμοστεί και για την κανονικοποίηση των δεδομένων μικροσυστοιχιών της Illumina.

Ο στόχος της **ανάλυσης διαφορικής έκφρασης** είναι ο εντοπισμός γονιδίων των οποίων η έκφραση διαφέρει (διαφορικά εκφραζόμενα γονίδια) υπό διαφορετικές συνθήκες. Μια σημαντική παρατήρηση για την ανάλυση διαφορικής έκφρασης είναι η διόρθωση των στατιστικών τεστ για τους ελέγχους πολλαπλών υποθέσεων (multiple testing correction). Πρόκειται για ένα στατιστικό φαινόμενο που εμφανίζεται όταν πραγματοποιούνται χιλιάδες συγκρίσεις (π.χ. σύγκριση της έκφρασης πολλών γονιδίων σε πολλαπλές συνθήκες) για μικρό αριθμό δειγμάτων (τα περισσότερα πειράματα μικροσυστοιχιών έχουν λιγότερες από πέντε βιολογικές επαναλήψεις ανά συνθήκη). Αυτό οδηγεί σε αυξημένη πιθανότητα ψευδώς θετικών αποτελεσμάτων.

Το πιο ευρέως χρησιμοποιούμενο πακέτο της R για τον εντοπισμό των γονιδίων που εκφράζονται διαφορετικά, το *limma*, ενσωματώνει μια μέθοδο διόρθωσης των πολλαπλών δοκιμών. Η μέθοδος αυτή δημιουργεί το λογαριθμισμένο ($\log_2$) λόγο των πολλαπλασίων μεταβολών (fold change ratio) μεταξύ της συνθήκης δοκιμής και της συνθήκης ελέγχου και μια «προσαρμοσμένη» τιμή-p (adjusted p-value) που εκφράζει τη σημαντικότητα της διαφοράς.

Έπειτα, ιδιαίτερο ουσιαστικό στάδιο αποτελεί η **βιολογική ερμηνεία των αποτελεσμάτων**, η οποία αποτελεί συνήθως και τον απώτερο στόχο μιας μελέτης. Πολλές από τις μεθόδους οπτικοποίησης και βιολογικής ερμηνείας των δεδομένων γονιδιακής έκφρασης μπορούν να χρησιμοποιηθούν τόσο για τα πειράματα μικροσυστοιχιών όσο και για τα πειράματα RNA-seq. Μερικές από τις πιο συνηθισμένες μεθόδους αποτελούν οι χάρτες θερμότητας, η ανάλυση εμπλουτισμού και η ανάλυση μονοπατιών. Μια κοινή μέθοδος οπτικοποίησης των δεδομένων γονιδιακής έκφρασης είναι η απεικόνισή τους με τη μορφή **θερμικού χάρτη** (Εικόνα 22). Ο θερμικός χάρτης μπορεί επίσης να συνδυαστεί με μεθόδους ομαδοποίησης / συσταδοποίησης (clustering), οι οποίες ομαδοποιούν γονίδια ή / και δείγματα με βάση την ομοιότητα του μοτίβου της γονιδιακής έκφρασής τους. Η τεχνική αυτή είναι ιδιαίτερα χρήσιμη για τον εντοπισμό γονιδίων που ρυθμίζονται από κοινού ή βιολογικών υπογραφών που σχετίζονται με μια συγκεκριμένη κατάσταση (π.χ. μια ασθένεια ή μια περιβαλλοντική κατάσταση).



Εικόνα 22: Ένα παράδειγμα ενός θερμικού χάρτη στον οποίο τα γονίδια έχουν ομαδοποιηθεί με βάση το μοτίβο της γονιδιακής τους έκφρασης.

(Πηγή: <https://www.ebi.ac.uk/training/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/biological-interpretation-of-gene-expression-data-2/>)

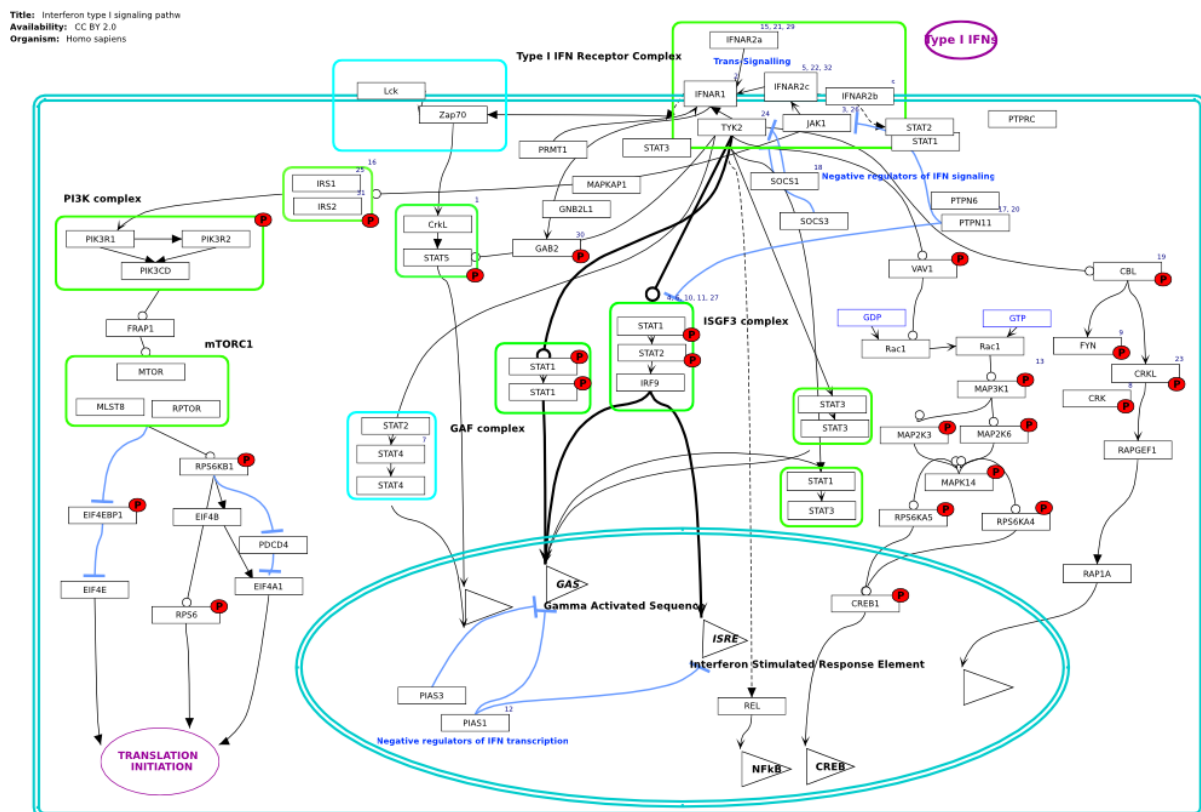
Στους θερμικούς χάρτες τα δεδομένα εμφανίζονται σε ένα πλέγμα όπου κάθε γραμμή αντιπροσωπεύει ένα γονίδιο και κάθε στήλη αντιπροσωπεύει ένα δείγμα. Το χρώμα και η ένταση των πλαισίων χρησιμοποιούνται για να αναπαραστήσουν τις αλλαγές (όχι τις απόλυτες τιμές) της γονιδιακής έκφρασης (υπερ-έκφραση ή υπο-έκφραση). Στην παραπάνω εικόνα, το κόκκινο χρώμα αντιπροσωπεύει γονίδια με αυξημένη έκφραση (up-regulated genes) και το μπλε χρώμα γονίδια με μειωμένη έκφραση (down-regulated genes). Το μαύρο χρώμα αντιπροσωπεύει γονίδια με αμετάβλητη έκφραση.

Επιπρόσθετα, μια συνήθης προσέγγιση για την ερμηνεία των δεδομένων γονιδιακής έκφρασης είναι η **ανάλυση εμπλουτισμού σε σύνολα γονιδίων** (gene set enrichment analysis) και η **λειτουργική ανάλυση εμπλουτισμού** (functional enrichment analysis) με βάση τον λειτουργικό σχολιασμό / ερμηνεία των γονιδίων που εκφράζονται διαφορετικά (Εικόνα 23). Αυτό είναι χρήσιμο για να διαπιστωθεί εάν τα γονίδια που εκφράζονται διαφορετικά σχετίζονται με μια συγκεκριμένη βιολογική διαδικασία ή μοριακή λειτουργία.

Η Οντολογία των Γονιδίων (Gene Ontology), που περιέχει τυποποιημένο σχολιασμό των γονιδιακών προϊόντων, χρησιμοποιείται συνήθως για το σκοπό αυτό. Λειτουργεί συγκρίνοντας τη συχνότητα των μεμονωμένων σχολιασμών στον κατάλογο των γονιδίων ενδιαφέροντος (π.χ. διαφορεικά εκφραζόμενα γονίδια) με έναν κατάλογο αναφοράς (συνήθως όλα τα γονίδια της μικροσυστοιχίας ή του γονιδιώματος). Ο **εμπλουτισμός βιολογικών μονοπατιών** (enrichment of biological pathways), που παρέχονται από την Εγκυκλοπαίδεια του Κιότο για τα γονίδια και τα γονιδιώματα (KEGG), το Ingenuity, το Reactome ή το WikiPathways, μπορεί να πραγματοποιηθεί με παρόμοιο τρόπο.

Τα δημοφιλέστερα εργαλεία για την ανάλυση εμπλουτισμού σε σύνολα γονιδίων και την ανάλυση βιολογικών μονοπατιών περιλαμβάνουν:

- την πλατφόρμα Database for Annotation, Visualization and Integrated Discovery – DAVID (δωρεάν διαδικτυακό εργαλείο),
- την υπολογιστική μέθοδο Gene Set Enrichment Analysis – GSEA (δωρεάν),
- το λογισμικό Ingenuity Pathway Analysis της QIAGEN (απαιτείται άδεια χρήσης), και
- την πλατφόρμα / βάση γνώσης Reactome (δωρεάν).

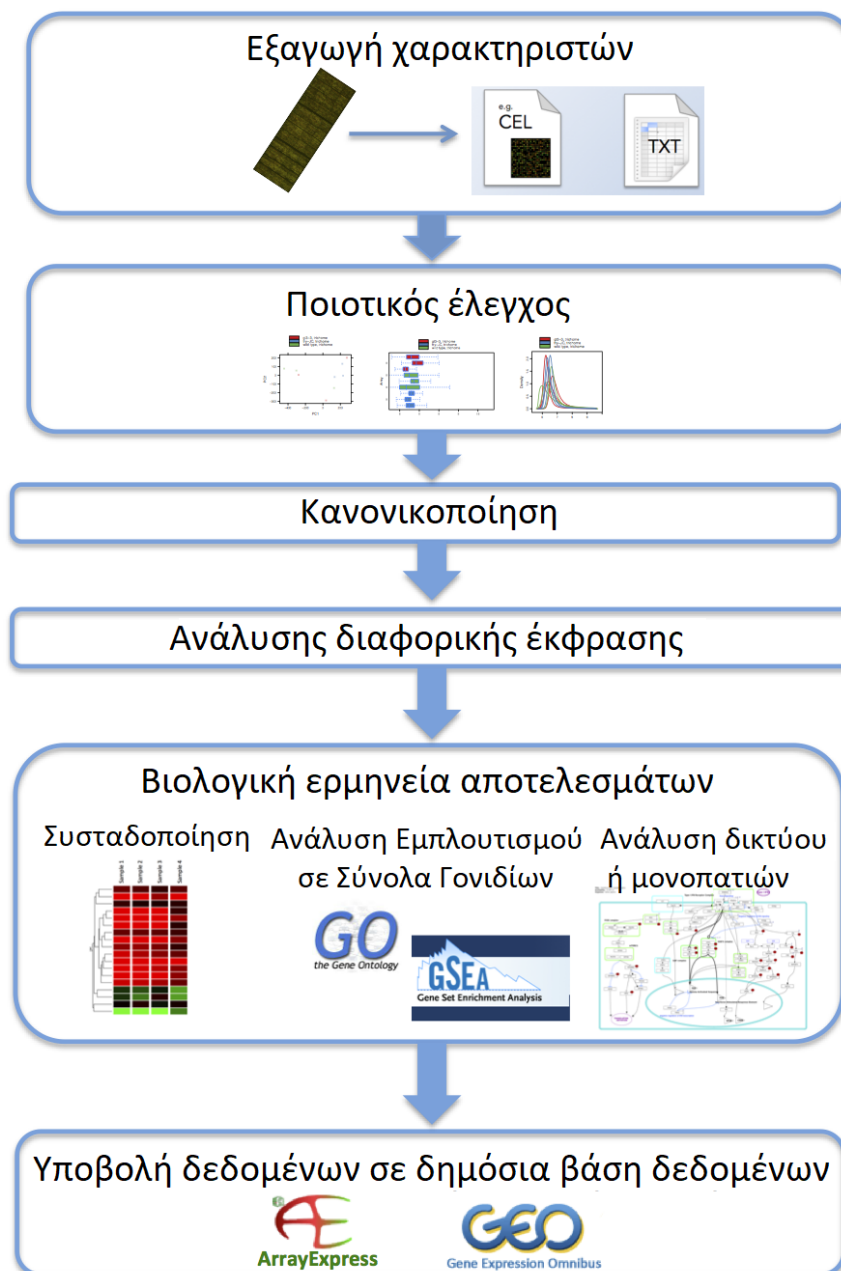


Εικόνα 23: Ένα παράδειγμα ανάλυσης μονοπατιών από το WikiPathways.

(Πηγή: <https://www.ebi.ac.uk/training/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/biological-interpretation-of-gene-expression-data-2/>)

Τέλος, η **ανάλυση δικτύου** είναι συμπληρωματική της ανάλυσης μονοπατιών και μπορεί να χρησιμοποιηθεί για να δείξει πως αλληλεπιδρούν τα βασικά συστατικά διαφορετικών μονοπατιών. Αυτό μπορεί να είναι χρήσιμο για τον εντοπισμό ρυθμιστικών γεγονότων που επηρεάζουν πολλαπλές βιολογικές διεργασίες και μονοπάτια.

Μόλις ολοκληρωθεί η δημιουργία των δεδομένων μικροσυστοιχιών (π.χ. τα αρχεία πρωτογενών δεδομένων και τα αρχεία των επεξεργασμένων / κανονικοποιημένων δεδομένων είναι έτοιμα), είναι σημαντικό να υποβληθούν όλα τα αρχεία δεδομένων μαζί με τα μετα-δεδομένα σε μια **δημόσια βάση δεδομένων**, όπως το Gene Expression Omnibus (GEO) του NCBI ή το ArrayExpress του European Bioinformatics Institute (EMBL-EBI). Αυτό συμβάλλει στη διασφάλιση της αναπαραγωγιμότητας του πειράματος και αποτελεί πλέον απαίτηση πολλών περιοδικών και φορέων χρηματοδότησης, με σκοπό τη δυνατότητα διάχυσης και επαναχρησιμοποίησης των δεδομένων και αποτελεσμάτων.

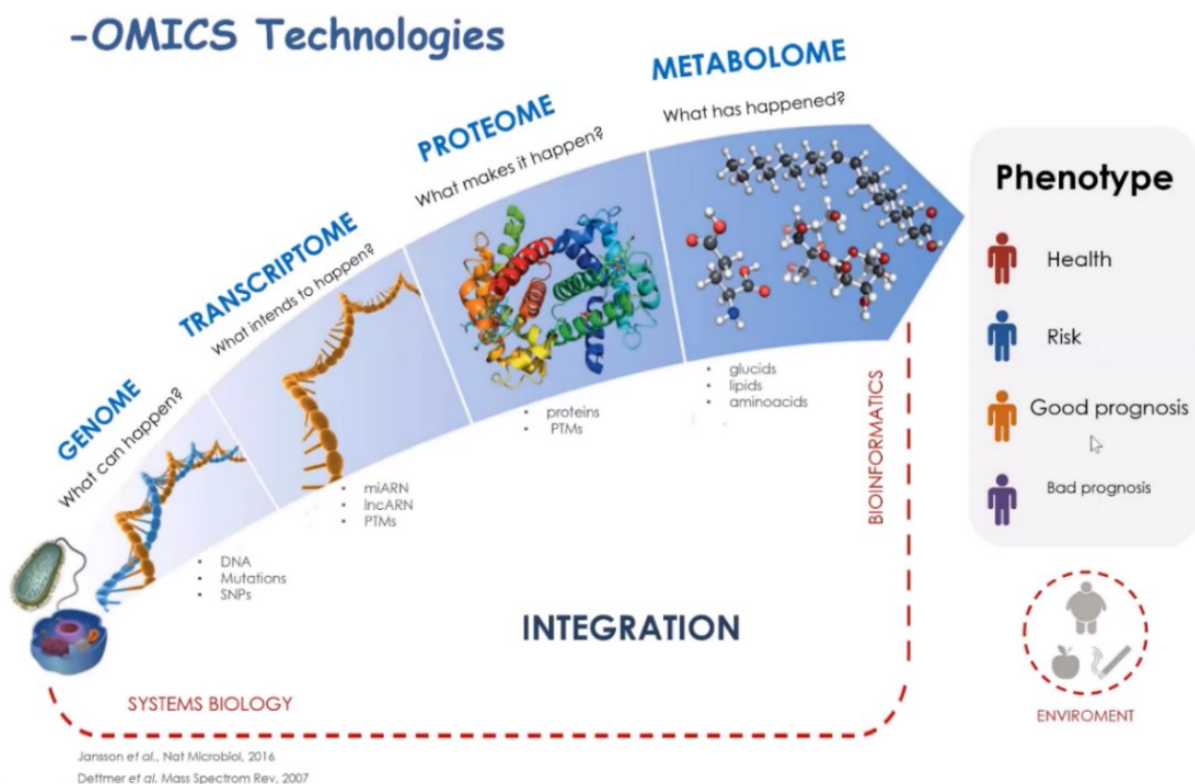


Εικόνα 24: Επισκόπηση της διοχέτευσης (pipeline) ανάλυσης δεδομένων μικροσυστοιχιών.

(Πηγή: <https://www.ebi.ac.uk/training/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/microarrays/analysis-of-microarray-data/>)

1.7 Ομικές τεχνολογίες και βιολογικά δεδομένα

Οι κλάδοι της επιστήμης που είναι ανεπίσημα γνωστοί ως **ομικές τεχνολογίες (-omics)** είναι οι διάφοροι κλάδοι της βιολογίας των οποίων τα ονόματα τελειώνουν στην κατάληξη -omics, όπως η γονιδιωματική (genomics), η μεταγραφομική (transcriptomics), η πρωτεωμική (proteomics), η μεταβολομική (metabolomics), η μεταγονιδιωματική (meta-genomics), και η φαινομική (phenomics) (Εικόνα 25). Οι ομικές τεχνολογίες στοχεύουν στον συλλογικό χαρακτηρισμό και στην ποσοτικοποίηση ομάδων βιολογικών μορίων που μεταφράζονται στη δομή, τη λειτουργία και τη δυναμική ενός οργανισμού ή μιας ομάδας οργανισμών. Η σχετική κατάληξη -ome χρησιμοποιείται για να απευθύνεται στο σύνολο των αντικειμένων μελέτης αυτών των πεδίων, όπως το γονιδίωμα (genome), το πρωτέωμα (proteome) ή το μεταβόλωμα (metabolome), αντίστοιχα. Η λειτουργική γονιδιωματική αποσκοπεί στον προσδιορισμό των λειτουργιών όσο το δυνατόν περισσότερων γονιδίων ενός συγκεκριμένου οργανισμού. Συνδυάζει διάφορες ομικές τεχνολογίες, όπως η μεταγραφομική και η πρωτεωμική.

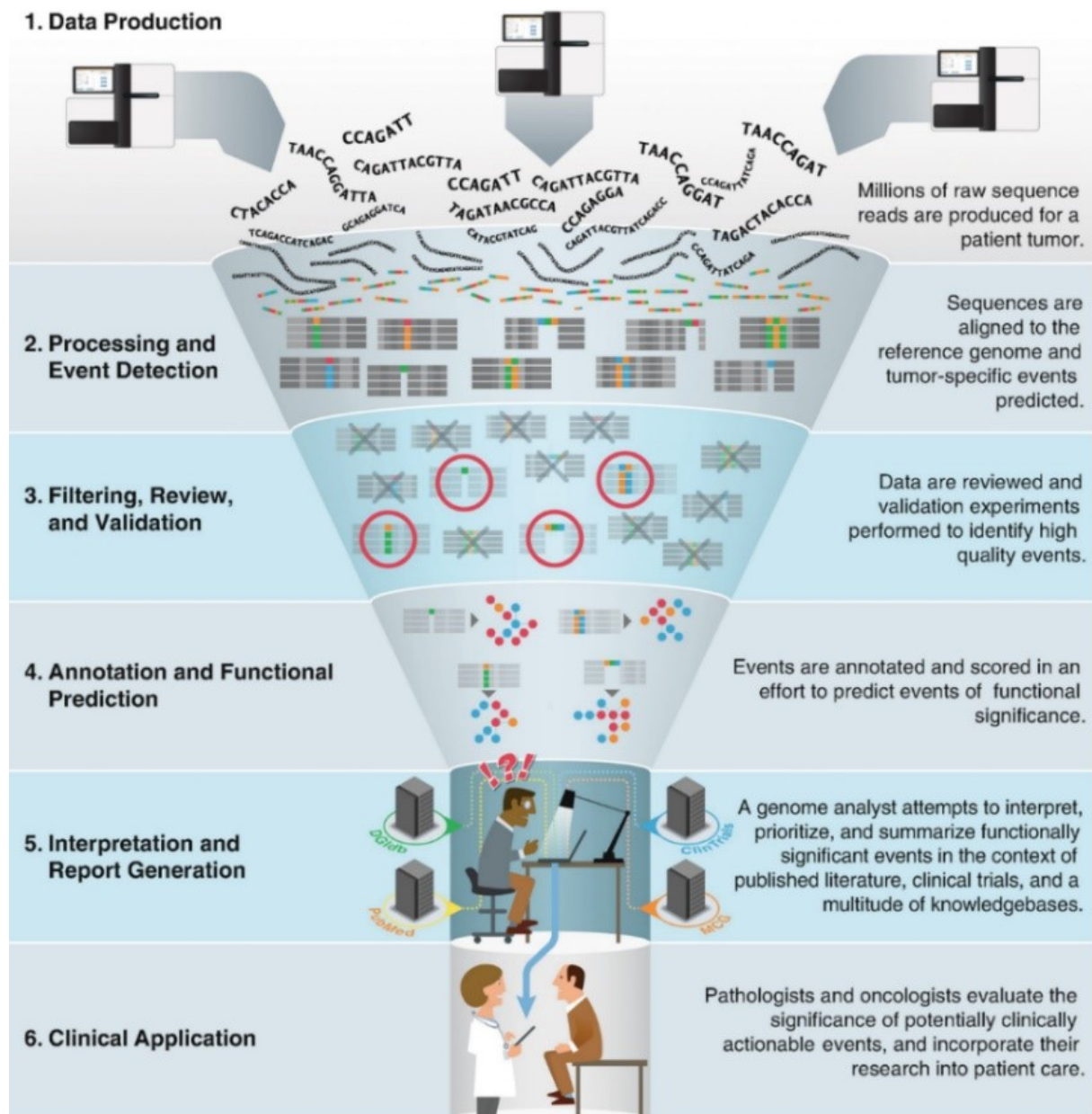


Εικόνα 25: Οι σύγχρονες ομικές (-omics) τεχνολογίες αποτελούν τις υψηλού επιπέδου τεχνολογίες που χρησιμοποιούνται για την ολιστική ανάλυση των μορίων που συνθέτουν τα κύτταρα των ζωντανών οργανισμών.

(Πηγή: <https://www.pnnl.gov/projects/soil-microbiome/research>)

Έτσι, οι ομικές τεχνολογίες αποτελούν βιοχημικές αναλύσεις υψηλής απόδοσης (high-throughput) που μετρούν συνολικά και ταυτόχρονα μόρια του ίδιου τύπου από ένα βιολογικό δείγμα. Η έννοια των "omics" αναφέρεται στο γεγονός ότι όλες ή σχεδόν όλες οι υπάρξεις του υπό μελέτη μοριακού χώρου μετρώνται στη δοκιμασία, και ως εκ τούτου παρέχουν ολιστικές απόψεις του υπό μελέτη βιολογικού συστήματος. Αρχικά, τα ομικά πειράματα επικεντρώνονταν σε έναν τύπο ανάλυσης (π.χ. transcriptomics) και παρείχαν δεδομένα single-omics. Ωστόσο, πιο πρόσφατα οι ερευνητές συνδύασαν πολλαπλές δοκιμασίες για το ίδιο σύνολο δειγμάτων ώστε να δημιουργήσουν σύνολα δεδομένων multi-omics. Οι γνώσεις των πρώτων μεμονωμένων ομικών έργων, όπως το Human Genome Project, και ο πολλαπλασιασμός των εγκαταστάσεων που προσφέρουν οικονομικά προσιτές υπηρεσίες ομικών

αναλύσεων, οδήγησαν στην ανάπτυξη πολλών νέων multi-omics έργων. Η πρόσθετη ισχύς των multi-omics είναι εμφανής εδώ και αρκετό καιρό, αλλά η πολυπλοκότητα της διαχείρισης και της ενσωμάτωσης τέτοιων πολυδιάστατων δεδομένων εξακολουθεί να αποτελεί σημαντική πρόκληση. Η αποθήκευση των δεδομένων, ο έλεγχος της ποιότητάς τους και η στατιστική ανάλυσή τους είναι πολύπλοκες διαδικασίες και η συμμόρφωση με τις «Κατευθυντήριες Αρχές FAIR για την διαχείριση και προστασία επιστημονικών δεδομένων» είναι εγγενώς πιο δύσκολη.



Εικόνα 26: Η πληθώρα των σύγχρονων τεχνολογιών παραγωγής βιολογικών δεδομένων δημιουργηθεί συμφόρηση (bottleneck) ως προς την ανάλυση, ερμηνεία και κλινική αξιοποίηση των δεδομένων αυτών. [110]

Διάφοροι τύποι αναλύσεων που εκτελούνται σε δεδομένα multi-omics βασίζονται σήμερα σε τεχνικές αλληλούχισης επόμενης γενιάς (Next Generation Sequencing – NGS) οι οποίες παράγουν μεγάλους όγκους αλληλουχιών. Η αυξανόμενη δημοτικότητα και η προσβασιμότητα των τεχνολογιών αλληλούχισης επόμενης γενιάς έχουν δημιουργήσει υπερβολικά μεγάλο όγκο δεδομένων, απαιτώντας ταυτόχρονη πρόοδο στις διαδικασίες της επεξεργασίας και της ανάλυσής τους. Η τεχνολογία NGS

περιγράφει μια συλλογή πλατφορμών που επιτρέπουν την ταχεία σκιαγράφιση του συνόλου των αλληλουχιών των νουκλεϊκών οξέων. Κατά την τελευταία δεκαετία, διάφορες πρωτοποριακές τεχνολογίες που εμφανίζουν ιδιαίτερα καλή αποτελεσματικότητα και μειωμένο κόστος διευκόλυναν μεγάλα έργα αλληλούχισης γονιδιώματος και μεταγραφώματος σε διαφορετικά γεωγραφικά και εθνικά περιβάλλοντα. Ως εκ τούτου, ο τεράστιος παραγόμενος όγκος δεδομένων έχει φέρει επανάσταση στη βιοϊατρική έρευνα παρέχοντας μια πληρέστερη κατανόηση των βιολογικών συστημάτων και των μοριακών μηχανισμών ανάπτυξης ασθενειών.

Το άμεσο κόστος που σχετίζεται με την εκτέλεση αντιδράσεων αλληλούχισης έχει μειωθεί δραματικά, ενώ το κόστος επεξεργασίας, αποθήκευσης, διαχείρισης και ερμηνείας των δεδομένων που παράγουν αυτές οι τεχνολογίες έχει αυξηθεί εκθετικά. Η υιοθέτηση των τεχνολογιών NGS σε ερευνητικά και κλινικά περιβάλλοντα παρήγαγε petabytes γονιδιωματικών δεδομένων, δημιουργώντας ένα σημείο συμφόρησης (bottleneck) της διαδικασίας (Εικόνα 26) και ιδιαίτερες προκλήσεις για την επιστήμη της βιοπληροφορικής. Η διαχείριση, η επεξεργασία, η ενημέρωση, η διάδοση και η ενοποίηση των δεδομένων αποτελούν τους κύριους τομείς σύγχρονων προκλήσεων. Έτσι, η ραγδαία αύξηση των διαφόρων ομικών τεχνολογιών προκαλεί δύο μείζονα ζητήματα, τα οποία μπορεί και να φαίνονται αντίθετα: τη δημοσίευση ανεπαρκώς επιμελημένων συνόλων δεδομένων, που εξακολουθούν να έχουν τη μορφή ακατέργαστων / πρωτογενών δεδομένων ή προκαταρκτικών αποτελεσμάτων, και την ταχεία διάχυση και δημοσίευση πληροφοριών που, σαν συνέπεια, επηρεάζουν δυσμενώς τη δημιουργία στέρεων και αξιόπιστων πηγών. Τα ζητήματα αυτά οφείλονται κυρίως στον χαμηλότερο ρυθμό που εμφανίζει η επιστήμη της βιοπληροφορικής στην εξαγωγή πληροφοριών προστιθέμενης αξίας από τον μεγάλο όγκο δεδομένων, σε σύγκριση με τις ταχύτατα αναπτυσσόμενες τεχνολογίες που εμπλέκονται στην παραγωγή δεδομένων. Επιπλέον, η επέκταση της ιατρικής ακριβείας έχει καταστήσει αναγκαίο το δύσκολο έργο του συνδυασμού των γονιδιωματικών δεδομένων με κλινικά δεδομένα, όπως τα δεδομένα του ηλεκτρονικού φακέλου υγείας, για την απόκτηση νέων πληροφοριών και την εξαγωγή συμπερασμάτων.

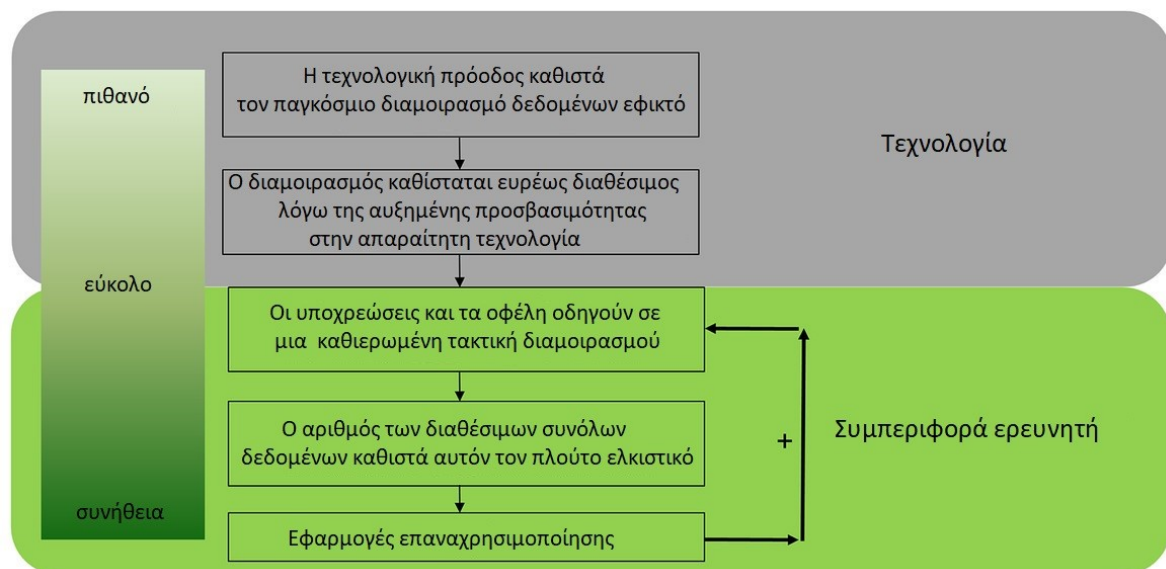
1.8 Επαναχρησιμοποίηση Δημόσιων Δεδομένων Γονιδιακής Έκφρασης

Η αυξανόμενη χρήση των τεχνολογιών υψηλής απόδοσης για την ανάλυση της γονιδιακής έκφρασης τις τελευταίες δύο δεκαετίες και η κατάθεση της συντριπτικής πλειοψηφίας των ερευνητικών δεδομένων σε δημόσια αποθετήρια έχουν δημιουργήσει μία πληθώρα δημόσια διαθέσιμων δεδομένων [111]. Όλα αυτά τα δεδομένα προσφέρουν έναν ανεκτίμητο πλούτο για την ερευνητική κοινότητα ο οποίος μπορεί να επαναχρησιμοποιηθεί μέσω μετα-αναλυτικών μεθόδων με σκοπό να την ανακάλυψη επιστημονικών ευρημάτων και νέας γνώσης. Ειδικότερα, η προσέγγιση της μετα-ανάλυσης και ενοποίησης δεδομένων από πολλαπλές πειραματικές μελέτες επιτρέπει την αύξηση του μεγέθους του δείγματος, της στατιστικής ισχύος και της ευρωστίας των πειραματικών αποτελεσμάτων [111, 112], καθώς και τη βελτίωση της αναπαραγωγιμότητας [113] του πειράματος αλλά και της εγκυρότητας των βιολογικών πληροφοριών που εξάγονται.

Επαναχρησιμοποίηση δεδομένων ως μέρος της επανάστασης των «μεγάλων δεδομένων»

Επιχειρώντας μια πιο σφαιρική προσέγγιση του θέματος, η επαναχρησιμοποίηση των δεδομένων αποτελεί βασικό στοιχείο της ανοικτής επιστήμης και διευκολύνεται από τη σύγχρονη επανάσταση των «μεγάλων δεδομένων» (big data). Η μετάβαση από (χειρόγραφες) σημειώσεις σε σύνολα δεδομένων αποθηκευμένα σε σκληρούς δίσκους μπορεί να θεωρηθεί ως το πρώτο βήμα στο δρόμο για την αποτελεσματική επαναχρησιμοποίηση δεδομένων στις βιοεπιστήμες, επιτρέποντας τη δημιουργία πολλαπλών αντιγράφων χωρίς σχεδόν κανένα πρόσθετο κόστος. Το δεύτερο βήμα ήταν η βελτίωση της

συνδεσιμότητας, η οποία κατέστη δυνατή χάρη στο διαδίκτυο. Έτσι, αυτές οι τεχνολογικές εξελίξεις στην αποθήκευση και τη μεταφορά δεδομένων, επέτρεψαν την παγκόσμια ανταλλαγή των «μεγάλων δεδομένων», τα οποία είναι συνηθισμένα στη βιολογία (π.χ. δεδομένα αλληλουχιών). Αυτές οι τεχνολογικές εξελίξεις κατέστησαν δυνατό τον εύκολο και άμεσο διαμοιρασμό των δεδομένων. Στη συνέχεια, ήταν ιδιαίτερα σημαντικό να καταστεί εύκολο για τους ερευνητές να μοιράζονται δεδομένα μέσω της αυξημένης προσβασιμότητας. Οι υποχρεώσεις και τα οφέλη που απορρέουν από αυτή τη διαδικασία οδηγούν στην καθιέρωση μιας συμπεριφοράς και κάποιων κανόνων κοινής χρήσης, τα οποία έχουν ως αποτέλεσμα να καθίστανται διαθέσιμα περισσότερα σύνολα δεδομένων, τα οποία μπορούν με τη σειρά τους να επαναχρησιμοποιηθούν. Τέλος, γίνεται κοινή πρακτική και συνήθεια η κοινή χρήση όλων των δεδομένων, με αποτέλεσμα έναν θετικό βρόχο ανατροφοδότησης (Εικόνα 27) [114]. Ο διαμοιρασμός των δεδομένων είναι ήδη συνήθης πρακτική σε διάφορους επιστημονικούς κλάδους, όπως η γονιδιωματική, οι νευροεπιστήμες, οι γεωεπιστήμες και η αστρονομία. Επίσης, ένας αυξανόμενος αριθμός μελετών επαναχρησιμοποιεί διαμοιρασμένα και ανοιχτά δεδομένα.



Εικόνα 27: Η εξέλιξη της συμπεριφοράς διαμοιρασμού δεδομένων. (1) Η τεχνική πρόοδος καθιστά δυνατή τον παγκόσμιο διαμοιρασμό μεγάλων συνόλων δεδομένων, (2) η αυξημένη προσβασιμότητα στην απαιτούμενη τεχνολογία τον καθιστά ευρέως διαθέσιμο, (3) οι υποχρεώσεις και τα οφέλη για τους ερευνητές καθιερώνουν μία συμπεριφορά / κουλτούρα διαμοιρασμού, (4) ο αριθμός των συνόλων δεδομένων αυξάνεται και τα καθιστά ελκυστικά, (5) η επαναχρησιμοποίηση αναπτύσσεται με την πάροδο του χρόνου - γεγονός που οδηγεί σε θετικό βρόχο ανατροφοδότησης και στη συνήθεια να μοιράζονται όλα τα δεδομένα. [114]

Για λόγους σαφήνειας, είναι σημαντικό να γίνει μια διάκριση μεταξύ της δίκαιης επαναχρησιμοποίησης (για νέους σκοπούς, π.χ. μετα-ανάλυση), της αναπαραγωγής προηγούμενων μελετών με διαθέσιμα δεδομένα (ζωτικό στοιχείο της ανοικτής επιστήμης) και της άδικης επαναχρησιμοποίησης (διπλή δημοσίευση και λογοκλοπή). Παράλληλα, μαζί με την αναπαραγωγή των μελετών για τον έλεγχο των ευρημάτων, μόνο η δίκαιη επαναχρησιμοποίηση θα πρέπει να θεωρείται αποδεκτή συνιστώσα της ανοικτής επιστήμης και είναι το είδος της επαναχρησιμοποίησης για το οποίο γίνεται λόγος εδώ.

Για να καθιερωθεί η κοινή χρήση των δεδομένων ως κανόνας έπρεπε να εισαχθεί μέσω υποχρεώσεων και να προωθηθεί μέσω κάποιων πλεονεκτημάτων και οφελιών για τους ερευνητές. Πολυάριθμοι φορείς χρηματοδότησης, εκδότες και πανεπιστημιακά ιδρύματα απαιτούν όλα τα δεδομένα να δημοσιοποιούνται εντός συγκεκριμένου χρονικού πλαισίου, υπό το σκεπτικό ότι αυτό οδηγεί σε αυξημένη διαφάνεια στον τομέα. Πολλές διεθνείς κατευθυντήριες γραμμές για την κοινή χρήση

δεδομένων, όπως η Findable Accessible Interoperable Reusable (FAIR), η TOP (The Transparency and Openness Promotion), τα ανοικτά δεδομένα σε έναν κόσμο μεγάλων δεδομένων (Open Data in a Big Data World, 2016) και η Διακήρυξη του Πεκίνου (CODATA, 2019) προέκυψαν από την ανάγκη της επανάστασης των «μεγάλων δεδομένων». Όπως προαναφέρθηκε, η κοινή χρήση των συνόλων δεδομένων οδηγεί σε στατιστική ευρωστία και επιτρέπει την εκ νέου ανάλυση των υφιστάμενων συνόλων δεδομένων που διέπουν τους ισχυρισμούς των συγγραφέων, ενώ επιτρέπει ακόμη την ανακάλυψη νέων μοτίβων μέσω της μετα-ανάλυσης. Μια τέτοια επαναχρησιμοποίηση δεδομένων οδηγεί στη μείωση του κόστους, στην αναπαραγωγικότητα και στον αποδοτικότερο έλεγχο της έρευνας, στην επιστημονική ανακάλυψη και στην ανίχνευση νέων βιολογικών πληροφοριών. Τέλος, μπορεί να φανεί ιδιαίτερα χρήσιμη σε άλλους τομείς όπως η εκπαίδευση, οι επιχειρήσεις και οι κυβερνητικοί φορείς.

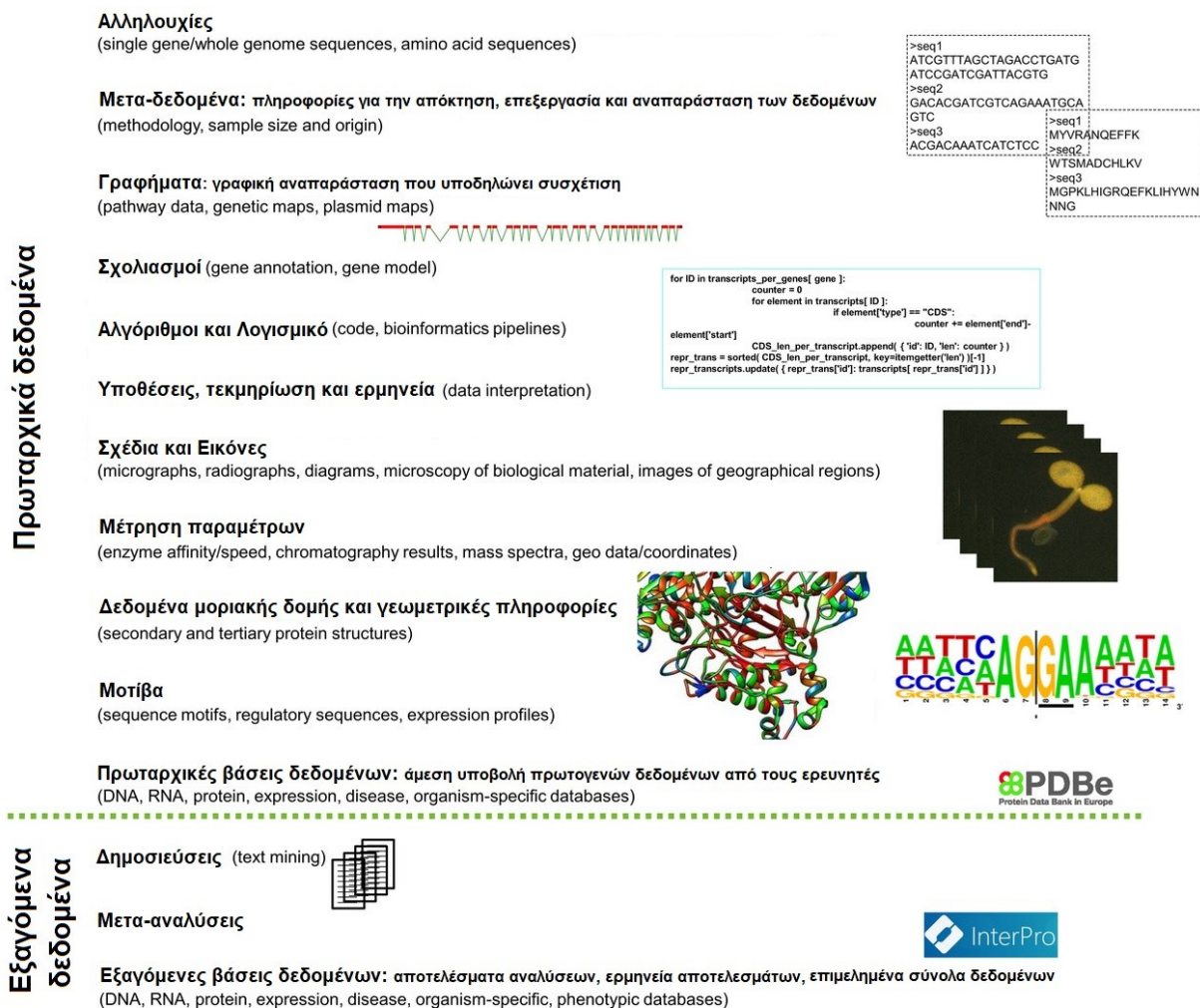
Παρόλες τις υποχρεώσεις και τα οφέλη, η επαναχρησιμοποίηση δεδομένων δεν είναι ακόμη πανταχού παρούσα, γεγονός για το οποίο μπορεί να υπάρχουν διάφορες εξηγήσεις. Το 2017, μια ανάλυση 318 βιοϊατρικών περιοδικών αποκάλυψε ότι μόνο το 11,9% από αυτά ανέφεραν ρητά ότι η κοινή χρήση δεδομένων απαιτείται ως προϋπόθεση για τη δημοσίευση [115]. Επιπλέον, μια έρευνα σε 100 σύνολα δεδομένων που σχετίζονται με την οικολογική και εξελικτική έρευνα έδειξε ότι το 56% των βάσεων δεδομένων ήταν ελλιπείς και το 64% είχαν αρχειοθετηθεί με τρόπο που εμποδίζει εν μέρει ή εξ ολοκλήρου την επαναχρησιμοποίηση [116]. Έτσι, εάν τα δημόσια διαθέσιμα σύνολα δεδομένων δεν επανεξετάζονται ευρέως (είτε ελέγχονται για την ποιότητα είτε/και επανααναλύονται), οι υποχρεώσεις της ανοικτής επιστήμης μέσω πολιτικών μπορεί να μην είναι επαρκής για την πλήρη αξιοποίηση της δύναμης του παγκόσμιου διαμοιρασμού. Οι κύριες αιτίες που κάνουν τους ερευνητές να απέχουν από την επαναχρησιμοποίηση δημοσίως διαθέσιμων συνόλων δεδομένων είναι: (1) η ανησυχία για την ποιότητα και την αξιοπιστία των δεδομένων (συχνά δικαιολογημένη), (2) η έλλειψη ενημέρωσης για τις δυνατότητες που κρύβουν τα «μεγάλα δεδομένα», και (3) η ανεπαρκής γνώση βιοπληροφορικής για την εξόρυξη των δεδομένων. Ανεξάρτητα από την αιτία, η προκύπτουσα «συσσώρευση» ανεπαρκώς επιμελημένων και αξιόπιστων συνόλων δεδομένων οδηγεί σε περιττά πειράματα (π.χ. εκτεταμένη επαναλαμβανόμενη αλληλούχιση που αυξάνει το κόστος) και πιθανότατα αλλοιώνει χρήσιμα ανεξερευνήτα μοτίβα.

Τύποι επαναχρησιμοποιήσιμων δεδομένων

Υπάρχουν πολυάριθμοι διαφορετικοί τύποι συνόλων δεδομένων με δυνατότητες επαναχρησιμοποίησης. Σε αυτά περιλαμβάνονται, ενδεικτικά: (1) δημοσιεύσεις που είναι προσβάσιμες στην εξόρυξη κειμένου, (2) αλληλουχίες γονιδιωμάτων, μεμονωμένων γονιδίων ή πλασμιδίων ή ολόκληρα σύνολα αναγνώσεων αλληλουχιών, (3) σχολιασμοί αλληλουχιών (π.χ. μοτίβα αλληλουχιών που συλλέγονται στη βάση δεδομένων JASPAR), (4) αποτελέσματα χρωματογραφίας και φασματομετρία μάζας, (5) πληροφορίες σχετικά με τη δομή των πρωτεϊνών, (6) βιοχημικές παράμετροι ενζύμων (π.χ. συγγένεια ή ταχύτητα αντίδρασης), (7) γεωγραφικά δεδομένα (π.χ. συντεταγμένες παρατηρήσεων), (8) εικόνες βιολογικού υλικού ή γεωγραφικών περιοχών καθώς και γραφικές παραστάσεις και διαγράμματα, (9) αλγόριθμοι και λογισμικό (π.χ. κώδικας), μετρήσεις αυτόματων αισθητήρων και δικτύων παρατήρησης παγκόσμιας κλίμακας, και (11) φαινοτυπικά δεδομένα.

Γενικά, τα σύνολα δεδομένων μπορούν να ταξινομηθούν σε πρωτογενή και εξαγόμενα (Εικόνα 28). Τα πρωτογενή μπορούν να οριστούν ως αυτά που περιλαμβάνουν τα «άμεσα, πειραματικά λαμβανόμενα δεδομένα» και τα εξαγόμενα ως αυτά που περιλαμβάνουν μετα-αναλύσεις και επεξεργασμένα αποτελέσματα. Τα στατιστικά δεδομένα αποτελούν παράδειγμα δεδομένων τα οποία δεν μπορούν να ταξινομηθούν σε πρωτογενή ή παραγόμενα. Τα αποτελέσματα ενός συγκεκριμένου πειράματος μπορούν να αξιολογηθούν άμεσα στατιστικά (πρωτογενή), ενώ τα αποτελέσματα πολλαπλών μελετών

μπορούν επίσης να αξιολογηθούν και να συγκριθούν με στατιστικές μεθόδους (παραγόμενα). Επίσης, είναι σημαντικό να αναφερθεί ότι σε συγκεκριμένους τομείς απαιτείται η ενοποίηση δεδομένων διαφόρων τύπων, μορφών και μεγεθών, κάτι που είναι δύσκολο να επιτευχθεί από μία μόνο βάση δεδομένων, και επομένως απαιτείται συνεργασία για την ενθάρρυνση της επαναχρησιμοποίησης των δεδομένων.



Εικόνα 28: Βασικοί τύποι επαναχρησιμοποιούμενων βιολογικών δεδομένων. [114]

1.9 Βιοπληροφορικές μετα-αναλύσεις για την ανεύρεση γονιδίων – κόμβων στον KOK

Η παραδοσιακή μετα-ανάλυση συνδυάζει τα αποτελέσματα από έναν αριθμό μελετών (ιδανικά όλων) που έχουν διεξαχθεί για το ίδιο ερευνητικό ερώτημα, με σκοπό τη στατιστική σύνοψη των ευρημάτων, την αξιολόγηση των αποκλίσεων και την ανίχνευση γενικεύσιμων αποτελεσμάτων. Η ικανότητα ανίχνευσης γενικών προτύπων καθιστά τις μετα-αναλύσεις εξαιρετικά σημαντικές για την ασθένεια του καρκίνου, η οποία χαρακτηρίζεται από εξαιρετική πολυπλοκότητα, γονιδιακή ετερογένεια και φαινοτυπική ποικιλομορφία. Οι βιοπληροφορικές μετα-αναλύσεις οι οποίες επιχειρούν την ενοποίηση πρωτογενών δεδομένων από πειράματα μικροσυστοιχιών με σκοπό την ανίχνευση σημαντικών γονιδίων για τον KOK είναι δυσεύρετες. Στη συνέχεια αναλύονται κάποιες πρόσφατες μελέτες οι οποίες διεξάγουν διερευνητικές μετα-αναλύσεις σε αποτελέσματα προγενέστερων πειραμάτων μικροσυστοιχιών για τον εντοπισμό γονιδίων – κόμβων.

Σε μία μελέτη του Wang κ.α. μετα-αναλύθηκαν δύο σύνολα δεδομένων από το αποθετήριο GEO με σκοπό την εύρεση γονιδίων που εκφράζονται διαφορετικά στον KOK και τη λειτουργική αξιολόγησή τους [117]. Η ανάλυση αποκάλυψε 619 κοινά διαφορεικά εκφραζόμενα γονίδια στα δύο σύνολα δεδομένων. Έπειτα από την ανάλυση του δικτύου αλληλεπίδρασης των πρωτεϊνών προέκυψαν τέσσερα γονίδια – κόμβοι, τα οποία ήταν τα *PGR*, *MAFG*, *CDC6*, και *MCMs* (οικογένεια γονιδίων). Οι συγγραφείς θεωρούν πως τα παραπάνω γονίδια συσχετίζονται στενά με την ανάπτυξη του KOK μέσω μονοπατιών που σχετίζονται με τη συγκρότηση των μιτοχονδρίων και των χρωμοσωμάτων.

Ο Zhang κ.α. διεξήγαγαν μία μελέτη με σκοπό τον εντοπισμό υποψήφιων γονιδίων στην καρκινογένεση και την εξέλιξη του KOK [118]. Οι συγγραφείς μετα-ανάλυσαν πέντε σύνολα δεδομένων από το αποθετήριο GEO και το σύνολο δεδομένων των ασθενών με KOK του προγράμματος TCGA και εντόπισαν 418 διαφορετικά εκφραζόμενα γονίδια. Έπειτα από την κατασκευή του δικτύου αλληλεπίδρασης πρωτεΐνης – πρωτεΐνης (PPIs) εντόπισαν 14 γονίδια – κόμβους και συγκεκριμένα τα *CDK1*, *CCNB1*, *CCNA2*, *KIF11*, *CDC20*, *UBE2C*, *MAD2L1*, *AURKA*, *KIF20A*, *KIF2C*, *KPNA2*, *TPM1*, *CASQ2*, και *CRYAB*. Η παρούσα μελέτη αναδεικνύει έναν, έως τώρα μη αναγνωρισμένο, πιθανό ρόλο του γονιδίου *CASQ2* στον KOK. Επιπλέον, το γονίδιο *CRYAB*, το οποίο δεν είχε αναφερθεί ποτέ ότι σχετίζεται με τον KOK, φάνηκε να αποτελεί υποψήφιο βιοδείκτη για αυτή την ασθένεια.

Ο Mo κ.α. διεξήγαγαν μία μετα-ανάλυση με σκοπό τον προσδιορισμό των γονιδίων – κόμβων που σχετίζονται με την καρκινογένεση και την εξέλιξη του KOK [119]. Στο πλαίσιο αυτό, ελήφθησαν τρία σύνολα δεδομένων από το αποθετήριο GEO και, μέσω της ανάλυσής τους, εντοπίστηκαν 141 διαφορετικά εκφραζόμενα γονίδια. Τα γονίδια αυτά τροφοδοτήθηκαν για τη δημιουργία του δικτύου αλληλεπίδρασης πρωτεΐνης – πρωτεΐνης. Η ανάλυση του δικτύου αποκάλυψε δέκα γονίδια – κόμβους, τα οποία ήταν τα γονίδια *TYMS*, *AURKB*, *CDK1*, *CCNB2*, *CEP55*, *KIF20A*, *KIF11*, *CENPE*, *PRC1*, και *CDC20*. Μεταξύ των γονιδίων αυτών, το γονίδιο *KIF11* αναγνωρίστηκε ως ένας από τους σημαντικότερους προγνωστικούς βιοδείκτες για τον KOK.

Ο Yan κ.α. χρησιμοποίησαν ένα σύνολο δεδομένων από το αποθετήριο GEO με σκοπό να εντοπίσουν σημαντικούς γονιδιακούς βιοδείκτες για τον KOK [120]. Από την ανάλυση του συνόλου δεδομένων προέκυψαν 976 διαφορετικά εκφραζόμενα γονίδια και, έπειτα από τις λειτουργικές αναλύσεις, κατασκευάστηκε το δίκτυο αλληλεπίδρασης πρωτεϊνών μέσω του οποίου προσδιορίστηκαν τα γονίδια – κόμβοι. Τα πιο σημαντικά από αυτά ήταν δώδεκα και συγκεκριμένα τα γονίδια *CCL5*, *SLPI*, *S100A9*, *CXCL1*, *TNFRSF1A*, *CCL20*, *MMP1*, *NOD2*, *IL1B*, *CXCR2*, *CD83*, και *MMP1*. Οι συγγραφείς ισχυρίστηκαν ότι τα παραπάνω γονίδια θα μπορούσαν να αποτελέσουν δυνητικά θεραπευτικούς στόχους για τον KOK.

Ο Xu κ.α. επιχειρώντας να προσδιορίσουν βιοδείκτες «κλειδιά» για τον KOK, μετα-ανάλυσαν τρία σύνολα δεδομένων, συμπεριλαμβανομένων τριών συνόλων από το αποθετήριο GEO και το σύνολο των ασθενών με KOK του προγράμματος TCGA [121]. Έπειτα από την ανάλυση διαφορικής έκφρασης, η οποία εντόπισε 291 γονίδια, έγινε ανάλυση του δικτύου αλληλεπίδρασης πρωτεΐνης – πρωτεΐνης, η οποία εντόπισε 14 γονίδια – κόμβους. Τα γονίδια αυτά ήταν τα *ASPM*, *CCNB2*, *CDC20*, *CENPF*, *CEP55*, *HJURP*, *KIF20A*, *NCAPG*, *NUSAP1*, *SPAG5*, *TOP2A*, *TRIP13*, *TROAP*, και *TTK*. Μεταξύ αυτών τα *CDC20* και *ASPM* φάνηκε να αποτελούν δυνητικούς προγνωστικούς βιοδείκτες και δυνητικούς ανοσοθεραπευτικούς στόχους για τον KOK.

Ο Gao κ.α. διεξήγαγαν μία μελέτη στην οποία μετα-ανέλυσαν δεδομένα μικροσυστοιχιών από τέσσερις μελέτες που λήφθηκαν από τη βάση δεδομένων GEO και οι οποίες περιλάμβαναν συνολικά 304 δείγματα ιστών ουροδόχου κύστης [122]. Οι συγγραφείς εντόπισαν τα διαφορικά εκφραζόμενα γονίδια σε κάθε ένα από αυτά τα σύνολα δεδομένων και στη συνέχεια, εφαρμόζοντας τη μέθοδο RobustRankAggreg (RRA), βρέθηκε μία συναινετική λίστα 343 διαφορικά εκφραζόμενων γονιδίων. Τα γονίδια αυτά αξιοποιήθηκαν για την ανάλυση του δικτύου αλληλεπίδρασης πρωτεΐνης – πρωτεΐνης από την οποία προέκυψαν 10 κορυφαία γονίδια – κόμβοι τα οποία περιλάμβαναν τα *VEGFA*, *TOP2A*, *CCNB1*, *CDC20*, *AURKA*, *AURKB*, *ACTA2*, *UBE2C*, *CEP55* και *CCNB2*. Ακόμη, διεξήγαγαν ανάλυση επιβίωσης η οποία αποκάλυψε ότι τα επίπεδα έκφρασης των γονιδίων *ACTA2*, *CCNB1*, *CDC20* και *VEGFA* σχετίζονταν με την πρόγνωση των ασθενών με KOK. Οι συγγραφείς ισχυρίστηκαν ότι αυτά τα γονίδια – κλειδιά ενδέχεται να αποτελέσουν δυνητικούς θεραπευτικούς στόχους και βιοδείκτες για τη θεραπεία του KOK.

Τέλος, ο Shiv Verma κ.α., σε μία πρόσφατη μελέτη τους, αξιοποίησαν τρία σύνολα δεδομένων από το αποθετήριο GEO με σκοπό τη μελέτη σημαντικών γονιδίων κατά τη μετάβαση από μη μυοδιηθητικό σε μυοδιηθητικό KOK [123]. Αρχικά, προσδιόρισαν τα διαφορικά εκφραζόμενα γονίδια σε κάθε σύνολο δεδομένων με τη χρήση της online πλατφόρμας GEO2R και στη συνέχεια εντόπισαν τα κοινά γονίδια για τα παραπάνω σύνολα, καταλήγοντας σε 1516 διαφορικά εκφραζόμενα γονίδια μεταξύ των δειγμάτων μη μυοδιηθητικού και μυοδιηθητικού KOK. Συνολικά επτά γονίδια, συμπεριλαμβανομένων των *CDKN2A*, *CDC20*, *CTSV*, *FOXMI*, *MAGEA6*, *KRT23* και *SI00A9*, βρέθηκαν να κατέχουν ισχυρή προγνωστική αξία στον KOK. Έπειτα, οι συγγραφείς προσδιόρισαν ένα σύνολο τριών γονιδίων, τα γονίδια *CDKN2A*, *CTSV* και *FOXMI*, με υψηλή αποτελεσματικότητα στην πρόβλεψη της εξέλιξης του KOK. Συνοπτικά, οι συγγραφείς ισχυρίστηκαν ότι οι τρεις αυτοί βιοδείκτες θα μπορούσαν να προβλέψουν με ακρίβεια την εξέλιξη και την πρόγνωση του KOK.

1.10 Περιγραφή της παρούσας μελέτης

Το κίνητρο της παρούσας μελέτης ήταν ο εντοπισμός **βασικών γονιδίων – κόμβων** που μπορούν να αποτελέσουν δυνητικούς **διαγνωστικούς, προγνωστικούς και προβλεπτικούς βιοδείκτες** για τον KOK. Σε αυτή τη βάση, ο στόχος του πρώτου μέρους της ανάλυσης ήταν να επαναχρησιμοποιηθούν όλα τα διαθέσιμα **δεδομένα γονιδιακής έκφρασης** που βασίζονται σε **μικροσυστοιχίες** από το μεγαλύτερο και πιο έγκυρο παγκόσμιο αποθετήριο, με σκοπό να πραγματοποιηθεί μια **ενοποιητική μετα-ανάλυση** (integrative meta-analysis) προκειμένου να αξιολογηθούν οι μεταβολές της γονιδιακής έκφρασης στον ιστό της ουροδόχου κύστης και να εντοπιστούν τα γονίδια – «κλειδιά» στον KOK. Στο πλαίσιο αυτό, ενσωματώθηκαν επίσης στην ανάλυση δεδομένα γονιδιακής έκφρασης που προέρχονταν από δείγματα ούρων και αίματος ασθενών με KOK, προκειμένου να διερευνηθεί περαιτέρω η πιθανή μεταβαλλόμενη έκφραση των προσδιορισθέντων γονιδίων – «κλειδιών» σε αυτά τα βιολογικά υγρά. Στη συνέχεια, πραγματοποιήθηκε ανάλυση επιβίωσης προκειμένου να αξιολογηθεί η **προγνωστική αξία** των γονιδίων – κλειδιών και κατασκευάστηκε ένα προγνωστικό μοντέλο για τον KOK. Η απόδοση του μοντέλου που αναπτύχθηκε επικυρώθηκε με τη δοκιμή του σε δύο ανεξάρτητα σύνολα δεδομένων

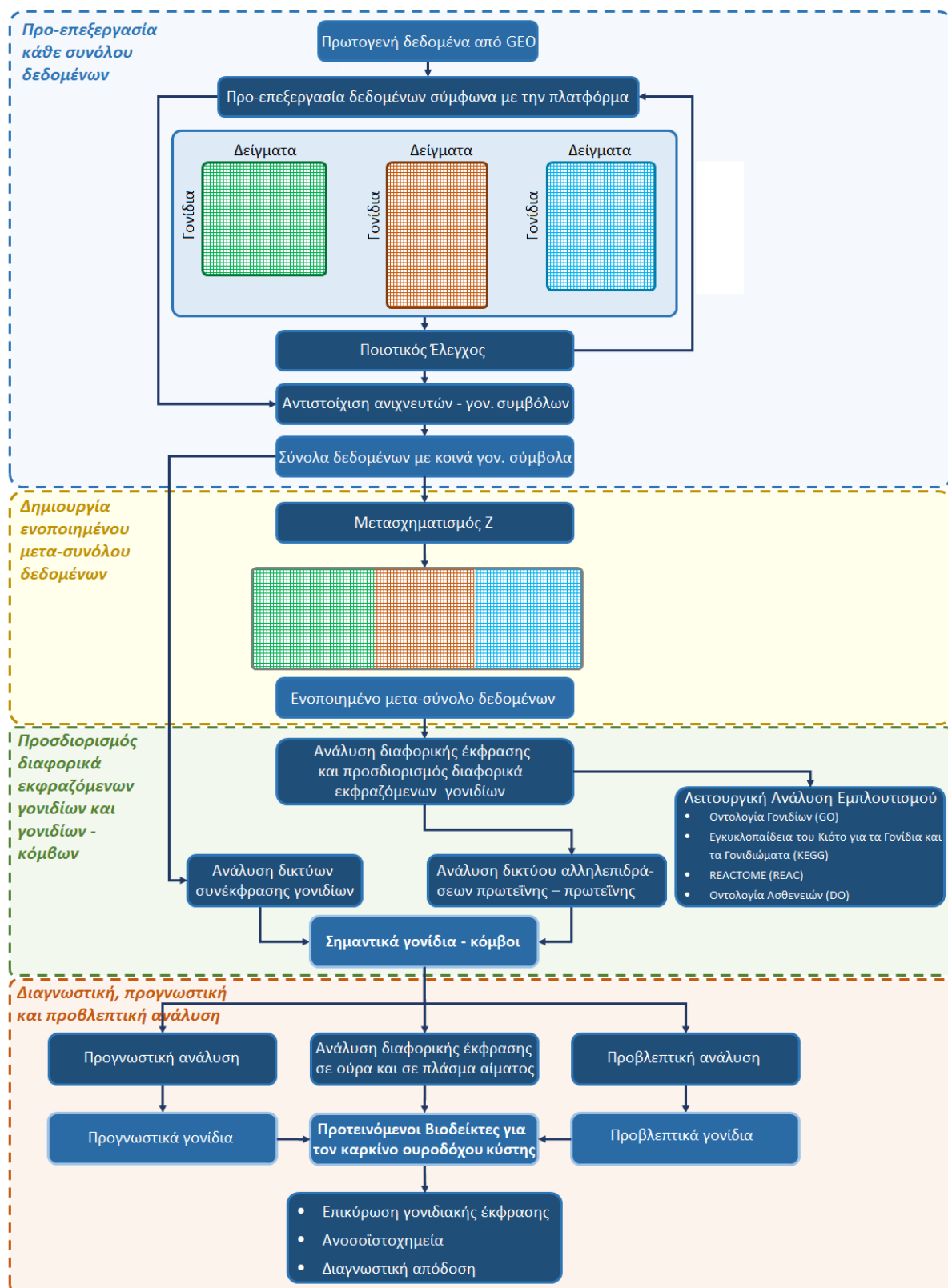
καθώς και μιας διαδικτυακής βιοπληροφορικής πλατφόρμας. Επιπλέον, στην ανάλυση συμπεριλήφθηκαν δεδομένα από ασθενείς με μυοδιαθητικό ΚΟΚ που έλαβαν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη με σκοπό να διερευνηθεί η **προβλεπτική** αξία των γονιδίων – κλειδιών όσον αφορά την **ανταπόκριση των ασθενών στη θεραπεία**. Βάσει αυτής της ανάλυσης, κατασκευάστηκε ένα προβλεπτικό μοντέλο για τον ΚΟΚ, το οποίο δοκιμάστηκε σε δύο εξωτερικά σύνολα δεδομένων. Τέλος, προσδιορίστηκε ένα **πάνελ εννέα γονιδίων** με δυνητικούς σημαντικούς βιοδείκτες για τον ΚΟΚ και αξιοποιήθηκαν αλγόριθμοι μηχανικής μάθησης [124] προκειμένου να γίνει μεγαλύτερη εμβάθυνση στα ερευνητικά αποτελέσματα και να επικυρωθεί η **διαγνωστική απόδοση** του συγκεκριμένου γονιδιακού πάνελ. Απώτερη πεποίθηση του συγγραφέα της παρούσας μελέτης είναι ότι αυτοί οι βιοδείκτες θα μπορούσαν να αποτελέσουν υποσχόμενους διαγνωστικούς και προγνωστικούς στόχους για τη διαχείριση και τη θεραπεία του ΚΟΚ.

Η δομή της υπόλοιπης διδακτορικής διατριβής διαρθρώνεται ως εξής: Στη δεύτερη ενότητα αναλύονται η διαδικασία με την οποία προσδιορίστηκαν τα σύνολα δεδομένων προς ανάλυση, οι μέθοδοι και αλγόριθμοι που χρησιμοποιήθηκαν, καθώς και η πλήρης ροή εργασίας της παρούσας μελέτης. Στην τρίτη ενότητα, παρατίθενται όλα τα αποτελέσματα της ανάλυσης, τα οποία συνοδεύονται από τους αντίστοιχους πίνακες, διαγράμματα, εικόνες και αριθμητικά αποτελέσματα. Στη συνέχεια, η τέταρτη ενότητα περιλαμβάνει τη συζήτηση σχετικά τα αποτελέσματα της έρευνας, τη σύγκριση με αποτελέσματα της βιβλιογραφίας και τον εντοπισμό πιθανών περιορισμών. Τέλος, στην πέμπτη ενότητα περιλαμβάνονται τα συμπεράσματα της παρούσας μελέτης και ακολούθως, στην έκτη και τελευταία ενότητα, παρατίθεται η πλήρης λίστα των βιβλιογραφικών αναφορών.

2. Υλικό και Μέθοδοι

2.1 Σχεδιασμός Μελέτης και Ροή Εργασίας

Ο συνολικός σχεδιασμός της μελέτης και η ροή εργασίας της παρούσας ενοποιητικής βιοπληροφορικής μετα-ανάλυσης παρουσιάζονται στην Εικόνα 29. Στο πρώτο στάδιο της έρευνας, εντοπίζονται και ενοποιούνται πολλαπλά σύνολα δεδομένων μικροσυστοιχιών και δημιουργείται ένα ενοποιημένο μετα-σύνολο δεδομένων (merged meta-dataset). Το σύνολο αυτό αξιοποιείται για τον προσδιορισμό των **διαφορικά εκφραζόμενων γονιδίων** μεταξύ των υγιών και των καρκινικών δειγμάτων της ουροδόχου κύστης. Επίσης, στη μελέτη προσδιορίζονται τα **γονίδια – κόμβοι** (hub genes) μέσω της ανάλυσης του δικτύου αλληλεπιδράσεων πρωτεΐνης – πρωτεΐνης (protein – protein interaction network analysis) και της ανάλυσης δικτύων συνέκφρασης γονιδίων (weighted gene co-expression network analysis – WGCNA). Επιπρόσθετα, πραγματοποιήθηκε **λειτουργική ανάλυση εμπλουτισμού της γονιδιακής έκφρασης** (functional enrichment analysis) των διαφορικά εκφραζόμενων γονιδίων αξιοποιώντας την Οντολογία Γονιδίων (Gene Ontology – GO), την Εγκυκλοπαίδεια του Κιότο για τα Γονίδια και τα Γονιδιώματα (Kyoto Encyclopedia of Genes and Genomes – KEGG), τη βάση γνώσης REACTOME (Reactome Pathway Database – REAC), καθώς και την Οντολογία Ασθενειών (Disease Ontology – DO). Εν συνεχεία, τα προσδιορισθέντα γονίδια – κόμβοι αξιολογήθηκαν για τη **διαγνωστική, προγνωστική και προβλεπτική** τους – ως προς την απόκριση στη θεραπεία – αξία. Στο πλαίσιο αυτό, αξιοποιήθηκαν επίσης δεδομένα γονιδιακής έκφρασης από **ούρα** και **πλάσμα αίματος** ασθενών με ΚΟΚ, δεδομένα επιβίωσης ασθενών με ΚΟΚ διαφόρων σταδίων, καθώς και δεδομένα επιβίωσης και υποτροπής ασθενών με μυοδιηθητικό ΚΟΚ, οι οποίοι έλαβαν νεοεπικουρική χημειοθεραπεία πριν από τη χειρουργική επέμβαση. Τα γονίδια – κόμβοι τα οποία εντοπίστηκαν να εκφράζονται διαφορικά στα ούρα ή στο αίμα των ασθενών με ΚΟΚ και ταυτόχρονα βρέθηκε πως κατέχουν προβλεπτική ή προγνωστική αξία θεωρήθηκαν ως πιθανοί **βιοδείκτες** για τον ΚΟΚ. Τέλος, η γονιδιακή έκφραση των συγκεκριμένων βιοδεικτών ταυτοποιήθηκε, μέσω έγκυρων βιοπληροφορικών πλατφορμών, καθώς και πειραμάτων ανοσοϊστοχημείας (immunohistochemistry), μέσω του Άτλαντα των Ανθρώπινων Πρωτεϊνών (Human Protein Atlas – HPA), και, τέλος, το πάνελ των εννέα προτεινόμενων βιοδεικτών ελέγχθηκε για την διαγνωστική του αξία σε πολλαπλά σύνολα δεδομένων με τη χρήση τεχνικών μηχανικής μάθησης.



Εικόνα 29: Ο συνολικός σχεδιασμός της μελέτης και η ροή εργασίας της ενοποιητικής μετα-ανάλυσης της παρούσας εργασίας.

2.2 Πηγή Δεδομένων, Συστηματική Αναζήτηση και Επιλογή Συνόλων

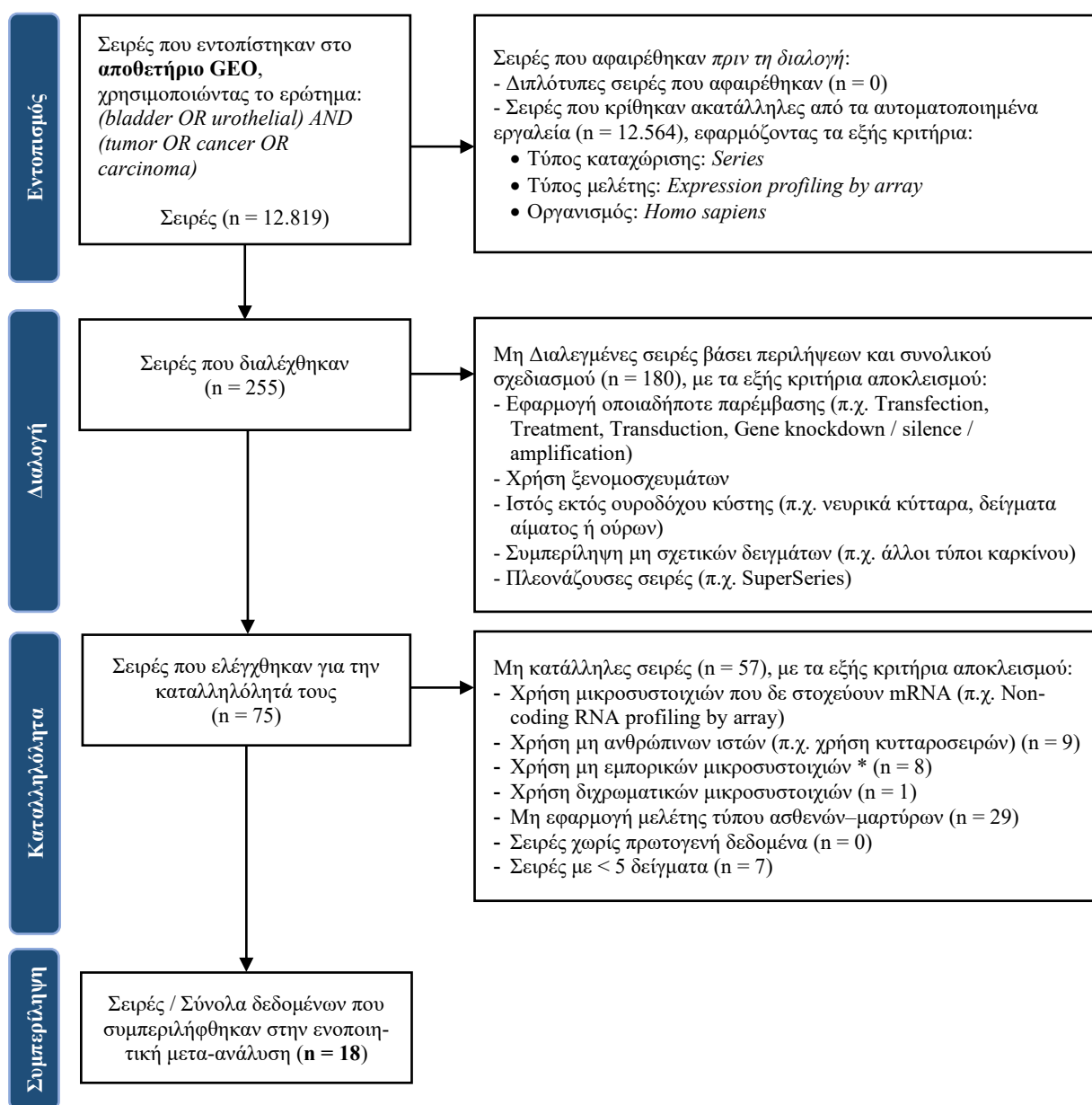
Δεδομένων

Τα σύνολα δεδομένων μικροσυστοιχιών που αναλύθηκαν στην παρούσα μελέτη είναι ελεύθερα διαθέσιμα και αντλήθηκαν από το Gene Expression Omnibus (GEO) του National Center for Biotechnology Information (NCBI) των ΗΠΑ, το οποίο είναι το μεγαλύτερο δημόσιο αποθετήριο ειδικά σχεδιασμένο για την αρχειοθέτηση και τη διανομή δεδομένων μικροσυστοιχιών, δεδομένων αλληλούχισης επόμενης γενιάς, καθώς και άλλων λειτουργικών γονιδιακών δεδομένων υψηλής απόδοσης [125].

Με σκοπό τον εντοπισμό των κατάλληλων συνόλων δεδομένων για την παρούσα μελέτη πραγματοποιήθηκε μία συστηματική αναζήτηση στο αποθετήριο GEO χρησιμοποιώντας το ακόλουθο ερώτημα: (*bladder OR urothelial*) AND (*tumor OR cancer OR carcinoma*). Επίσης, εφαρμόστηκαν τα ακόλουθα φίλτρα: “Series”, “Expression profiling by array”, και “Homo sapiens” ως τύπος καταχώρησης, τύπος μελέτης και οργανισμός, αντίστοιχα. Συνολικά, ελήφθησαν 255 σύνολα δεδομένων από την αναζήτηση έως και τις 10 Ιανουαρίου 2022. Ένα σύνολο δεδομένων συμπεριλαμβάνονταν στην μετα-ανάλυση εάν πληρούνταν τα ακόλουθα βασικά κριτήρια συμπερίληψης: 1) εφαρμόστηκε πρωτόκολλο μελέτης ασθενών – μαρτύρων, 2) χρησιμοποιήθηκαν εμπορικές πλατφόρμες μικροσυστοιχιών μονοχρωματικού καναλιού, 3) αναλύθηκαν δείγματα από ανθρώπινους ιστούς προερχόμενους από την κατώτερη ουροποιητική οδό (δηλ. από ουροδόχο κύστη ή ουρήθρα) και 4) χρησιμοποιήθηκαν μονάχα ανέπαφα δείγματα (δηλαδή δείγματα τα οποία δεν είχαν επιμολυνθεί με οποιαδήποτε ξένη ή φαρμακευτική ουσία). Η μετα-ανάλυση πραγματοποιήθηκε σύμφωνα με τις οδηγίες του πρωτοκόλλου Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) που δημοσιεύτηκε το 2020 [126]. Τα αναλυτικά κριτήρια για τη διαδικασία επιλογής και τα επιμέρους στάδια της αναζήτησης περιγράφονται στην Εικόνα 30. Κάθε σύνολο δεδομένων ελέγχθηκε ανεξάρτητα από δύο ερευνητές όσον αφορά την ικανοποίηση των ανωτέρω κριτηρίων συμπερίληψης.

Οι μελέτες που συμπεριλήφθηκαν σε αυτήν τη μετα-ανάλυση εξέταζαν ανθρώπινους υγιείς και καρκινικούς ιστούς από την ουροδόχο κύστη και πραγματοποιήθηκαν χρησιμοποιώντας εμπορικές πλατφόρμες μικροσυστοιχιών (Affymetrix, Illumina and Agilent) για λόγους επαναληψιμότητας, αναπαραγωγιμότητας και συνέπειας [127]. Επίσης, η επιλογή μονάχα των μονοχρωματικών μικροσυστοιχιών συνέβαλε στη διεξαγωγή μιας αποδοτικότερης ενοποιητικής μετα-ανάλυσης πρώιμου σταδίου («early stage» approach) [128], συμβάλλοντας στην αποφυγή της πολυπλοκότητας που θα είχε η συμπερίληψη δεδομένων από διχρωματικές μικροσυστοιχίες.

Για κάποια από τα σύνολα δεδομένων, ο αρχικός αριθμός δειγμάτων που αυτά περιείχαν ήταν μεγαλύτερος από τον αριθμό των δειγμάτων που τελικά συμπεριλήφθηκαν στην παρούσα μελέτη και οι λόγοι για την εξαίρεση κάποιων δειγμάτων εξηγούνται στη συνέχεια: Τα 24 δείγματα της σειράς GSE37815 συμπεριλαμβάνονταν όλα στην σειρά GSE13507 και, έτσι, αφαιρέθηκαν από την τελευταία. Επίσης, η σειρά GSE38264 περιείχε 13 δείγματα από τον οργανισμό *mus musculus*, και, έτσι, αυτά εξαιρέθηκαν από τη συγκεκριμένη σειρά. Τέλος, η σειρά GSE40355 περιλαμβάνει 24 δείγματα τα οποία ελήφθησαν χρησιμοποιώντας μία μικροσυστοιχία μη κωδικοποιημένου DNA (Agilent Human miRNA Microarray V2) και, έτσι, τα δείγματα αυτά αφαιρέθηκαν από τη συγκεκριμένη σειρά.



Εικόνα 30: Διάγραμμα ροής της συστηματικής αναζήτησης βάσει του πρωτοκόλλου Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020)

Μια πλατφόρμα μικροσυστοιχίας η οποία είναι πανομοιότυπη με μια εμπορική πλατφόρμα και στην οποία χρησιμοποιείται ένα προσαρμοσμένο, αναδιαμορφωμένο περιβάλλον CDF για την εξαγωγή δεδομένων, θεωρήθηκε ως εμπορική πλατφόρμα.

2.3 Προ-επεξεργασία με βάση τον κατασκευαστή της πλατφόρμας

Μετά από τον εντοπισμό των κατάλληλων συνόλων δεδομένων, τα πρωτογενή δεδομένα (raw data) μικροσυστοιχιών όλων των πειραμάτων ελήφθησαν από το GEO. Έπειτα, διενεργήθηκε ξεχωριστή κανονικοποίηση σε όλα τα σύνολα δεδομένων με σκοπό τη διόρθωση των τεχνικών και περιβαλλοντικών επιδράσεων στα δεδομένα. Η διαδικασία αυτή επιτρέπει τα δείγματα από μία μελέτη να προσαρμοστούν σε μία παρόμοια κλίμακα. Η διαδικασία της κανονικοποίησης εφαρμόστηκε με βάση τον τύπο της πλατφόρμας με την οποία διενεργήθηκε η εκάστοτε μελέτη και περιγράφεται αναλυτικά παρακάτω.

Η κανονικοποίηση των δεδομένων που πραγματοποιήθηκαν με πλατφόρμα της Affymetrix διενεργήθηκε με χρήση της μεθόδου Robust Multiarray Analysis (RMA), όπως αυτή υλοποιείται στα

R/Bioconductor πακέτα *affy* (version 1.72) [129], για τους τύπους πλατφορμών *HG-U133A* και *HG-U133 Plus 2*, και *oligo* (version 1.58) [130], για τους τύπους πλατφορμών *HuGene-1.0 ST*, *HuEx-1.0 ST* και *HTA-2.0*. Ο αλγόριθμος αυτός πραγματοποιεί διόρθωση υποβάθρου, $\log_2$ μετασχηματισμό, κανονικοποίηση ποσοστημορίων, και σύνοψη όλων των ομάδων ανιχνευτών σε μία μοναδική τιμή έκφρασης για κάθε γονίδιο. Η μέθοδος αυτή έχει αποδειχθεί ότι αποδίδει ιδιαίτερα καλά ως προς την ευαισθησία της στις βιολογικές διακυμάνσεις αλλά και ως προς τη βελτίωση της συγκρισιμότητας μεταξύ των πλατφορμών [131, 132].

Η κανονικοποίηση των συνόλων δεδομένων που πραγματοποιήθηκαν με πλατφόρμα της Illumina διενεργήθηκε με τη χρήση των συναρτήσεων *read.ilmn* και *neqc*, όπως αυτές υλοποιούνται στο R/Bioconductor πακέτο *limma* (version 3.50) [133]. Οι συναρτήσεις αυτές διαβάζουν τα δεδομένα έκφρασης και εκτελούν διόρθωση υποβάθρου, $\log_2$ μετασχηματισμό, κανονικοποίηση ποσοστημορίων, αξιοποιώντας τους αρνητικούς και θετικούς ανιχνευτές ελέγχου για την κανονικοποίηση και μόνο τους αρνητικούς ανιχνευτές ελέγχου για τη διόρθωση του υποβάθρου. Στις σειρές GSE13507 και GSE37815 οι ανιχνευτές ελέγχου είχαν ήδη αφαιρεθεί από τα πρωτογενή δεδομένα. Έτσι, αξιοποιήθηκαν οι συναρτήσεις *read.table* και *normalizeBetweenArrays* του ίδιου πακέτου της R έτσι ώστε να διαβαστούν σωστά τα δεδομένα έκφρασης και στην συνέχεια να υλοποιηθούν τα προηγούμενα βήματα.

Τέλος, η κανονικοποίηση των συνόλων δεδομένων που πραγματοποιήθηκαν με πλατφόρμα της Agilent διενεργήθηκε επίσης με τη χρήση του R/Bioconductor πακέτου *limma* και αντίστοιχων μεθόδων που αυτό περιλαμβάνει για τις πλατφόρμες του συγκεκριμένου κατασκευαστή. Τα δεδομένα έκφρασης διαβάστηκαν και εκτελέστηκε διόρθωση υποβάθρου, $\log_2$ μετασχηματισμός, και κανονικοποίηση ποσοστημορίων, αξιοποιώντας τις συναρτήσεις *read.maimages*, *backgroundCorrect* και *normalizeBetweenArrays* του R πακέτου *limma*. Στην προκειμένη περίπτωση ακολουθήθηκαν οι συστάσεις που εφαρμόζονται στις δίχρωμες πλατφόρμες της Agilent, καθώς η διαδικασία μπορεί να εφαρμοστεί εξίσου αποδοτικά στις μονόχρωμες πλατφόρμες και να αποδώσει παρόμοιες τιμές συστηματικών σφαλμάτων [134].

2.4 Ποιοτικός Έλεγχος Δειγμάτων

Για όλα τα σύνολα δεδομένων, εφαρμόστηκε ένα κοινό αναλυτικό πλαίσιο ελέγχου ποιότητας, το οποίο ήταν ανεξάρτητο της πλατφόρμας διενέργειας των πειραμάτων για λόγους αμεροληψίας και συνέπειας. Μετά την αντίστοιχη προ-επεξεργασία για κάθε τύπο πλατφόρμας, κάθε κανονικοποιημένο σύνολο δεδομένων υποβλήθηκε σε ποιοτικό έλεγχο, εφαρμόζοντας τη στρατηγική αφαίρεσης ακραίων τιμών. Οι ποιοτικοί έλεγχοι διεξήχθησαν αξιοποιώντας το R/Bioconductor πακέτο *arrayQualityMetrics* (version 3.50) [135] και επιθεωρώντας τις τρεις απεικονιστικές τεχνικές που περιλαμβάνονται στην τελική αναφορά του *arrayQualityMetrics*: τον θερμικό χάρτη (heatmap) της απόστασης μεταξύ των μικροσυστοιχιών, το θηκόγραμμα (boxplot) των λογαριθμικών λόγων, καθώς και το διάγραμμα MA (MA-plot), το οποίο αποτυπώνει τον λογάριθμο του λόγου των εντάσεων (M) έναντι των μέσων λογαριθμημένων εντάσεων (A). Αυτή η στρατηγική έχει αποδειχθεί ότι βελτιώνει την αποτελεσματικότητα της μετα-ανάλυσης και αυξάνει την στατιστική ισχύ στον προσδιορισμό των διαφορικά εκφραζόμενων γονιδίων [136]. Τα δείγματα που ταξινομήθηκαν ως ακραίες τιμές σε τουλάχιστον δύο από τις τρεις τεχνικές, κατά τη διαδικασία ποιοτικού ελέγχου, αφαιρέθηκαν από το αντίστοιχο σύνολο δεδομένων τους. Στη συνέχεια, τα πρωτογενή δεδομένα, χωρίς τα δείγματα με τις ακραίες τιμές, κανονικοποιήθηκαν εκ νέου, ακολουθώντας τη διαδικασία που περιγράφηκε στην προηγούμενη ενότητα, και τελικά χρησιμοποιήθηκαν για την καθοδική ανάλυση (downstream analysis).

2.5 Αντιστοίχιση Ανιχνευτών – Γονιδίων

Όλοι οι ανιχνευτές των μικροσυστοιχιών αντιστοιχίστηκαν στο γονιδιακό επίπεδο, χρησιμοποιώντας τα γονιδιακά σύμβολα ως κοινό αναγνωριστικό για όλες τις πλατφόρμες. Τα επίσημα γονιδιακά σύμβολα έχουν εγκριθεί από την HUGO Gene Nomenclature Committee (HGNC) [137]. Η χρήση της εγκεκριμένης ονοματολογίας από την HGNC συνίσταται να χρησιμοποιείται στις επιστημονικές μελέτες καθώς αποτελεί μία ονοματολογία η οποία ενημερώνεται και επιμελείται διαρκώς και η χρήση της έχει αποδειχθεί ότι ενισχύει την ακρίβεια στον επιστημονικό και δημόσιο διάλογο [138]. Εάν περισσότεροι από ένας ανιχνευτές αντιστοιχούν στο ίδιο γονιδιακό σύμβολο, η τελική τιμή έκφρασης αυτού του γονιδίου υπολογίστηκε ως η μέση τιμή έκφρασης όλων των ανιχνευτών. Οι ανιχνευτές που αντιστοιχίστηκαν σε περισσότερα από ένα σύμβολα γονιδίων ή σε κανένα σύμβολο αποκλείστηκαν από την παρούσα μελέτη.

Η αντιστοίχιση μεταξύ των ανιχνευτών και των γονιδιακών συμβόλων πραγματοποιήθηκε με τη χρήση συγκεκριμένων πακέτων αντιστοίχισης για την εκάστοτε πλατφόρμα τα οποία παρέχονται στο αποθετήριο Bioconductor. Πιο συγκεκριμένα, η μετατροπή μεταξύ ανιχνευτών και γονιδιακών συμβόλων γονιδίων πραγματοποιήθηκε με τη χρήση των R/Bioconductor πακέτων *AnnotationDbi* (version 1.56.2) and *org.Hs.eg.db* (version 3.14). Αναλυτικότερα, οι ανιχνευτές των συνόλων δεδομένων αντιστοιχίστηκαν χρησιμοποιώντας τα R πακέτα *hgu133a.db*, *hgu133plus2.db* (version 3.13), *hugene10sttranscriptcluster.db*, *huex10sttranscriptcluster.db*, *hta20transcriptcluster.db* (version 8.8), *illuminaHumanv2.db*, *illuminaHumanv3.db*, και *illuminaHumanv4.db* (version 1.26), με βάση την εκάστοτε πλατφόρμα. Για τις τρεις σειρές, GSE21142, GSE24152 και GSE42089, χρησιμοποιήθηκε το εκάστοτε προσαρμοσμένο αρχείο αντιστοίχισης Brainarray Chip Description File (CDF), το οποίο ουσιαστικά παρέχει μία ειδική τροποποιημένη αντιστοίχιση για τους ανιχνευτές της πλατφόρμας. Για τις πλατφόρμες της Agilent, η αντιστοίχιση ανιχνευτών και γονιδιακών συμβόλων πραγματοποιήθηκε χρησιμοποιώντας το R/Bioconductor πακέτο *biomaRt* (version 2.50.3) με σκοπό την πρόσβαση στην αντιστοίχιση *Ensembl* [139], καθώς αυτή δεν παρέχονταν από το Bioconductor. Η επιλογή των ανωτέρω πόρων γονιδιακής αντιστοίχισης βασίστηκε στις συνεχείς τους ενημερώσεις, τη συνέπεια και την αξιοπιστία τους [139-141]. Είναι πολύ σημαντικό να χρησιμοποιούνται ονοματολογίες και αντιστοιχίσεις μεταξύ ανιχνευτών και γονιδίων οι οποίες είναι ενημερωμένες και αξιόπιστες, έτσι ώστε οι βιολογικές ερμηνείες των αποτελεσμάτων να μπορούν να γίνουν με ακρίβεια κατά τη διάρκεια της καθοδικής ανάλυσης.

2.6 Φαινόμενα Παρτίδας και Διατεχνολογική Κανονικοποίηση

Τα επίπεδα γονιδιακής έκφρασης μπορεί να ποικίλλουν λόγω βιολογικών παραγόντων, οι οποίοι αποτελούν συνήθως τους παράγοντες ενδιαφέροντος, σε συνδυασμό με μη βιολογικούς, δηλαδή τεχνικών παραγόντων, που εξαρτώνται από το χρόνο και τον τόπο διεξαγωγής των πειραμάτων. Αυτοί οι παράγοντες απόκλισης, οι οποίοι είναι μη σχετικοί με τις πραγματικές διαφοροποιήσεις εντός ενός δείγματος αλλά και μεταξύ των δειγμάτων, αποτελούν ουσιαστικά συστηματικά σφάλματα, είναι σχεδόν πάντα αναπόφευκτοι και ονομάζονται συνοπτικά ως «φαινόμενα παρτίδας» (batch effects) [142]. Λόγω αυτών των φαινομένων, η ενοποίηση δεδομένων από διάφορα πειράματα μικροσυστοιχιών γονιδιακής έκφρασης, η οποία πρακτικά διεξάγεται σε αυτή τη μελέτη, μετατρέπεται σε μια πολύπλοκη διαδικασία που απαιτεί προσεκτικό χειρισμό [143].

Οι πληροφορίες για τα φαινόμενα παρτίδας ή την ημερομηνία των πειραμάτων δεν ήταν διαθέσιμες για την πλειοψηφία των 18 συνόλων δεδομένων της ενοποιητικής μετα-ανάλυσης της παρούσας μελέτης. Ως εκ τούτου, η εφαρμογή μιας μεθόδου που να διορθώνει τις τιμές της γονιδιακής έκφρασης με βάση γνωστά φαινόμενα παρτίδας είναι αδύνατη. Προκειμένου να πραγματοποιηθεί η ανίχνευση των

φαινομένων παρτίδας στα πειράματα, διεξήχθη οπτική επισκόπηση της μειωμένων διαστάσεων αναπαράστασης των δεδομένων, χρησιμοποιώντας την ανάλυση κύριων συνιστωσών (Principal Component Analysis – PCA). Είναι σημαντικό να αναφερθεί ότι, λόγω της ανίχνευσης ενός πολύ ισχυρού φαινομένου παρτίδας, τα δείγματα από τη σειρά GSE13507 διαιρέθηκαν σε δύο υποομάδες, GSE13507A και GSE13507B, αντίστοιχα. Αυτές οι δύο υποομάδες θεωρήθηκαν ως ξεχωριστά σύνολα δεδομένων κατά την παρούσα ενοποιητική μετα-ανάλυση.

Ο μετασχηματισμός Z , ή αλλιώς τυποποίηση (standardization), είναι αυτό που εφαρμόστηκε στα δεδομένα γονιδιακής έκφρασης, χρησιμοποιώντας τη συνάρτηση *scale* του R πακέτου *stats*. Η εφαρμογή αυτής της κλασικής μεθόδου κανονικοποίησης αποτελεί μια απλή προσέγγιση τυποποίησης δεδομένων για ένα ευρύ φάσμα πειραμάτων και επιτρέπει τη σύγκριση των δεδομένων μικροσυστοιχιών ανεξάρτητα από τις πρωτογενείς εντάσεις υβριδοποίησης [144]. Επιπλέον, ο μετασχηματισμός Z είναι απλός, έχει χαμηλή πολυπλοκότητα ως προς το χρόνο και την απαιτούμενη μνήμη, δεν απαιτεί καμία υπόθεση για την κατανομή των δεδομένων και έχει εφαρμοστεί με επιτυχία σε προηγούμενες μελέτες γονιδιακής έκφρασης, εμφανίζοντας υψηλή απόδοση [145-147]. Ο μετασχηματισμός Z εφαρμόστηκε σε όλα τα δείγματα αφαιρώντας τον μέσο όρο των εκφράσεων κάθε δείγματος και διαιρώντας με την τυπική του απόκλιση (standard deviation – SD) των εκφράσεων, σύμφωνα με τον τύπο:

$$Z_{score} = \frac{\text{ένταση } G - \text{μέση ένταση } G_1 \dots G_n}{\text{τυπική απόκλιση } G_1 \dots G_n}$$

όπου G είναι οποιοδήποτε γονίδιο και $G_1 \dots G_n$ είναι η συνολική τιμή των εκφράσεων όλων των γονιδίων για ένα δείγμα.

Μετά από αυτή την απλή ομογενοποίηση των δεδομένων, δεν επιχειρήθηκε περαιτέρω αφαίρεση ή διόρθωση των φαινομένων παρτίδας, καθώς αυτή θα μπορούσε να εισάγει συστηματικά σφάλματα στη σύγκριση των δύο ομάδων δειγμάτων, ειδικά για την ανάλυσή μας, όπου ο πειραματικός σχεδιασμός φαινομένων παρτίδας – φαινοτυπικών ομάδων δεν είναι ισορροπημένος. Αντίθετα, από τη διεθνή βιβλιογραφία συνιστάται, όταν είναι δυνατόν, τα φαινόμενα παρτίδας να ενσωματώνονται στην καθοδική ανάλυση [148]. Επομένως, κατά τη μελέτη των διαφορικά εκφραζόμενων γονιδίων, το σύνολο δεδομένων στο οποίο ανήκει κάθε δείγμα συμπεριλήφθηκε ως συμμεταβλητή στη στατιστική ανάλυση.

2.7 Ανάλυση Διαφορικής Έκφρασης Γονιδίων

Στην παρούσα ενοποιητική μετα-ανάλυση, ακολουθήθηκε μια μέθοδος ενοποίησης δεδομένων «πρώιμου σταδίου» [128]. Έτσι, δημιουργήθηκε ένα ενοποιημένο μετα-σύνολο δεδομένων μικροσυστοιχιών ενώνοντας τα Z -μετασχηματισμένα δεδομένα και αντιστοιχίζοντας τα 8.201 κοινά γονιδιακά σύμβολα μεταξύ των συνόλων. Αυτό το μετα-σύνολο δεδομένων περιείχε συνολικά 606 δείγματα, ενοποιώντας 410 δείγματα KOK και 196 υγιή δείγματα, από 19 διαφορετικά σύνολα δεδομένων.

Τα διαφορικά εκφραζόμενα γονίδια (ΔΕΓ) μεταξύ των δειγμάτων KOK και των υγιών δειγμάτων εντοπίστηκαν χρησιμοποιώντας το R/Bioconductor πακέτο *limma* (version 3.50) [133], με το σύνολο δεδομένων στο οποίο ανήκει κάθε δείγμα να συμπεριλαμβάνεται ως συμμεταβλητή (covariate) στο μοντέλο. Για την ανάλυση της στατιστικής σημαντικότητας, το βασικό στατιστικό τεστ που χρησιμοποιήθηκε ήταν το moderated t-statistic, το οποίο υπολογίστηκε για κάθε γονιδιακό σύμβολο μεταξύ των δειγμάτων KOK και των δειγμάτων ελέγχου. Προκειμένου να ελεγχθεί το ψευδές ποσοστό ανακάλυψης (False Discovery Rate – FDR) για τις πολλαπλές συγκρίσεις, η τιμή- p διορθώθηκε με βάση τη μέθοδο των Benjamini – Hochberg (BH).

Οι στατιστικές μέθοδοι που χρησιμοποιούνται για την αναγνώριση των διαφορικά εκφραζόμενων γονιδίων εξαρτώνται από τον καθορισμό αυθαίρετων ορίων για τις τιμές- p και τις πολλαπλάσιες

μεταβολές (Fold Change – FC) των τιμών έκφρασης των γονιδίων, γεγονός που μπορεί να αλλάξει σημαντικά τα εξαγόμενα αποτελέσματα και την ερμηνεία τους [149].

Στην παρούσα ανάλυση χρησιμοποιήθηκε ένα αυστηρό όριο για την διορθωμένη τιμή- $p = 0,01$. Το κατώτατο όριο για την τιμή $|\log_2 FC|$ επιλέγεται συνήθως στο διάστημα 1 με 2. Είναι προφανές πως για τις 11 διαφορετικές τιμές του $|\log_2 FC|$ από το 1 έως το 2 με βήμα 0,1, εξάγονται 11 διαφορετικά σύνολα διαφορεικά εκφραζόμενων γονιδίων. Είναι σημαντικό να αναφερθεί ότι η τιμή $\log_2 FC$ για κάθε γονίδιο προέκυψε από την τιμή FC που υπολογίζεται από το πακέτο *limma* διαιρεμένη με την τυπική απόκλιση των διαφορών των μέσων τιμών έκφρασης μεταξύ των δύο φαινοτυπικών ομάδων για όλα τα γονίδια, σύμφωνα με τον τύπο:

$$\log_2 FC_{\text{διορθωμένο}} = \frac{\log_2 FC}{\text{τυπική απόκλιση (διαφορές } Z - \text{τιμών } G_1 \dots G_n)}$$

όπου τυπική απόκλιση (διαφορές $Z - \text{τιμών } G_1 \dots G_n$) είναι η τυπική απόκλιση των διαφορών των μέσων τιμών έκφρασης μεταξύ των δύο φαινοτυπικών ομάδων (καρκινικών και υγιών δειγμάτων) για όλα τα γονίδια [144].

Προκειμένου να βρεθεί το βέλτιστο σύνολο διαφορετικά εκφραζόμενων γονιδίων που διακριτοποιεί τα δείγματα με τον αποδοτικότερα δυνατό τρόπο, δημιουργήθηκε ένα μοντέλο βασισμένο στον αλγόριθμο μηχανών διανυσμάτων υποστήριξης (Support Vector Machines – SVM), το οποίο χρησιμοποιούσε τα διαφορετικά εκφραζόμενα γονίδια ως χαρακτηριστικά (features). Ο SVM είναι ένας από τους πιο δημοφιλείς αλγόριθμους εποπτευόμενης μάθησης που έχουν επιδείξει σημαντική ικανότητα στο χειρισμό δεδομένων πολλαπλών διαστάσεων και υψηλή απόδοση στην ταξινόμηση καρκινικών και υγιών δειγμάτων από δεδομένα μικροσυστοιχιών [150]. Στην περίπτωση μας, κατασκευάστηκε ένα μοντέλο SVM για κάθε ένα από τα 11 σύνολα διαφορετικά εκφραζόμενων γονιδίων, που υλοποιήθηκαν χρησιμοποιώντας το R πακέτο *caret* (version 6.0). Για κάθε σύνολο χαρακτηριστικών, το ενοποιημένο μετα-σύνολο δεδομένων χωρίστηκε με τυχαίο τρόπο σε ένα σύνολο εκπαίδευσης και ένα σύνολο δοκιμής με αναλογία 90:10 και αυτή η διαδικασία επαναλήφθηκε 10 φορές, εφαρμόζοντας τη μέθοδο 10-πλης διασταυρούμενης επικύρωσης (10-fold cross validation). Επιλέχθηκε η τιμή του $|\log_2 FC|$, και κατ' επέκταση το σύνολο των διαφορετικά εκφραζόμενων γονιδίων, το οποίο είχε ως αποτέλεσμα την υψηλότερη εκτίμηση περιοχής κάτω από την καμπύλη (Area Under the Curve – AUC) λειτουργικού χαρακτηριστικού δέκτη (Receiver Operating Characteristic – ROC) του ταξινομητή. Η καμπύλη ROC είναι μια γραφική παράσταση πιθανοτήτων που παρουσιάζει τον πραγματικό θετικό ρυθμό (ευαισθησία) έναντι του ψευδώς θετικού ρυθμού (1 - ειδικότητα) για διάφορες τιμές κατωφλίου και αποτελεί ένα διάγραμμα αξιολόγησης των ταξινομητών για τα προβλήματα δυαδικής ταξινόμησης. Η τιμή AUC είναι μια μέτρηση που εκφράζει την ικανότητα του ταξινομητή να κάνει διάκριση μεταξύ των κλάσεων και την απόδοση της ταξινόμησης. Πρακτικά, αποτελεί την πιθανότητα ενός μοντέλου δυαδικού ταξινομητή να ταξινομήσει σωστά ένα δείγμα.

Μετά τον καθορισμό των κριτηρίων επιλογής (τιμών κατωφλίου), λήφθηκε το τελικό σύνολο των διαφορετικά εκφραζόμενων γονιδίων και σχεδιάστηκε η γραφική παράσταση ηφαιστείου (volcano plot) καθώς και ο θερμικός χάρτης (heatmap) για τα πρώτα 100 διαφορετικά εκφραζόμενα γονίδια, με τη χρήση των R πακέτων *ggplot2* (version 3.3.5) and *ComplexHeatmap* (version 2.10), αντίστοιχα.

2.8 Λειτουργική Ανάλυση Εμπλουτισμού της Γονιδιακής Έκφρασης

Προκειμένου να αναλυθούν και να οπτικοποιηθούν τα λειτουργικά προφίλ των προσδιορισθέντων διαφορετικά εκφραζόμενων γονιδίων, χρησιμοποιήθηκε το R/Bioconductor πακέτο *clusterProfiler* (version 4.2.2) [151]. Αναλυτικότερα, διεξήχθησαν λειτουργικές αναλύσεις εμπλουτισμού με βάση την Οντολογία Γονιδίων (Gene Ontology – GO), την Εγκυκλοπαίδεια του Κιότο για τα γονίδια και τα

γονιδιώματα (Kyoto Encyclopedia of Genes and Genomes – KEGG), τη βάση γνώσης REACTOME (Reactome Pathway Database – REAC), καθώς και την Οντολογία Ασθενειών (Disease Ontology – DO). Πριν από την εκτέλεση των αναλύσεων εμπλουτισμού, τα γονιδιακά σύμβολα αντιστοιχίστηκαν στη βάση δεδομένων Entrez Gene του NCBI με σκοπό να ανακτηθούν τα αναγνωριστικά γονιδίων Entrez, χρησιμοποιώντας το πακέτο αντιστοίχισης *org.Hs.eg.db* που παρέχεται από το Bioconductor. Για όλες τις ακόλουθες αναλύσεις, τα όρια αποκοπής για τον προσδιορισμό της στατιστικής σημαντικότητας ήταν τιμή- $p = 0,01$ και διορθωμένη τιμή- $p = 0,05$. Οι διορθώσεις για την τιμή- p , λόγω των πολλαπλών υποθέσεων, πραγματοποιήθηκαν χρησιμοποιώντας τη μέθοδο BH.

Η βάση γνώσης GO αποτελεί την πιο εκτεταμένη πηγή πληροφοριών σχετικά με τις λειτουργίες των γονιδίων [152]. Η ανάλυση εμπλουτισμού GO καλύπτει τρεις τομείς λειτουργιών: τα κυτταρικά συστατικά (cell components – CC), τη μοριακή λειτουργία (molecular function – MF) και τις βιολογικές διεργασίες (biological processes – BP), οι οποίες περιλαμβάνονται όλες στην παρούσα μελέτη. Οι όροι της βάσης GO για τα υπο- και υπερ-εκφραζόμενα γονίδια εμπλουτίστηκαν χρησιμοποιώντας τη συνάρτηση *enrichGO* του πακέτου *clusterProfiler*. Επίσης, οι όροι της βάσης GO εμπλουτίστηκαν με τη χρήση του πακέτου αντιστοίχισης *org.Hs.eg.db*, κατά την εκτέλεση της συνάρτησης *enrichGO*. Οι πλεονάζοντες εμπλουτισμένοι όροι της βάσης GO αφαιρέθηκαν χρησιμοποιώντας τη συνάρτηση *simplify*, εφαρμόζοντας όριο αποκοπής = 0,7 και τη μέθοδο Wang για τη μέτρηση της ομοιότητας [153]. Στη συνέχεια, οι πιο σημαντικοί εμπλουτισμένοι όροι αποτυπώθηκαν γραφικά με τη μορφή ενός ραβδογράμματος.

Η KEGG είναι μια ολοκληρωμένη βάση δεδομένων για την κατανόηση και τη συσχέτιση των λειτουργικών πληροφοριών ανώτερης τάξης των βιολογικών συστημάτων με γονιδιωματικές πληροφορίες [154]. Η ανάλυση εμπλουτισμού μονοπατιών KEGG πραγματοποιήθηκε χρησιμοποιώντας τη συνάρτηση *enrichKEGG* του πακέτου *clusterProfiler*. Εισήχθησαν τα αντίστοιχα αναγνωριστικά Entrez Gene Ids των διαφορεικά εκφραζόμενων γονιδίων και εφαρμόστηκαν τα προαναφερθέντα όρια αποκοπής. Η ανάλυση εμπλουτισμού αποτυπώθηκε γραφικά με τη μορφή ενός διαγράμματος κουκίδων (dot plot).

Η REAC είναι μια δημόσια βάση δεδομένων, ανοιχτού κώδικα, επιμελημένη και αξιολογημένη από ομότιμους, η οποία συσχετίζει με συστηματικό τρόπο τις ανθρώπινες πρωτεΐνες με τις μοριακές λειτουργίες τους [155]. Η ανάλυση μονοπατιών REAC πραγματοποιήθηκε με τη χρήση της έκδοσης 79 της πλατφόρμας και χρησιμοποιήθηκε το R/Bioconductor πακέτο *ReactomePA* [156] (version 1.38). Η ανάλυση μονοπατιών αποτυπώθηκε γραφικά με τη μορφή ενός διαγράμματος κουκκίδων και ενός χάρτη εμπλουτισμού των αποτελεσμάτων, με βάση τις ομοιότητες ανά ζεύγη των εμπλουτισμένων όρων [157]. Η DO αντιπροσωπεύει μια ολοκληρωμένη βάση γνώσεων με περισσότερες από 10.000 κληρονομικές, αναπτυξιακές και επίκτητες ανθρώπινες ασθένειες [158]. Για την ανάλυση εμπλουτισμού DO, χρησιμοποιήθηκε το R/Bioconductor πακέτο *DOSE* (version 3.20.1) [159]. Οι όροι DO με περισσότερα από 5 και με λιγότερο από 500 αντιστοιχούμενα γονίδια ήταν αυτοί που επιλέχθηκαν, και από αυτούς μόνο αυτοί που ικανοποιούσαν τα όρια αποκοπής θεωρήθηκαν ότι ήταν σημαντικά εμπλουτισμένοι. Τα αποτελέσματα παρουσιάστηκαν με τη μορφή διαγράμματος κουκίδων.

2.9 Ανάλυση του Δικτύου Αλληλεπιδράσεων Πρωτεΐνης – Πρωτεΐνης

Με βάση τα διαφορεικά εκφραζόμενα γονίδια που εντοπίστηκαν από την ενοποιητική μετα-ανάλυση των διαφορετικών συνόλων δεδομένων, διεξήχθη μια ανάλυση του δικτύου αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης – ΑΠΠ (protein-protein interaction – PPI), με σκοπό να διερευνηθεί περαιτέρω η πιθανή αλληλεπίδραση μεταξύ των γονιδίων αυτών και να εντοπιστούν τα βασικά γονίδια – κόμβοι μεταξύ αυτών. Η βάση δεδομένων Search Tool for Retrieval of InteractiNg Genes (STRING) (version 11.5) [160], η οποία εμπεριέχει τόσο γνωστές όσο και προβλεπόμενες αλληλεπιδράσεις μεταξύ των διαφόρων

πρωτεϊνών, χρησιμοποιήθηκε για την πρόβλεψη της λειτουργικής αλληλεπίδρασης μεταξύ των πρωτεϊνών που παράγονται από τα διαφορετικά εκφραζόμενα γονίδια. Το δίκτυο των πρωτεϊνών των 815 διαφορετικά εκφραζόμενων γονιδίων δημιουργήθηκε και οπτικοποιήθηκε μέσω της διεπαφής χρήστη του STRING, εφαρμόζοντας ως ελάχιστη απαιτούμενη βαθμολογία αλληλεπίδρασης την τιμή 0,4.

Στη συνέχεια, το δίκτυο πρωτεϊνών εισήχθη στο λογισμικό *Cytoscape* (version 3.9.1) [161]. Οι κόμβοι του δικτύου πρωτεϊνών ταξινομήθηκαν βάσει 10 μεθόδων τοπολογικής ανάλυσης δικτύων, με τη χρήση του πρόσθετου *cytoHubba* [162] του *Cytoscape*. Αυτές οι μέθοδοι περιλάμβαναν τρεις μεθόδους που βασίζονται σε τοπικό επίπεδο, οι οποίες είναι η Maximal Clique Centrality (MCC), η Maximum Neighborhood Component (MNC) και η Degree, καθώς και επτά μέθοδοι που βασίζονται σε καθολικό επίπεδο, οι οποίες είναι η Edge Percolated Component (EPC), η BottleNeck, η EcCentricity, η Closeness, η Radiality, η Betweenness and η Stress. Με βάση τις 10 ανωτέρω μεθόδους, εξάχθηκε μια τελική ταξινόμηση των γονιδίων – κόμβων του δικτύου πρωτεϊνών, από την ανάλυση του πρόσθετου *CytoHubba*, κάνοντας χρήση της ανθεκτικής συνθετικής μεθόδου (Robust Rank Aggregation – RRA) από το R πακέτο *RobustRankAggreg* (version 1.1). Στην τελική ταξινόμηση, διατηρήθηκαν μόνο τα γονίδια – κόμβοι με τιμή-p μικρότερη του 0,01.

Επίσης, το πρόσθετο *MCODE* (Molecular Complex Detection) [163] του *Cytoscape* χρησιμοποιήθηκε για τον προσδιορισμό συστάδων γονιδίων στο κατασκευασμένο δίκτυο πρωτεϊνών. Οι παράμετροι επιλογής ορίστηκαν ως εξής: *MCODE scores* ≥ 7 , *degree cut-off* = 2, *node score cut-off* = 0.2, *max depth* = 100, *k-score* = 2 and *haircut* = true. Έτσι, προέκυψε μία λίστα με συστάδες γονιδίων τα οποία ανήκαν σε ομάδες που πληρούσαν τα παραπάνω κριτήρια.

Τέλος, υπολογίστηκε η τομή των δύο ταξινομημένων λιστών γονιδίων, προκειμένου να ληφθεί μια τελική λίστα γονιδίων – κόμβων με βάση τα δύο πρόσθετα του *Cytoscape*. Η τελική αυτή λίστα γονιδίων περιέχει τα τελικά γονίδια – κόμβους, όπως προσδιορίστηκαν στη συγκεκριμένη μελέτη. Κατασκευάστηκε ακόμη το δίκτυο πρωτεϊνών της τελικής λίστας των γονιδίων – κόμβων για να μελετηθεί η αλληλεπίδραση μεταξύ τους.

2.10 Ανάλυση Δικτύων Συνέκφρασης Γονιδίων

Η σταθμισμένη ανάλυση δικτύων συνέκφρασης γονιδίων (Weighted Gene Co-expression Network Analysis – WGCNA) μπορεί να χρησιμοποιηθεί για την κατασκευή ενός σταθμισμένου δικτύου συνέκφρασης γονιδίων, τον ορισμό συστάδων (ενοτήτων) υψηλά συσχετιζόμενων γονιδίων, τη συσχέτιση των συστάδων με τα κλινικά χαρακτηριστικά και τον προσδιορισμό των γονιδίων – κόμβων της κάθε συστάδας [164]. Σε αυτή τη μελέτη, πραγματοποιήθηκε συναινετική WGCNA, χρησιμοποιώντας το R/Bioconductor πακέτο *WGCNA* [165], προκειμένου να βρεθούν οι σημαντικές συστάδες γονιδίων που σχετίζονται σε μεγάλο βαθμό με τον ΚΟΚ. Για το σκοπό αυτό, αξιοποιήθηκαν τα μεμονωμένα σύνολα δεδομένων που χρησιμοποιήθηκαν στην παρούσα ενοποιητική μετα-ανάλυση. Ωστόσο, χρησιμοποιήθηκαν μόνο οι σειρές που περιείχαν περισσότερα από 20 δείγματα, καθώς τα σύνολα δεδομένων με λιγότερα δείγματα θα ήταν απλώς πολύ θορυβώδη για να προκύψει κάποιο δίκτυο με ιδιαίτερο βιολογικό νόημα. Λόγω του γεγονότος ότι τα δεδομένα προέρχονταν από διαφορετικές παρτίδες, οι οποίες είναι άγνωστες, ελέγχθηκαν και διορθώθηκαν τα φαινόμενα παρτίδας χρησιμοποιώντας το R/Bioconductor πακέτο *sna* (version 3.42) [166].

Οι τιμές της γονιδιακής έκφρασης ομαδοποιήθηκαν ιεραρχικά για καθένα από τα σύνολα δεδομένων, προκειμένου να εντοπιστούν οι ακραίες τιμές και να αφαιρεθούν από την περαιτέρω ανάλυση. Για κάθε σύνολο δεδομένων, πραγματοποιήθηκε η ανάλυση τοπολογίας δικτύου. Το επιδιωκόμενο κατώφλι για τον δείκτη προσαρμογής της ελεύθερης κλίμακας τοπολογίας (R^2) ορίστηκε πάνω από 0,77 και η διάμεση συνδεσιμότητα ορίστηκε κάτω από 30. Μετά την επιλογή της δύναμης ελαφριού κατωφλιού, οι συναινετικές συστάδες των συνόλων δεδομένων υπολογίστηκαν χρησιμοποιώντας τη συνάρτηση

blockwiseConsensusModules. Οι παράμετροι ορίστηκαν ως *soft threshold power* = 7, *minModuleSize* = 30, *deepSplit* = 2 και *mergeCutHeight* = 0,25, για τη συγχώνευση των όμοιων συστάδων και όλα τα γονίδια υποβλήθηκαν σε επεξεργασία σε ένα μπλοκ. Στη συνέχεια, για κάθε σύνολο δεδομένων, ο πίνακας συσχέτισης μετατράπηκε σε πίνακα γειτνίασης, ο οποίος αναλύθηκε περαιτέρω για τον υπολογισμό του πίνακα τοπολογικής επικάλυψης (topological overlap matrix – TOM), χρησιμοποιώντας τον αλγόριθμο ομοιότητας TOM. Με βάση τη μέθοδο υπολογισμού τοπολογικής επικάλυψης ανομοιότητας, τα 8.201 γονίδια αντιστοιχίστηκαν σε διακριτές συστάδες γονιδίων οι οποίες υποδεικνύονταν με διάφορα χρώματα.

Συνεπώς, ο βαθμός συσχέτισης μεταξύ του ιδιογονιδίου κάθε συστάδας (ME) και του φαινοτύπου των δειγμάτων για κάθε σύνολο δεδομένων υπολογίστηκε με τον συντελεστή συσχέτισης Pearson, χρησιμοποιώντας τη συνάρτηση *cor* και τη συνάρτηση *corPvalueFisher* για τον υπολογισμό της αντίστοιχης τιμής-p. Προκειμένου να βρεθεί μια συναινετική συσχέτιση συστάδας – φαινοτυπικού χαρακτηριστικού, ορίστηκε ένα μέτρο της σχέσης συστάδας – φαινοτυπικού χαρακτηριστικού που συνόψιζε όλα τα σύνολα δεδομένων σε ένα μέτρο: για κάθε ζεύγος συστάδας – φαινοτυπικού χαρακτηριστικού εξάχθηκε μία συσχέτιση με βάση το πρόσημο των συσχετίσεων για κάθε σύνολο δεδομένων. Ειδικότερα, η χαμηλότερη απόλυτη τιμή συσχέτισης αποδόθηκε σε κάθε ζεύγος συστάδας – φαινοτυπικού χαρακτηριστικού, εάν οι συσχετίσεις είχαν το ίδιο πρόσημο για όλα τα σύνολα δεδομένων, και η μηδενική συσχέτιση αποδόθηκε για εκείνα τα ζεύγη που στα εκάστοτε σύνολα δεδομένων εμφάνιζαν αντίθετα πρόσημα. Ως εκ τούτου, μόνο οι συστάδες με σταθερό πρόσημο του συντελεστή συσχέτισης, είτε θετικό είτε αρνητικό, μεταξύ όλων των συνόλων δεδομένων θεωρήθηκαν ως βασικές συστάδες. Οι βασικές γονιδιακές συστάδες προσδιορίστηκαν με βάση τον συντελεστή συσχέτισης και τη στατιστική σημαντικότητα μεταξύ των τιμών ME των συστάδων και των φαινοτυπικών χαρακτηριστικών των δειγμάτων (καρκινικά και υγιή δείγματα).

Στη συνέχεια, με σκοπό να προσδιοριστούν ποια γονίδια στις βασικές συστάδες συσχετίζονται σε μεγάλο βαθμό με τα φαινοτυπικά χαρακτηριστικά, αξιολογήθηκε η συσχέτιση μεταξύ του φαινοτύπου των δειγμάτων, της γονιδιακής σημασίας (Gene Significance – GS) και της ιδιότητας μέλους συστάδας (Module Membership – MM). Η MM αντιπροσωπεύει τη συσχέτιση μεταξύ ME και του προφίλ της γονιδιακής έκφρασης και η GS αντιπροσωπεύει τη συσχέτιση μεταξύ γονιδίων και των φαινοτυπικών χαρακτηριστικών. Έτσι, για κάθε ME υπολογίστηκαν τα GS και MM σε κάθε σύνολο δεδομένων και, στη συνέχεια, συνδυάστηκαν οι Z-τιμές των συσχετίσεων από κάθε σύνολο δεδομένων για να προκύψει μία συναινετική μετα-Z-τιμή και η αντίστοιχη μετα-τιμή-p για κάθε συστάδα. Τα γονίδια με υψηλή Z-τιμή τόσο για τη MM όσο και για τη GS στην αντίστοιχη συστάδα συσχετίζονται σε μεγάλο βαθμό με τον φαινότυπο του καρκίνου. Ειδικότερα, γονίδια για τα οποία οι τιμές MM και GS ήταν στο άνω ή κάτω τεταρτημόριο όλων των γονιδίων της συστάδας προσδιορίστηκαν ως γονίδια – κόμβοι για τον KOK. Τέλος, συγκρίθηκαν αυτά τα γονίδια – κόμβοι με τα γονίδια – κόμβους που προέκυψαν από την ανάλυση του δικτύου πρωτεϊνών προκειμένου να προσδιοριστούν τα βασικά γονίδια – κόμβοι (key hub genes) αυτής της μελέτης που συνδέονταν σε μεγάλο βαθμό με τον KOK.

2.11 Διαφορική Έκφραση Γονιδίων στα Ούρα και στο Πλάσμα Αίματος

Με σκοπό να διερευνηθεί η πιθανή διαφορική έκφραση των βασικών γονιδίων – κόμβων, που προέκυψαν από την ανάλυση γονιδιακής έκφρασης των ιστών της ουροδόχου κύστης, σε άλλα βιολογικά υγρά, συμπεριλήφθηκαν δείγματα ούρων και πλάσματος αίματος στη συγκεκριμένη ενοποιητική μετα-ανάλυση. Τα νέα αυτά δείγματα προέρχονταν από ανάλυση γονιδιακής έκφρασης, με χρήση μικροσυστοιχιών. Ακόμη, αφορούσαν παρόμοιες μελέτες, τύπου ασθενών – μαρτύρων, όπως τα μετα-αναλυμένα σύνολα δεδομένων που εντοπίστηκαν αρχικά, και λήφθηκαν επίσης από το αποθετήριο GEO.

Για τον προσδιορισμό του προφίλ της γονιδιακής έκφρασης στα ούρα των ασθενών με ΚΟΚ, αξιοποιήθηκαν δύο σύνολα δεδομένων από το αποθετήριο GEO, και συγκεκριμένα οι σειρές GSE51843 και GSE68020. Το πρώτο σύνολο δεδομένων (GSE51843) περιλαμβάνει συνολικά 11 δείγματα εξωκυτταρικών κυστιδίων που περιέχουν mRNA, τα πέντε από τα οποία αποτελούν δείγματα ούρων από ασθενείς με ΚΟΚ και τα έξι δείγματα από μη ασθενείς, τα οποία υβριδοποιήθηκαν με χρήση της πλατφόρμας Illumina Human HT12 v4 BeadChip (GPL10558) [167]. Το δεύτερο σύνολο δεδομένων (GSE68020) περιέχει συνολικά 50 δείγματα ούρων, συμπεριλαμβανομένων ούρα από 30 ασθενείς με ουροθηλιακό καρκίνωμα υψηλού βαθμού και από 20 υγιείς μάρτυρες, τα οποία υβριδοποιήθηκαν σε τον ίδιο τύπο πλατφόρμας (GPL10558).

Τα πρωτογενή δεδομένα λήφθηκαν και για τα δύο σύνολα δεδομένων και πραγματοποιήθηκε προεπεξεργασία με βάση τη συγκεκριμένη πλατφόρμα, όπως περιγράφηκε στην Ενότητα 2.3 για τα άλλα σύνολα δεδομένων. Εφαρμόζοντας τη συνάρτηση *num.sv* από το R/Bioconductor πακέτο *sna*, δεν εντοπίστηκαν συγχυτικοί παράγοντες και, ως εκ τούτου, δεν χρειάστηκε διόρθωση των φαινομένων παρτίδας. Κάθε ένα από τα βασικά γονίδια – κόμβοι ελέγχθηκε για τη στατιστικά σημαντική διαφορά του μεταξύ των δειγμάτων ΚΟΚ και των δειγμάτων ελέγχου σε κάθε ένα από τα δύο σύνολα δεδομένων, χρησιμοποιώντας το μη παραμετρικό τεστ Wilcoxon που αποτελεί έναν εύχρηστο και ακριβή τρόπο αναγνώρισης διαφορικά εκφραζόμενων γονιδίων [168].

Το σύνολο δεδομένων με πλάσμα αίματος (GSE138118) περιλαμβάνει συνολικά 75 δείγματα, 11 από νεοδιαγνωσθέντες ασθενείς με ΚΟΚ, 18 από ασθενείς με παλιότερα διαγνωσμένη υποτροπή, 17 από ασθενείς με ενεργή υποτροπή ΚΟΚ και 29 από υγιείς εθελοντές χωρίς προηγούμενο ιστορικό ΚΟΚ ή οποιονδήποτε άλλου καρκίνου. Το συνολικό mRNA του πλάσματος απομονώθηκε από τα κλινικά δείγματα πλήρους αίματος και υβριδοποιήθηκε με χρήση της πλατφόρμας Affymetrix Human Gene 2.1 ST Array (GPL17692).

Τα πρωτογενή δεδομένα έκφρασης ελήφθησαν και πραγματοποιήθηκε προεπεξεργασία αυτών με βάση τη συγκεκριμένη πλατφόρμα, όπως περιγράφηκε προηγουμένως στην Ενότητα 2.3 για τα άλλα σύνολα δεδομένων. Πραγματοποιήθηκε διόρθωση των φαινομένων παρτίδας, για την ελαχιστοποίηση της ανεπιθύμητης διακύμανσης στα δεδομένα, χρησιμοποιώντας τη συνάρτηση *sna*, όπως υλοποιείται στο ομώνυμο R/Bioconductor πακέτο. Η συγκεκριμένη συνάρτηση μπορεί να χρησιμοποιηθεί χωρίς να είναι εκ των προτέρων γνωστές οι μεταβλητές των φαινομένων παρτίδας. Για αυτό το σύνολο δεδομένων, μόνο τα 28 δείγματα αίματος από ασθενείς με ΚΟΚ, δηλαδή ασθενείς που είχαν διαγνωσθεί πρόσφατα ή ήταν σε ενεργή υποτροπή, διατηρήθηκαν μαζί με τα 29 δείγματα αίματος από υγιή άτομα ελέγχου. Καθένα από τα βασικά γονίδια – κόμβους δοκιμάστηκε για τη στατιστικά σημαντική διαφορά του μεταξύ των δειγμάτων ΚΟΚ και των υγιών δειγμάτων, χρησιμοποιώντας το μη παραμετρικό τεστ Wilcoxon.

2.12 Εύρεση Προγνωστικών Γονιδίων για τον Καρκίνο της Ουροδόχου Κύστης

Προκειμένου να προσδιοριστούν ποια από τα βασικά γονίδια – κόμβους έχουν προγνωστική αξία, διεξήχθη ανάλυση επιβίωσης ασθενών. Στο πλαίσιο αυτό, χρησιμοποιήθηκε το σύνολο δεδομένων GSE13507, το οποίο περιέχει δεδομένα γονιδιακής έκφρασης από 165 ασθενείς με ΚΟΚ διαφόρων σταδίων (102 δείγματα μη μυοδιηθητικού ΚΟΚ και 63 δείγματα μυοδιηθητικού ΚΟΚ) [169]. Τα κλινικά δεδομένα των ασθενών είναι επίσης διαθέσιμα και περιέχουν πληροφορίες για τα συμβάντα που σχετίζονται με τον ΚΟΚ και τον συνολικό χρόνο επιβίωσης (ΧΕ) των ασθενών. Διεξήχθη μια μονο-μεταβλητή ανάλυση παλινδρόμησης κινδύνου Cox στα βασικά γονίδια – κόμβους για να αξιολογηθεί η συσχέτιση μεταξύ του σχετικού με τον καρκίνο χρόνου επιβίωσης κάθε ασθενούς και των τιμών

έκφρασης των γονιδίων, λαμβάνοντας υπόψη μόνο γονίδια με τιμή- $p < 0,05$. Το R πακέτο *survival* (version 3.2) χρησιμοποιήθηκε για τη διεξαγωγή της μονο-μεταβλητής ανάλυσης παλινδρόμησης Cox [170]. Με σκοπό να επιλεγεί ένα πάνελ γονιδίων και στη συνέχεια να δημιουργηθεί ένα πολυ-γονιδιακό προγνωστικό μοντέλο, εφαρμόστηκε ο ελάχιστος απόλυτος χειριστής συρρίκνωσης και επιλογής (Least Absolute Shrinkage and Selection Operator – LASSO) παλινδρόμησης Cox, εφαρμόζοντας τη 10-πλη διασταυρούμενη επικύρωση (10-fold cross validation) για 100 επαναλήψεις, χρησιμοποιώντας το R πακέτο *glmnet* (version 4.1) [171]. Με στόχο την εξάλειψη της συσχέτισης των επιλεγμένων γονιδίων και την αποτροπή της υπερ-προσαρμογής του μοντέλου, οι συντελεστές γονιδίων συρρικνώθηκαν προς το μηδέν, εφαρμόζοντας την ελάχιστη απόκλιση *lambda.min* σε κάθε επανάληψη και χρησιμοποιώντας τον *Harrell's C-index* (δείκτης συμφωνίας) ως μέτρο της προσαρμογής. Επιλέχθηκαν τα γονίδια με μη μηδενικούς συντελεστές για το 75% των επαναλήψεων. Προκειμένου να περιοριστεί περαιτέρω η λίστα γονιδίων και να βελτιστοποιηθεί το μοντέλο, πραγματοποιήθηκε πολυ-μεταβλητή ανάλυση Cox για τον εντοπισμό των ανεξάρτητων προγνωστικών παραγόντων όσον αφορά την πορεία των ασθενών με ΚΟΚ και την κατασκευή ενός μοντέλου προγνωστικού δείκτη (ΠΔ). Ο προγνωστικός δείκτης υπολογίστηκε με βάση τον τύπο:

$$ΠΔ = \sum_{i=1}^n c_i \cdot X_i$$

όπου c_i είναι ο συντελεστής του $i^{ου}$ γονιδίου, X_i είναι η έκφραση του $i^{ου}$ γονιδίου και n είναι ο αριθμός των επιλεγμένων γονιδίων στο βέλτιστο μοντέλο. Ο προγνωστικός δείκτης υπολογίστηκε για κάθε ασθενή και η διάμεση τιμή ορίστηκε ως η τιμή αποκοπής που διαχωρίζει τους ασθενείς με ΚΟΚ σε ομάδες χαμηλού και υψηλού κινδύνου, όσον αφορά την επιβίωσή τους. Οι καμπύλες ROC για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση σχεδιάστηκαν μαζί με τις τιμές AUC (εμβαδό κάτω από την καμπύλη) για την αξιολόγηση της απόδοσης του μοντέλου, χρησιμοποιώντας τα R πακέτα *survivalROC* (έκδοση 1.3.0) [172] και *plotROC* (έκδοση 2.2.1) [173]. Για να διερευνηθεί η σχέση μεταξύ των προγνωστικών γονιδίων αυτού του πάνελ, προσδιορίστηκε ο συντελεστής συσχέτισης Pearson μεταξύ όλων των πιθανών ζευγών. Τέλος, το R πακέτο *survminer* (έκδοση 0.4.9) χρησιμοποιήθηκε για την εκτέλεση της ανάλυσης του μοντέλου αναλογικών κινδύνων του Cox.

Με σκοπό τη διερεύνηση της πιθανής εφαρμογής του προγνωστικού μοντέλου σε άλλα σύνολα δεδομένων και της επικύρωσης της προγνωστικής του αξίας, λήφθηκαν δύο ανεξάρτητα σύνολα δεδομένων μικροσυστοιχιών από το αποθετήριο GEO (GSE32894 και GSE32548), τα οποία περιείχαν δεδομένα γονιδιακής έκφρασης μαζί με πληροφορίες επιβίωσης ασθενών με ΚΟΚ. Το GSE32894 περιέχει συνολικά 224 πρωτογενή δείγματα ΚΟΚ διαφόρων σταδίων, τα οποία χαρακτηρίστηκαν από την πλατφόρμα Illumina HumanHT-12 V3.0 (GPL6947) [174]. Ωστόσο, το αρχικό σύνολο δεδομένων περιέχει περισσότερα δείγματα, αλλά οι πληροφορίες σχετικά με τα συμβάντα επιβίωσης είναι διαθέσιμες μόνο για ένα υποσύνολο των αρχικών δειγμάτων, το οποίο και αξιοποιήθηκε. Το σύνολο δεδομένων GSE32548 περιλαμβάνει συνολικά 131 δείγματα πρωτογενών όγκων ΚΟΚ, τα οποία χαρακτηρίστηκαν από την ίδια πλατφόρμα (GPL6947) [175]. Στη συνέχεια, υπολογίστηκε ο προγνωστικός δείκτης για κάθε ασθενή από τα δύο σύνολα δεδομένων. Με βάση αυτόν τον δείκτη, οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου και δημιουργήθηκαν οι καμπύλες επιβίωσης των Kaplan – Meier για να συγκριθεί η επιβίωση μεταξύ των δύο ομάδων με χρήση του τεστ log-rank, θεωρώντας την τιμή- $p < 0,05$ ως στατιστικά σημαντική. Υπολογίστηκαν επίσης οι λόγοι κινδύνου (Hazard Ratios – HR) και τα διαστήματα εμπιστοσύνης 95% (Confidence Intervals – CI). Τέλος, διεξήχθησαν αναλύσεις των καμπυλών ROC εξαρτώμενων από το χρόνο με σκοπό να αξιολογηθεί η προγνωστική ικανότητα και η απόδοση του μοντέλου προγνωστικού κινδύνου.

Με σκοπό την περαιτέρω διερεύνηση και επικύρωση των αποτελεσμάτων, χρησιμοποιήθηκαν δημόσια διαδικτυακά διαθέσιμα εργαλεία βιοπληροφορικής για τη διερεύνηση της προγνωστικής αξίας των προσδιορισθέντων γονιδίων – κόμβων. Η πλατφόρμα «Διαδραστική Ανάλυση του Προφίλ της Γονιδιακής Έκφρασης» (Gene Expression Profiling Interactive Analysis – GEPIA2) [176] αποτελεί ένα διαδικτυακό εργαλείο ανοιχτής πρόσβασης για τη διαδραστική εξερεύνηση δεδομένων αλληλούχισης RNA, το οποίο βασίζεται στα ερευνητικά προγράμματα «Άτλας του Καρκινικού Γονιδιώματος» (The Cancer Genome Atlas – TCGA) [177] και «Γονότυπος – Ιστοειδική έκφραση» (Genotype-Tissue Expression – GTEx) [178]. Το GEPIA2 χρησιμοποιήθηκε για την αξιολόγηση της προγνωστικής αξίας των μεμονωμένων γονιδίων – κόμβων, όσον αφορά τη συνολική επιβίωση (Overall Survival – OS) και την ελεύθερης νόσου επιβίωση (Disease-Free Survival – DFS) των ασθενών με ΚΟΚ του προγράμματος TCGA. Η κοόρτη των ασθενών με ΚΟΚ του TCGA αποτελείται από 404 ασθενείς με ΚΟΚ και 19 άτομα ελέγχου. Η διαφορά μεταξύ των ποσοστών επιβίωσης των ομάδων υψηλής και χαμηλής έκφρασης για κάθε ένα γονίδιο – κόμβου εκλέχθηκε χρησιμοποιώντας το τεστ log-rank, λαμβάνοντας υπόψη ως στατιστικά σημαντική διαφορά τις περιπτώσεις με τιμή- $p < 0,05$ και χρησιμοποιώντας τη διάμεση τιμή ή τα άνω και κάτω τεταρτημόρια ως κριτήρια αποκοπής. Επίσης, σχεδιάστηκαν οι καμπύλες επιβίωσης με τον υπολογισμένο λόγο κινδύνου και την τιμή- p log-rank. Τέλος, η πλατφόρμα GEPIA2 χρησιμοποιήθηκε για να επιβεβαιώσει την προγνωστική εγκυρότητα της γονιδιακής υπογραφής που δημιουργήθηκε παραπάνω, από την πολυμεταβλητή ανάλυση παλινδρόμησης Cox, και να εξάγει τις αντίστοιχες καμπύλες επιβίωσης.

2.13 Εύρεση Προβλεπτικών Γονιδίων για τον Καρκίνο της Ουροδόχου

Κύστης

Με στόχο την περαιτέρω διερεύνηση της προβλεπτικής αξίας των βασικών γονιδίων – κόμβων, συμπεριλήφθηκαν στην παρούσα εργασία δείγματα από ασθενείς με μυοδιηθητικό ΚΟΚ, οι οποίοι έλαβαν προεγχειρητική χημειοθεραπεία που περιείχε σισπλατίνη (cisplatin). Τα δείγματα αυτά είχαν υποβληθεί σε ανάλυση γονιδιακής έκφρασης με μικροσυστοιχίες και προήλθαν από το αποθετήριο GEO. Ένας από τους στόχους αυτής της μελέτης ήταν να διερευνήσει σε ποιο βαθμό οι υπογραφές γονιδιακής έκφρασης μπορούν να προβλέψουν την ανταπόκριση των ασθενών στη χημειοθεραπεία. Αξίζει να σημειωθεί ότι επί του παρόντος το τρέχον θεραπευτικό πρωτόκολλο για τον μυοδιηθητικό ΚΟΚ είναι η νεοεπικουρική χημειοθεραπεία που περιέχει λευκόχρυσο, γνωστό και ως πλατίνα (π.χ. σισπλατίνη), ακολουθούμενη από ριζική κυστεκτομή. Ωστόσο, για πολλούς ασθενείς υπάρχει χαμηλό ποσοστό επιτυχίας της χημειοθεραπείας και ως εκ τούτου διερευνώνται αρκετοί υποψήφιοι βιοδείκτες οι οποίοι να μπορούν να προβλέψουν την ανταπόκριση των ασθενών στη θεραπεία [179].

Στο πλαίσιο αυτής της ανάλυσης, επιλέχθηκε το σύνολο δεδομένων GSE169455, το οποίο περιλαμβάνει συνολικά 149 δείγματα, τα οποία προέρχονται όλα από ασθενείς με μυοδιηθητικό ΚΟΚ, οι οποίοι έλαβαν νεοεπικουρική χημειοθεραπεία με βάση τη σισπλατίνη και υποβλήθηκαν σε ριζική κυστεκτομή [180]. Στα πειράματα αυτά, το mRNA εξήχθη από δείγματα διουρηθρικής εκτομής της ουροδόχου κύστης και υβριδοποιήθηκε στη μικροσυστοιχία Affymetrix Human Gene 1.0 ST (GPL6244). Τα πρωτογενή δεδομένα έκφρασης λήφθηκαν και πραγματοποιήθηκε προεπεξεργασία με βάση τη συγκεκριμένη πλατφόρμα, όπως περιγράφηκε και προηγουμένως (Ενότητα 2.3). Πραγματοποιήθηκε διόρθωση των φαινομένων παρτίδας, με σκοπό την ελαχιστοποίηση της ανεπιθύμητης (μη βιολογικής) διακύμανσης των δεδομένων, χρησιμοποιώντας τη συνάρτηση *Combat* από το πακέτο R/Bioconductor *sna* [181].

Το βασικό μέτρο έκβασης ήταν η παθολογική ανταπόκριση του εκάστοτε δείγματος κυστεκτομής στη νεοεπικουρική χημειοθεραπεία, η οποία κατηγοριοποιήθηκε ως «πλήρης απόκριση», «μερική

απόκριση» και «καμία απόκριση». Η έκφραση καθενός από τα βασικά γονίδια – κόμβους ελέγχθηκε για τη στατιστικά σημαντική διαφορά της μεταξύ των τριών κατηγοριών απόκρισης, χρησιμοποιώντας το μη παραμετρικό τεστ Wilcoxon. Επιπλέον, διεξήχθη μονο-μεταβλητή ανάλυση παλινδρόμησης Cox προκειμένου να υπολογιστεί η συσχέτιση μεταξύ κάθε γονιδίου – κόμβου και της επιβίωσης χωρίς υποτροπή (Recurrence Free Survival – RFS), της σχετικής με τον καρκίνο επιβίωσης (Cancer Specific Survival – CSS) και της συνολικής επιβίωσης (Overall Survival – OS) κάθε ασθενούς, χρησιμοποιώντας το R πακέτο *survival* και θεωρώντας την τιμή- $p < 0,05$ ως στατιστικά σημαντική. Καθώς η μονο-μεταβλητή ανάλυση οδήγησε σε περιορισμένο αριθμό εντοπισθέντων γονιδίων, πραγματοποιήθηκε η παλινδρόμηση LASSO Cox (όπως περιγράφηκε στην Ενότητα 2.12), με χρήση της διασταυρούμενης επικύρωσης 10 μερών, για τη διαλογή των βασικών γονιδίων – κόμβων. Τα επικρατέστερα γονίδια με μη μηδενικό συντελεστή ενσωματώθηκαν στην πολυ-μεταβλητή ανάλυση, εφαρμόζοντας το μοντέλο αναλογικής παλινδρόμησης των κινδύνων Cox, το οποίο οδήγησε σε μια προγνωστική γονιδιακή υπογραφή. Τέλος, κατασκευάστηκε ένα μοντέλο προβλεπτικού κινδύνου, χρησιμοποιώντας τους αντίστοιχους συντελεστές της ανωτέρω γονιδιακής υπογραφής (όπως στην Ενότητα 2.12). Η βαθμολογία κινδύνου χώρισε τους ασθενείς σε ομάδες χαμηλού και υψηλού κινδύνου με βάση τη διάμεση τιμή. Οι καμπύλες επιβίωσης Kaplan – Meier σχεδιάστηκαν για τις δύο ομάδες και πραγματοποιήθηκε ανάλυση της καμπύλης ROC εξαρτώμενης από το χρόνο, με βάση τη βαθμολογία κινδύνου πρόβλεψης. Τέλος, υπολογίστηκαν οι τιμές AUC για την αξιολόγηση της απόδοσης πρόβλεψης.

Με σκοπό να ελεγχθεί η προβλεπτική απόδοση του μοντέλου των έξι γονιδίων, χρησιμοποιήθηκαν δύο ανεξάρτητα σύνολα δεδομένων μικροσυστοιχιών (GSE87304 και GSE69795) από τη βάση δεδομένων GEO, τα οποία περιλάμβαναν τις πληροφορίες επιβίωσης των ασθενών με KOK, οι οποίοι θα υποβάλλονταν σε νεοεπικουρική χημειοθεραπεία. Το σύνολο δεδομένων GSE87304 περιέχει 305 δείγματα από ασθενείς με μυοδιηθητικό KOK, τα οποία ελήφθησαν με διουρηθρική εκτομή πριν από την νεοεπικουρική χημειοθεραπεία, και τα οποία χαρακτηρίστηκαν με τη διάταξη Affymetrix Human Exon 1.0 ST Array [182]. Το σύνολο δεδομένων GSE69795 περιέχει 38 δείγματα μυοδιηθητικού KOK, που έχουν σταθεροποιηθεί με φορμαλίνη και έχουν ενσωματωθεί σε παραφίνη, τα οποία ελήφθησαν από διουρηθρική εκτομή από ασθενείς με KOK που έλαβαν νεοεπικουρική χημειοθεραπεία πυκνής δόσης με μεθοτρεξάτη (methotrexate), βινβλαστίνη (vinblastine), δοξορουβικίνη (doxorubicin) και σισπλατίνη μαζί με μπεβασιζουμάμπη (bevacizumab), και τα οποία χαρακτηρίστηκαν με Illumina HumanHT-12 WG-DASL V4.0 R2 [183]. Και για τα δύο σύνολα δεδομένων, οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου, σύμφωνα με τον προβλεπτικό δείκτη, και σχεδιάστηκαν οι καμπύλες επιβίωσης Kaplan – Meier, συμπεριλαμβανομένου του λόγου κινδύνου (HR) και του 95% διαστήματος εμπιστοσύνης (CI). Ακόμη, πραγματοποιήθηκε χρονοεξαρτώμενη ανάλυση ROC για την αξιολόγηση της προβλεπτικής ικανότητας του μοντέλου των έξι γονιδίων.

2.14 Ταυτοποίηση της Έκφρασης των Βιοδεικτών και Ανοσοϊστοχημεία

Με βάση τα βασικά γονίδια – κόμβους που προσδιορίστηκαν στην έρευνα και λαμβάνοντας υπόψη την παραπάνω ανάλυση, διακρίθηκαν **εννέα γονίδια – κόμβοι ως πιθανοί βιοδείκτες** που φαίνεται να διαδραματίζουν σημαντικό ρόλο στην ανάπτυξη και εξέλιξη του KOK. Οι βιοδείκτες αυτοί φάνηκε να εκφράζονται σημαντικά στα ούρα ή στο πλάσμα του αίματος ασθενών με KOK και να έχουν κάποια προγνωστική ή προγνωστική αξία.

Η πλατφόρμα GEPIA2 χρησιμοποιήθηκε για την επιβεβαίωση της διαφορικής έκφρασης των εννέα γονιδίων – κόμβων. Η επικύρωση έγινε συγκρίνοντας μεταγραφικά δεδομένα από τα σύνολα δεδομένων TCGA-BCa, TCGA normal και GTEx. Τα κριτήρια αποκοπής ορίστηκαν ως $|\log_2FC| > 1$ και τιμή- $p < 0,05$ για να θεωρηθεί στατιστικά σημαντική η διαφορά στη γονιδιακή έκφραση. Επιπλέον, μέσω της

πλατφόρμας GEPIA2, πραγματοποιήθηκε η συσχέτιση των εννέα γονιδίων – κόμβων με τα επιμέρους παθολογικά στάδια του ΚΟΚ.

Επίσης, η έκφραση πρωτεϊνών που κωδικοποιούνται από αυτά τα γονίδια – κόμβους επικυρώθηκε σε δείγματα ΚΟΚ αξιοποιώντας την πλατφόρμα Human Protein Atlas, η οποία ενσωματώνει δεδομένα ανάλυσης από ανοσοϊστοχημική έκφραση (IHC) του γονιδιώματος των ανθρώπινων πρωτεϊνών [184].

2.15 Ανάλυση Διαγνωστικής Απόδοσης των Βιοδεικτών

Χρησιμοποιώντας τα εννέα γονίδια – κόμβους ως χαρακτηριστικά (features), κατασκευάστηκαν και δοκιμάστηκαν διάφορα μοντέλα ταξινόμησης, ώστε να αξιολογηθεί η διαγνωστική τους απόδοση. Αρχικά, όλα τα σύνολα δεδομένων που συμπεριελήφθησαν στην τρέχουσα ανάλυση και περιείχαν περισσότερα από 10 δείγματα επαναχρησιμοποιήθηκαν ως σύνολα εκπαίδευσης / δοκιμών, προκειμένου να επικυρωθεί η διαγνωστική ικανότητα των εννέα γονιδίων – κόμβων / χαρακτηριστικών. Η διαγνωστική τους σημασία αξιολογήθηκε επίσης στο τελικό ενοποιημένο μετα-σύνολο δεδομένων.

Τέλος, ένα εξωτερικό σύνολο δεδομένων χρησιμοποιήθηκε για την περαιτέρω αξιολόγηση αυτών των γονιδίων – κόμβων / χαρακτηριστικών ως διαγνωστικών βιοδεικτών για τον ΚΟΚ. Αυτό το σύνολο δεδομένων ελήφθη από το αποθετήριο ArrayExpress του Ευρωπαϊκού Ινστιτούτου Βιοπληροφορικής (European Molecular Biology Laboratory – European Bioinformatics Institute, EMBL – EBI) [185] [88] και περιέχει 19 δείγματα ΚΟΚ (14 μη μυοδιηθητικού και 5 μυοδιηθητικού ΚΟΚ) και 11 δείγματα ελέγχου από βιοψίες ιστού ουροδόχου κύστης [186], τα οποία υβριδοποιήθηκαν με χρήση της πλατφόρμας Affymetrix Human Gene 1.0 ST (GPL6244).

Για κάθε σύνολο δεδομένων, εφαρμόστηκε η τεχνική διασταυρούμενης επικύρωσης 5 μερών και επαναλήφθηκε 10 φορές, προκειμένου να ληφθεί μια πιο ακριβής αξιολόγηση της απόδοσης των μοντέλων ταξινόμησης. Για το τελικό ενοποιημένο μετα-σύνολο δεδομένων, εφαρμόστηκε η διασταυρούμενη επικύρωση 10 μερών. Για όλα τα μοντέλα που αναπτύχθηκαν, η τιμή AUC του ταξινομητή χρησιμοποιήθηκε για την αξιολόγηση της διαγνωστικής απόδοσης του εκάστοτε μοντέλου. Οι καμπύλες ROC που προέκυψαν για όλα τα κατασκευασμένα μοντέλα μαζί με τις αντίστοιχες τιμές AUC και το 95% CI σχεδιάστηκαν. Κάθε σημείο στις καμπύλες ROC υποδηλώνει ένα ζεύγος ευαισθησίας / ειδικότητας που λαμβάνεται από ένα συγκεκριμένο όριο απόφασης και η τιμή AUC υποδεικνύει την αποτελεσματικότητα του αντίστοιχου μοντέλου. Όσο πιο κοντά είναι η τιμή AUC στο 1,0 τόσο καλύτερη είναι η απόδοση του μοντέλου ταξινόμησης.

3. Αποτελέσματα

3.1 Συστηματική αναζήτηση και Επιλογή Συνόλων Δεδομένων

Συνολικά, 18 μελέτες από το αποθετήριο GEO πληρούσαν τα κριτήρια συμπερίληψης (Εικόνα 30) και επιλέχθηκαν για την ενοποιημένη μετα-ανάλυση. Το τελικό σύνολο δεδομένων περιελάμβανε 619 δείγματα (417 δείγματα ΚΟΚ και 202 δείγματα ελέγχου). Τα πειράματα των συνόλων δεδομένων που περιελήφθησαν σε αυτή τη μετα-ανάλυση ακολουθούσαν παρόμοιο πειραματικό σχεδιασμό και συνέκριναν ανθρώπινους ιστούς από δείγματα ΚΟΚ με υγιή δείγματα. Συγκεκριμένα, στα πειράματα των συνόλων δεδομένων χρησιμοποιήθηκαν 13 διαφορετικές πλατφόρμες μικροσυστοιχιών. Ο Πίνακας 2 παρέχει λεπτομερείς πληροφορίες για κάθε σύνολο δεδομένων που περιελήφθηκε στην ενοποιητική μετα-ανάλυση και εμπεριέχει τον τύπο των δειγμάτων, τα φαινοτυπικά χαρακτηριστικά τους, το έτος και την αναφορά της μελέτης καθώς και την πλατφόρμα μικροσυστοιχίας που χρησιμοποιήθηκε.

Πίνακας 2: Τα χαρακτηριστικά των 18 σειρών / συνόλων δεδομένων που συμπεριλήφθηκαν στην παρούσα ενοποιητική μετα-ανάλυση.

Αριθμός πρόσβασης	Δείγματα (n)			Έτος	Πλατφόρμα	Χαρακτηριστικά δειγμάτων	Αναφορά
	Σύνολο	ΚΟΚ	Υγιή				
GSE3167	60	46	14	2005	GPL96 [HG-U133A] Affymetrix Human Genome U133A Array	· 13 superficial transitional cell carcinomas with surrounding CIS · 15 without surrounding CIS lesions · 13 muscle invasive carcinomas · 5 CISs · 14 normal bladder tissues	[187]
GSE7476	12	9	3	2007	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	· 3 groups of 5 pooled Ta tumours · 1 group of 5 pooled T1 tumours · 2 groups of 4 pooled T1 tumours · 3 groups of 5 pooled T2+ tumours · 3 groups of 4 pooled normal bladder tissues	[188]
GSE13507	232	170	62	2010	GPL6102 Illumina human-6 v2.0 expression beadchip	· 24 primary Ta cancer tissues · 80 primary T1 cancer tissues · 31 primary T2 cancer tissues · 19 primary T3 cancer tissues · 11 primary T4 cancer tissues · 23 recurrent NMIBC tissues · 68 normal bladder tissues	[169]
GSE21142	24	12	12	2013	GPL10274 Affymetrix GeneChip Human Genome U133 Plus 2.0 Array (Brainarray CustomCDF, HGU133Plus2_Hs_UG_Version 12.cdf)	· 6 superficial urothelial carcinomas · 6 invasive urothelial carcinomas · 12 normal bladder tissues	[189]
GSE23732	8	7	1	2012	GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	· 7 muscle invasive bladder cancer · 1 normal bladder tissue	-
GSE24152	17	10	7	2010	GPL6791 Affymetrix GeneChip Human Genome U133 Plus 2.0 Array [CDF: Hs_ENTREZG_10]	· 10 muscle invasive urothelial bladder carcinomas · 7 benign bladder tissues	[190]

GSE31189	92	52	40	2013	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	· 52 urothelial cancer cells · 40 normal urothelial cells	[191]
GSE37815	24	18	6	2013	GPL6102 Illumina human-6 v2.0 expression beadchip	· 18 NMIBC tissues · 6 normal bladder tissues	[192]
GSE38264	38	28	10	2014	GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	· 28 Ta and T1 tumours · 10 normal bladder tissues	[193]
GSE40355	24	16	8	2013	GPL13497 Agilent-026652 Whole Human Genome Microarray 4x44K v2 (Probe Name version)	· 8 Ta urothelial carcinoma tissues · 5 T1 urothelial carcinoma tissues · 3 T2 urothelial carcinoma tissues · 8 normal bladder tissue samples	[194]
GSE41614	10	5	5	2013	GPL5175 [HuEx-1_0-st] Affymetrix Human Exon 1.0 ST Array [transcript (gene) version]	· 2 samples with blood vessels from T1 bladder cancer tissue · 2 samples with blood vessels from T2 bladder cancer tissue · 1 sample with blood vessels from T3 bladder cancer tissue · 5 samples with blood vessels from normal bladder	[195]
GSE42089	18	10	8	2013	GPL9828 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array [CDF: Brainarray Hs133P_Hs_ENTREZG version 10)	· 10 urothelial cell carcinomas · 8 normal bladder tissues	[196]
GSE45184	6	3	3	2013	GPL14550 Agilent-028004 SurePrint G3 Human GE 8x60K Microarray (Probe Name Version)	· 3 bladder cancer tissues · 3 normal adjacent tissues	[197]
GSE52519	12	9	3	2013	GPL6884 Illumina HumanWG-6 v3.0 expression beadchip	· 1 T1 cancerous tissue sample · 2 T2 cancerous tissue samples · 2 T3 cancerous tissue samples · 4 T4 cancerous tissue samples · 3 normal bladder tissues	[198]
GSE65635	12	8	4	2015	GPL14951 Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	· 5 T1 bladder cancer tissues · 2 T3 bladder cancer tissues · 1 T4 bladder cancer tissues · 4 normal bladder tissues	[198]
GSE76211	6	3	3	2017	GPL17586 [HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]	· 3 T3 bladder cancer tissues · 3 normal bladder tissues	[199]
GSE100926	6	3	3	2017	GPL14550 Agilent-028004 SurePrint G3 Human GE 8x60K Microarray (Probe Name Version)	· 2 T2 MIBC tissues · 1 T3 MIBC tissue · 3 normal bladder tissues	[200]
GSE121711	18	8	10	2019	GPL17586 [HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]	· 3 Ta primary tumours · 2 T1 primary tumours · 3 T2 primary tumours · 10 normal bladder tissues	[201]
Σύνολο	619	417	202	--	--	--	--

3.2 Ποιοτικός Έλεγχος Δειγμάτων

Συνολικά, οι τιμές έκφρασης 13 δειγμάτων χαρακτηρίστηκαν ως ακραίες τιμές, σύμφωνα με το εφαρμοσμένο πλαίσιο ποιοτικού ελέγχου, και κατά συνέπεια τα δείγματα αυτά εξαιρέθηκαν από την περαιτέρω ανάλυση. Πιο συγκεκριμένα, αφαιρέθηκαν δύο δείγματα από το σύνολο δεδομένων GSE3167 (ένα δείγμα KOK και ένα δείγμα ελέγχου), ένα δείγμα από το σύνολο δεδομένων GSE13507

(δείγμα KOK), ένα δείγμα από το σύνολο δεδομένων GSE21142 (δείγμα KOK), πέντε δείγματα από το σύνολο δεδομένων GSE31189 (δύο δείγματα KOK και τρία δείγματα ελέγχου), ένα δείγμα από το σύνολο δεδομένων GSE38264 (δείγμα ελέγχου), δύο δείγματα από το σύνολο δεδομένων GSE40355 (δύο δείγματα KOK) και ένα δείγμα από το σύνολο δεδομένων GSE42089 (δείγμα ελέγχου). Μετά την εξαίρεση κάποιων δειγμάτων βάσει του ποιοτικού ελέγχου, η μετα-ανάλυση περιελάμβανε **606 δείγματα**, αποτελούμενα από **410 δείγματα KOK** και **196 δείγματα ελέγχου**.

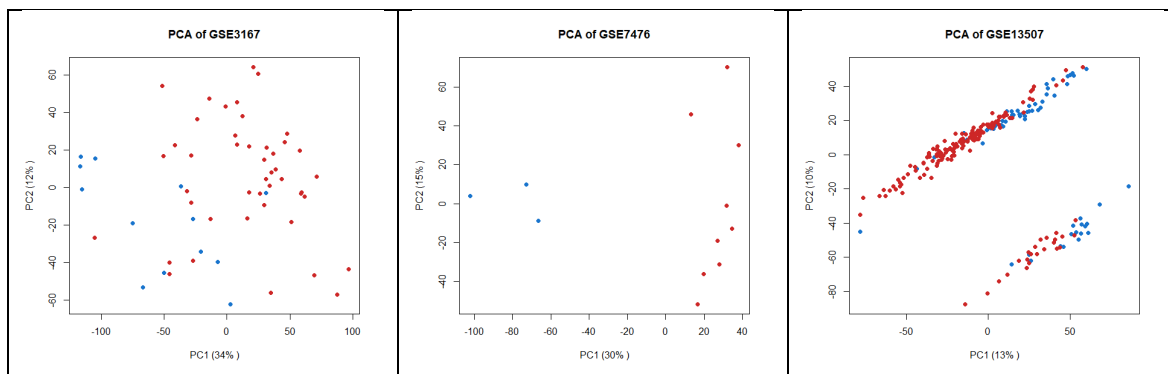
3.3 Αντιστοίχιση Ανιχνευτών – Γονιδίων

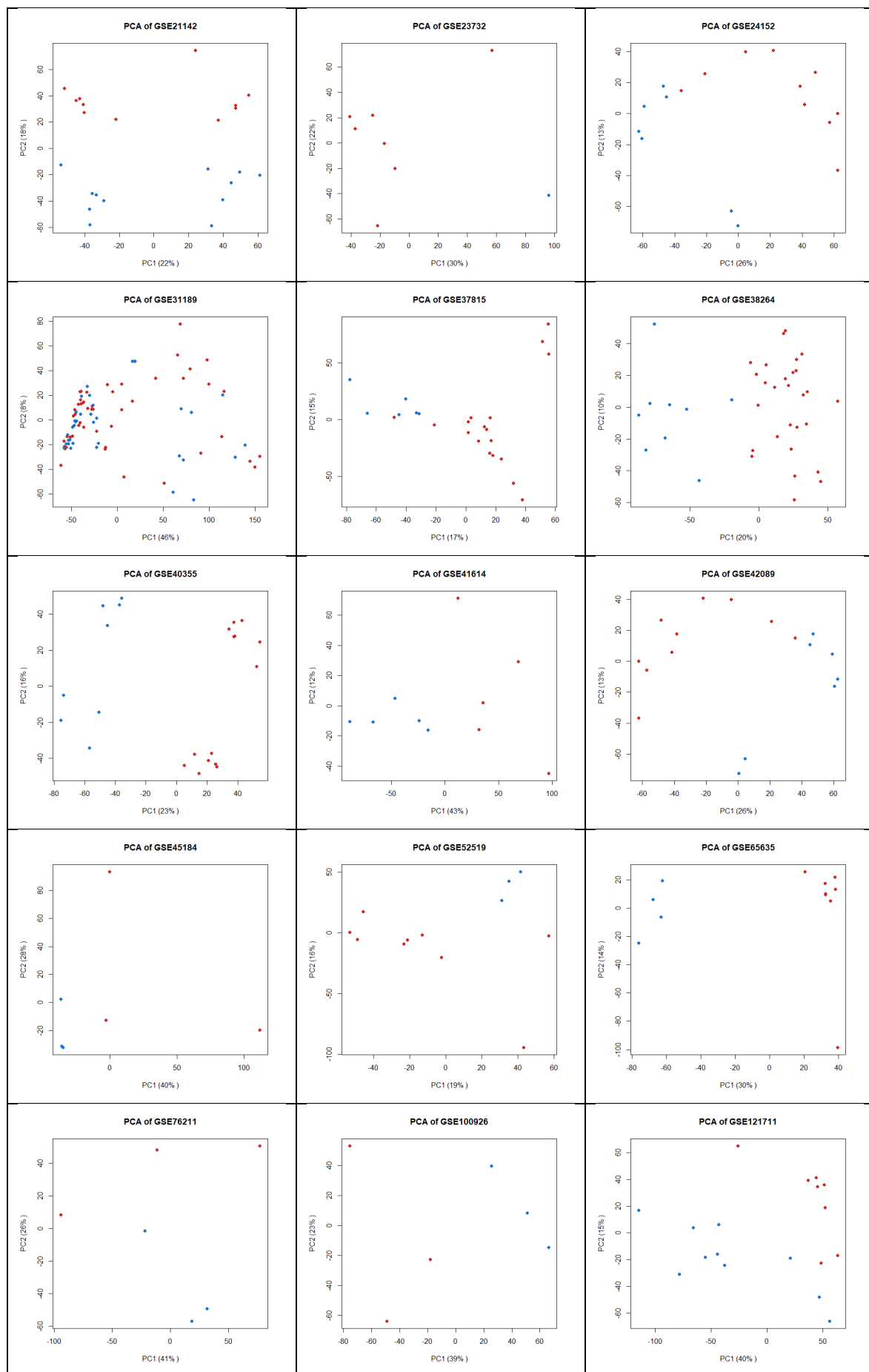
Μετά από την κατάλληλη αντιστοίχιση των ανιχνευτών με τα γονιδιακά σύμβολα για κάθε πλατφόρμα, αντιπαρατέθηκαν τα 13 διαφορετικά σύνολα γονιδιακών συμβόλων των αντίστοιχων μικροσυστοιχιών (*Affy HG U133A*, *Affy HG U133 Plus 2*, *Illu Human-6 V2*, *Affy HuGene 1 ST*, *Agi WHG 4x44K V2*, *Affy HuEx 1 ST*, *Agi G3 GE 8x60K*, *Illu Human-6 V3*, *Illu Human-12 V4*, *Affy HTA 2* και τρία προσαρμοσμένα αρχεία CDF Brainarray για *Affy HG U133 Plus 2*). Συνολικά, οι ανιχνευτές στις 13 πλατφόρμες μικροσυστοιχιών στόχευαν σε 27.579 μοναδικά γονιδιακά σύμβολα, από τα οποία **8.201 σύμβολα γονιδίων ήταν κοινά** και στις 13 πλατφόρμες μικροσυστοιχιών. Ως εκ τούτου, η ενοποιημένη μετα-ανάλυση διεξήχθη μόνο για αυτά τα 8.201 κοινά γονιδιακά σύμβολα σε όλα τα σύνολα δεδομένων.

3.4 Φαινόμενα Παρτίδας και Διατεχνολογική Κανονικοποίηση

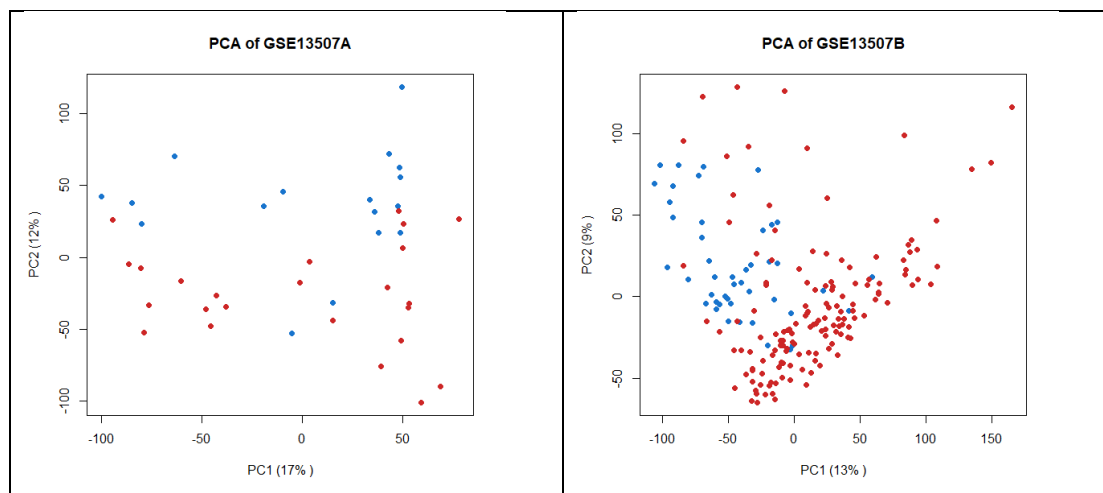
Τα φαινόμενα παρτίδας που παρουσιάζονται σε κάθε σύνολο δεδομένων επιθεωρήθηκαν οπτικά χρησιμοποιώντας την ανάλυση κύριων συνιστωσών (PCA) (Εικόνα 31). Λόγω ενός πολύ ισχυρού φαινομένου παρτίδας που ανιχνεύτηκε στο σύνολο δεδομένων GSE13507, τα δείγματα από το σύνολο αυτό διαχωρίστηκαν περαιτέρω σε δύο υποομάδες, την GSE13507A και GSE13507B, με 41 (24 δείγματα KOK και 17 δείγματα ελέγχου) και 190 (145 δείγματα KOK και 45 δείγματα ελέγχου) δείγματα, αντίστοιχα (Εικόνα 31). Τα νέα διαγράμματα PCA για καθένα από αυτά τα δύο υποσύνολα δεδομένων παρουσιάζονται στην Εικόνα 32. Αυτά τα δύο νέα υποσύνολα δεδομένων θεωρήθηκαν ως διακριτά και ανεξάρτητα υποσύνολα δεδομένων για την περαιτέρω ανάλυση.

Ο μετασχηματισμός Z εφαρμόστηκε για τη διόρθωση των τιμών της γονιδιακής έκφρασης εντός του δείγματος αλλά και για την προσαρμογή των συστηματικών σφαλμάτων μεταξύ των συνόλων δεδομένων που προέρχονται από διαφορετικές πλατφόρμες μικροσυστοιχιών και διαφορετικά εργαστήρια. Επομένως, οι τιμές έκφρασης για κάθε γονίδιο ενός δείγματος εκφράζονται σε μονάδες τυπικής απόκλισης από τη μέση τιμή (που είναι μηδενική). Με αυτόν τον τρόπο, οι συγκρίσεις μεταξύ δειγμάτων πραγματοποιήθηκαν σε ομοιόμορφα μετασχηματισμένα δεδομένα.





Εικόνα 31: Εφαρμογή της μεθόδου PCA σε καθένα από τα 18 σύνολα δεδομένων, χρησιμοποιώντας την ευκλείδεια απόσταση για τη μέτρηση των ανομοιοτήτων στις τιμές γονιδιακής έκφρασης μεταξύ των δειγμάτων. Τα κόκκινα σημεία υποδηλώνουν δείγματα KOK και τα μπλε σημεία υποδηλώνουν δείγματα ελέγχου.



Εικόνα 32: Εφαρμογή της μεθόδου PCA στα σύνολα δεδομένων GSE13507A και GSE13507B. Τα κόκκινα σημεία υποδηλώνουν δείγματα KOK και τα μπλε σημεία υποδηλώνουν δείγματα ελέγχου.

3.5 Ανάλυση Διαφορικής Έκφρασης Γονιδίων

Μετά την προεπεξεργασία και την τυποποίηση κάθε συνόλου δεδομένων, συνδυάστηκαν τα Z-μετασχηματισμένα δεδομένα έκφρασης για κάθε δείγμα σε ένα ενοποιημένο σύνολο δεδομένων χρησιμοποιώντας τα κοινά σύμβολα γονιδίων. Αυτό το ενοποιημένο μετα-σύνολο δεδομένων περιελάμβανε 606 δείγματα (410 καρκινικά δείγματα και 196 δείγματα ελέγχου) και 8.201 κοινά σύμβολα γονιδίων.

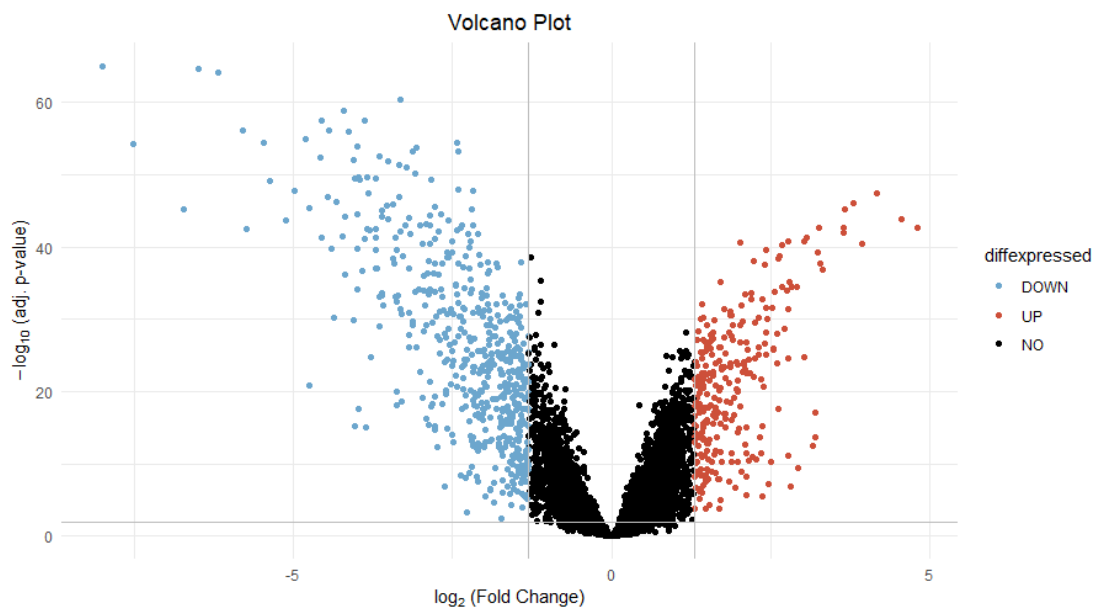
Τα **διαφορικά εκφραζόμενα γονίδια** μεταξύ των καρκινικών και των υγιών δειγμάτων ιστού των 19 ενοποιημένων συνόλων δεδομένων από το GEO προσδιορίστηκαν χρησιμοποιώντας το πακέτο *limma* του R/Bioconductor. Η τιμή-p διορθώθηκε χρησιμοποιώντας τη μέθοδο BH προκειμένου να ελεγχθεί το ψευδές ποσοστό ανακάλυψης και επιλέχθηκε ένα όριο αποκοπής διορθωμένης τιμής-p < 0,01. Με σκοπό να προσδιοριστεί η τιμή αποκοπής $|\log_2FC|$, υλοποιήθηκε ένα μοντέλο ταξινόμησης SVM για καθένα από τα διαφορετικά σύνολα διαφορικά εκφραζόμενων γονιδίων που αντιστοιχούν στις τιμές αποκοπής $|\log_2FC|$ από 1 έως 2 με βήμα 0,1. Έτσι, για κάθε τιμή αποκοπής $|\log_2FC|$ το αντίστοιχο σύνολο διαφορικά εκφραζόμενων γονιδίων χρησιμοποιήθηκε ως σύνολο χαρακτηριστικών του μοντέλου. Για κάθε μοντέλο, ο αριθμός των χαρακτηριστικών μαζί με την εκτιμώμενη περιοχή κάτω από την καμπύλη ROC (AUC), την ευαισθησία και την ειδικότητα του ταξινομητή παρουσιάζονται στον παρακάτω πίνακα (Πίνακας 3). Η τιμή AUC εκφράζει στην πραγματικότητα την πιθανότητα ένα δείγμα να έχει ταξινομηθεί σωστά. Όπως παρατηρεί κανείς, ο ταξινομητής επιτυγχάνει πολύ υψηλή ακρίβεια ταξινόμησης για τα διάφορα σύνολα χαρακτηριστικών και οι διαφορές είναι σχεδόν αμελητέες.

Πίνακας 3: Παράμετροι απόδοσης των διαφόρων μοντέλων ταξινόμησης για τα εκάστοτε σύνολα διαφορικά εκφραζόμενων γονιδίων.

$ \log_2FC $	Αριθμός χαρακτηριστικών (DEGs)	AUC	Ευαισθησία	Ειδικότητα
1	1295	0.9525	0.7964	0.9366
1.1	1099	0.9517	0.7934	0.9327
1.2	929	0.9527	0.7842	0.9346

1.3	815	0.9531	0.7985	0.9334
1.4	725	0.9510	0.7996	0.9322
1.5	625	0.9516	0.8022	0.9342
1.6	549	0.9487	0.7929	0.9278
1.7	495	0.9482	0.7844	0.9312
1.8	442	0.9510	0.7966	0.9298
1.9	407	0.9519	0.8001	0.9288
2.0	364	0.9507	0.7903	0.9356

Ωστόσο, η υψηλότερη τιμή για την τιμή AUC, η οποία υποδεικνύει και τον ταξινομητή με την υψηλότερη απόδοση, επιτεύχθηκε για την τιμή αποκοπής $|\log_2 FC| = 1,3$. Ως εκ τούτου, προέκυψαν συνολικά **815 διαφορεικά εκφραζόμενα γονίδια** μεταξύ των καρκινικών δειγμάτων και των δειγμάτων ελέγχου, όπως υπολογίστηκαν από τα προφίλ έκφρασης του πακέτου *limma*, εφαρμόζοντας ως κριτήρια αποκοπής $|\log_2 FC| \geq 1,3$ και προσαρμοσμένη τιμή- $p < 0,01$. Συνολικά, τα διαφορετικά εκφραζόμενα γονίδια περιείχαν **540 υπο-εκφραζόμενα γονίδια** και **275 υπερ-εκφραζόμενα γονίδια**. Το γράφημα ηφαιστείου (volcano plot) και ο θερμικός χάρτης (heatmap) των 100 κορυφαίων διαφορετικά εκφραζόμενων γονιδίων απεικονίζονται στην Εικόνα 33 και στην Εικόνα 34, αντίστοιχα. Στο θερμικό χάρτη που δημιουργήθηκε, πραγματοποιήθηκε ιεραρχική ομαδοποίηση σε επίπεδο γονιδίων αλλά και δειγμάτων. Μπορεί να παρατηρηθεί ότι στο χώρο γονιδιακής έκφρασης υπήρχαν έντονες επιδράσεις των φαινομένων παρτίδας, καθώς τα δείγματα ομαδοποιήθηκαν με βάση τον φαινότυπο αλλά και το πείραμα από το οποίο προέρχονταν.



Εικόνα 33: Διάγραμμα ηφαιστείου των διαφορετικά εκφραζόμενων γονιδίων μεταξύ των δειγμάτων ΚΟΚ και των δειγμάτων ελέγχου στο ενοποιημένο μετα-σύνολο δεδομένων. Τα διαφορετικά εκφραζόμενα γονίδια προσδιορίστηκαν με βάση τα κριτήρια $|\log_2(FC)| \geq 1,3$ και προσαρμοσμένη τιμή- $p < 0,01$, όπως απεικονίζονται με γκρίζες γραμμές. Τα μπλε και τα κόκκινα σημεία υποδηλώνουν τα γονίδια που υπο-εκφράζονται και υπερ-εκφράζονται, αντίστοιχα. Τα μαύρα σημεία υποδηλώνουν γονίδια που δεν παρουσιάζουν στατιστική σημαντική διαφορά στην έκφραση μεταξύ των δύο φαινοτύπων.



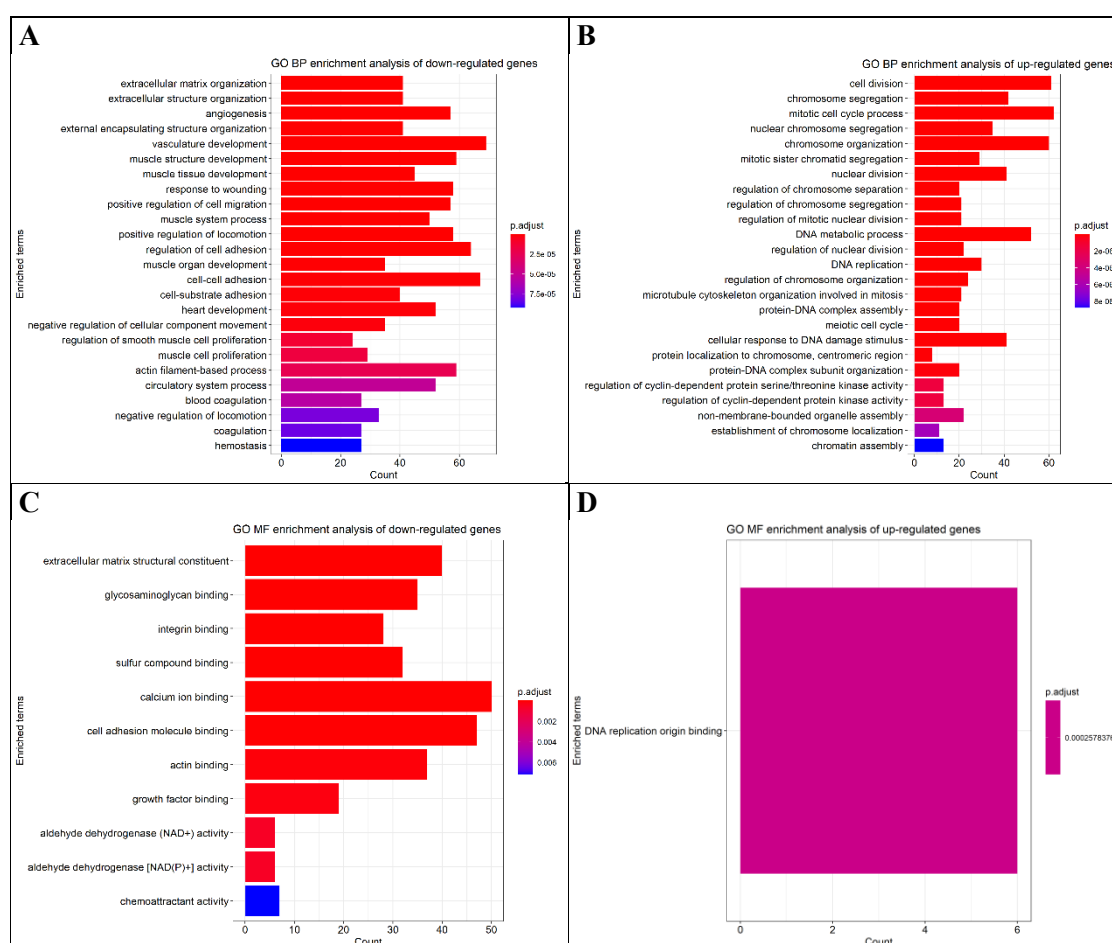
Εικόνα 34: Διάγραμμα θερμικού χάρτη των 100 κορυφαίων διαφορικά εκφραζόμενων γονιδίων μεταξύ των δειγμάτων ΚΟΚ και των υγιών δειγμάτων ελέγχου στο ενοποιημένο μετα-σύνολο δεδομένων. Το μπλε και το κόκκινο χρώμα αντιπροσωπεύουν τη σχετική μείωση και αύξηση της έκφρασης, αντίστοιχα, ενώ το λευκό χρώμα αντιπροσωπεύει ότι το γονίδιο δεν παρουσιάζει κάποια σημαντική αλλαγή στη γονιδιακή έκφραση. Οι οριζόντιοι και κάθετοι άξονες είναι ομαδοποιημένοι ανά γονίδια και δείγματα, αντίστοιχα.

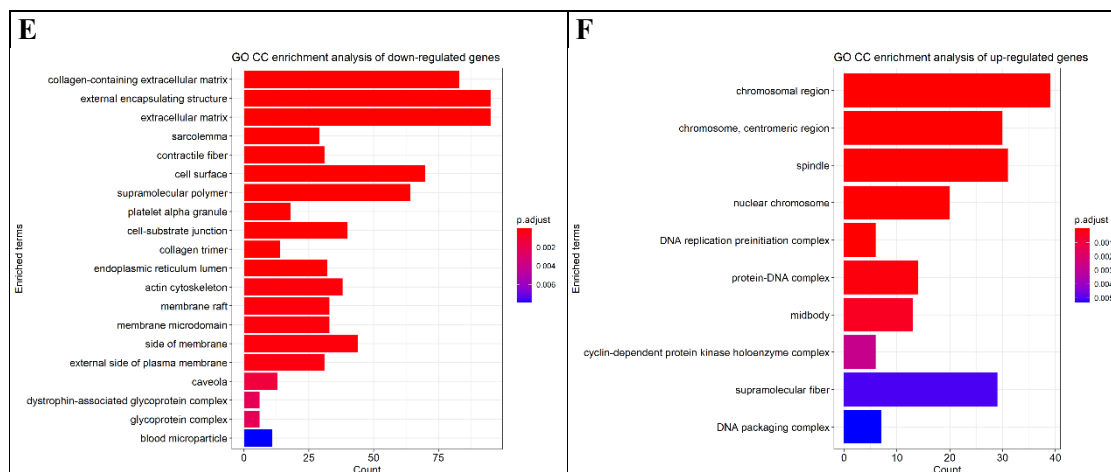
3.6 Λειτουργική Ανάλυση Εμπλουτισμού της Γονιδιακής Έκφρασης

Με σκοπό την εξαγωγή πληροφορίας για το λειτουργικό ρόλο των διαφορικά εκφραζόμενων γονιδίων και των μονοπατιών που εμπλέκονται στον ΚΟΚ, πραγματοποιήθηκε ολοκληρωμένη λειτουργική ανάλυση εμπλουτισμού της γονιδιακής έκφρασης σε διάφορες βάσεις βιολογικών δεδομένων. Πιο συγκεκριμένα, πραγματοποιήθηκε ανάλυση εμπλουτισμού των 815 διαφορικά εκφραζόμενων γονιδίων στην **Οντολογία Γονιδίων – ΟΓ (Gene Ontology – GO)**, στα μονοπάτια της **Εγκυκλοπαίδειας του Κιότο για τα Γονίδια και τα Γονιδιώματα – ΕΚΓΤ (Kyoto Encyclopedia of Genes and Genomes – KEGG)**, στα μονοπάτια της **Reactome (REAC)** και της **Οντολογίας Ασθενειών (Disease Ontology – DO)**, χρησιμοποιώντας το πακέτο *clusterProfiler* του R/Bioconductor. Για όλες τις λειτουργικές αναλύσεις που ακολούθησαν, οι παράμετροι των ορίων αποκοπής ήταν τιμή- $p = 0,01$ και τιμή- $q = 0,05$, διορθωμένες με βάση τη μέθοδο των Benjamini – Hochberg.

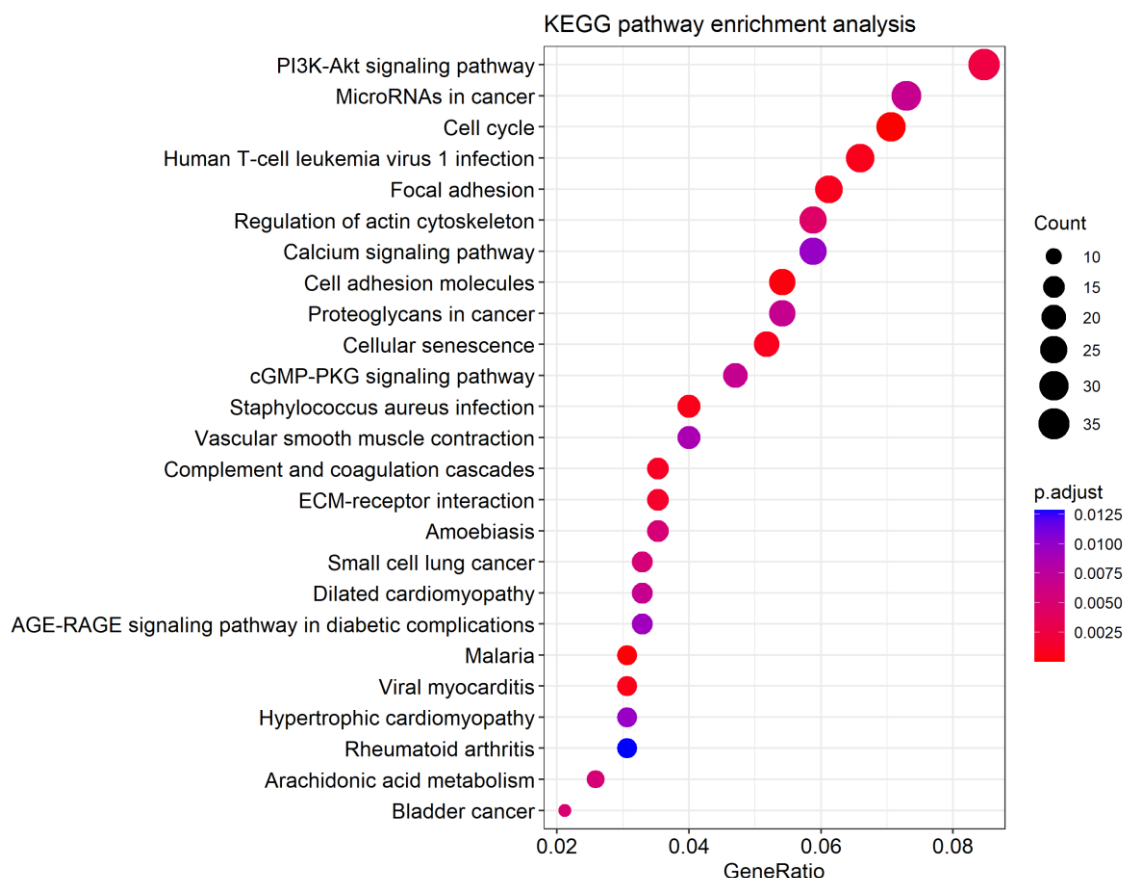
Αρχικά, πραγματοποιήθηκε λειτουργική ανάλυση εμπλουτισμού γονιδιακής έκφρασης στην Οντολογία Γονιδίων με βάση το σύνολο των αναγνωρισμένων διαφορικά εκφραζόμενων γονιδίων. Τα ραβδογράμματα των 25 κορυφαίων, εάν υπήρχαν τόσοι, εμπλουτισμένων όρων της Οντολογίας Γονιδίων, δημιουργήθηκαν και παρουσιάζονται στην Εικόνα 35. Τα γραφήματα αυτά απεικονίζουν τους κορυφαίους όρους οι οποίοι σχετίζονται με τις βιολογικές διεργασίες (biological processes – BP), τις μοριακές λειτουργίες (molecular functions – MF) και τα κυτταρικά συστατικά (cellular components – CC). Οι όροι της Οντολογίας Γονιδίων που σχετίζονταν με τα υπο-εκφραζόμενα γονίδια και τις βιολογικές διεργασίες περιλάμβαναν την οργάνωση της εξωκυτταρικής μήτρας (EMC), την οργάνωση της εξωκυτταρικής δομής, την αγγειογένεση, το αγγειακό σύστημα, τη δομή των μυών και την ανάπτυξη

του μυϊκού ιστού. Οι όροι της Οντολογίας Γονιδίων που σχετίζονταν με τα υπερ-εκφραζόμενα γονίδια και τις βιολογικές διεργασίες περιλάμβαναν την κυτταρική διαίρεση, τη διαδικασία του μιτωτικού κυτταρικού κύκλου, τον διαχωρισμό και την οργάνωση των χρωμοσωμάτων. Στην κατηγορία των μοριακών λειτουργιών, τα υπο-εκφραζόμενα γονίδια ήταν εμπλουτισμένα σε λειτουργίες σχετικές με το δομικό συστατικό της εξωκυτταρικής μήτρας, καθώς και σε λειτουργίες γλυκοζαμινογλυκάνης, ιντεγκρίνης, θειούχων ενώσεων και δέσμευσης ιόντων ασβεστίου. Τα υπερ-εκφραζόμενα γονίδια στην κατηγορία των μοριακών λειτουργιών, εμπλουτίστηκαν μόνο στη λειτουργία της δέσμευσης της προέλευσης αντιγραφής του DNA. Οι όροι της Οντολογίας Γονιδίων σχετικοί με τα κυτταρικά συστατικά των υπο-εκφραζόμενων γονιδίων αφορούσαν κυρίως την εξωκυτταρική μήτρα που περιέχει κολλαγόνο, την εξωκυτταρική δομή εγκλεισμού και την εξωκυτταρική μήτρα. Τέλος, οι όροι σχετικοί με τα κυτταρικά συστατικά των υπερ-εκφραζόμενων γονιδίων ήταν εμπλουτισμένοι στην άτρακτο, στην κεντρομερική περιοχή του χρωμοσώματος και στη χρωμοσωμική περιοχή.





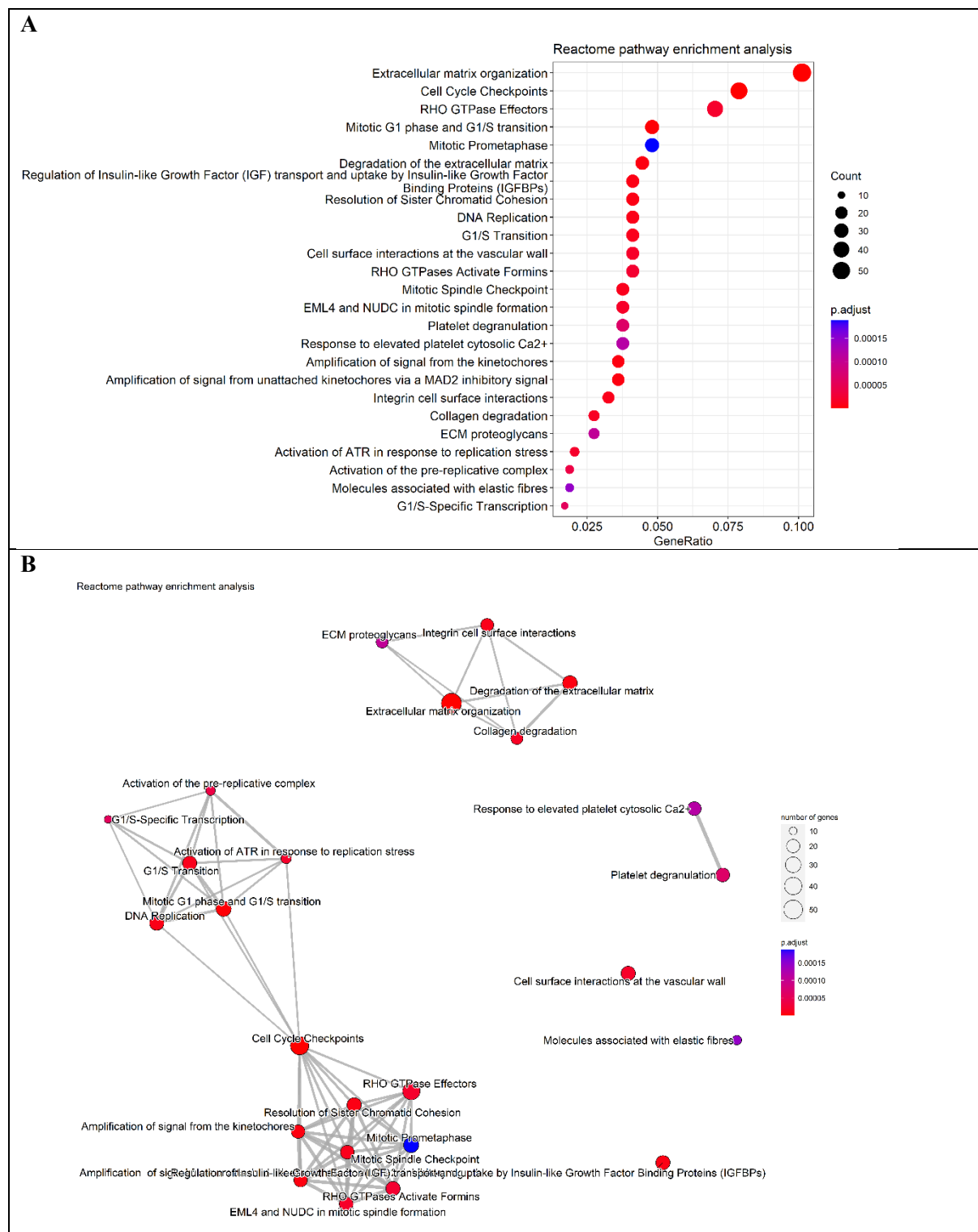
Εικόνα 35: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού με βάση την Οντολογία Γονιδίων (GO) όσον αφορά: (Α)-(Β) Τις βιολογικές διεργασίες (BP) των υποεκφραζόμενων και υπερεκφραζόμενων διαφορετικά εκφραζόμενων γονιδίων (C)-(D) Τις μοριακές λειτουργίες (MF) των υποεκφραζόμενων και υπερεκφραζόμενων διαφορετικά εκφραζόμενων γονιδίων (Ε)-(F) Τα κυτταρικά συστατικά (CC) των υποεκφραζόμενων και υπερεκφραζόμενων διαφορετικά εκφραζόμενων γονιδίων.



Εικόνα 36: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού μονοπατιών με βάση την KEGG.

Η ανάλυση εμπλουτισμού μονοπατιών με βάση την KEGG έδειξε ότι υπήρχαν 35 μονοπάτια εμπλουτισμένα στο σύνολο των διαφορετικά εκφραζόμενων γονιδίων. Οι πρώτοι 25 εμπλουτισμένοι όροι παρουσιάζονται στην Εικόνα 36. Η ανάλυση έδειξε ότι το σηματοδοτικό μονοπάτι PI3K-Akt, τα micro-RNAs στον καρκίνο, ο κυτταρικός κύκλος, η εστιακή προσκόλληση, τα μόρια κυτταρικής προσκόλλησης, η κυτταρική γήρανση, οι καταρράκτες συμπληρώματος και πήξης, η αλληλεπίδραση

ECM-υποδοχέα και ο καρκίνος της ουροδόχου κύστης συνδέονταν σε μεγάλο βαθμό με τα ανιχνευθέντα διαφορικά εκφραζόμενα γονίδια.

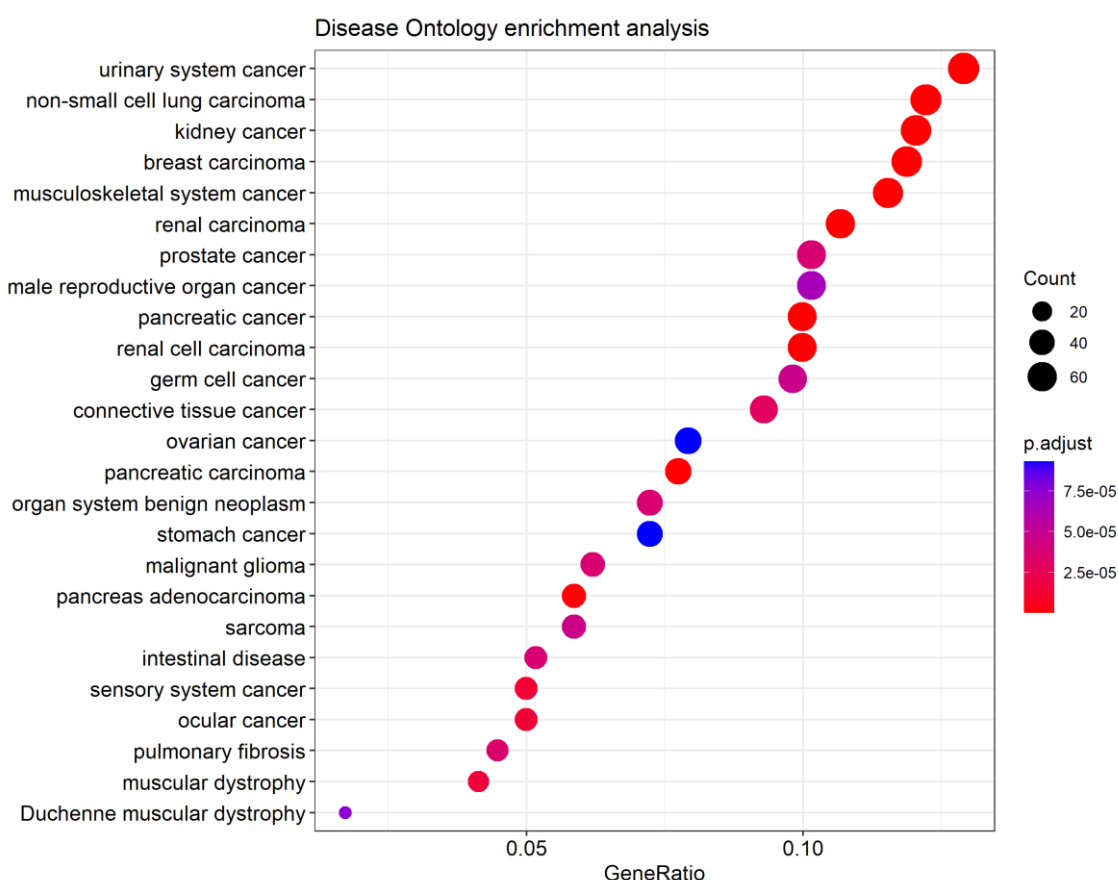


Εικόνα 37: (A) Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού μονοπατιών με βάση τη Reactome (B) Χάρτης των εμπλουτισμένων όρων της βάσης Reactome που παρουσιάζονται σε μορφή δικτύου.

Μέσω της ανάλυσης εμπλουτισμού με βάση τη REAC, εμπλουτίστηκαν 81 μονοπάτια. Και τα 25 κορυφαία μονοπάτια έδειξαν υψηλή στατιστική σημασία για τις αντίστοιχες οντότητές τους και

παρουσιάζονται στην Εικόνα 37Α. Οι πιο εμπλουτισμένοι όροι REAC ήταν η οργάνωση της εξωκυτταρικής μήτρας, τα σημεία ελέγχου του κυτταρικού κύκλου και οι τελεστές της Rho GTPάσης, οι οποίοι είναι βασικοί ρυθμιστές της δυναμικής του κυτταροσκελετού. Οι κατά ζεύγη ομοιότητες των εμπλουτισμένων όρων υπολογίστηκαν επίσης και απεικονίστηκαν (Εικόνα 37Β) σε χάρτη εμπλουτισμού. Σε αυτόν τον χάρτη ορίστηκαν δύο κύριες συστάδες, οι οποίες εμπλέκονται στις διαδικασίες του κυτταρικού κύκλου και της αντιγραφής και της εξωκυττάριας μήτρας, αντίστοιχα.

Η ανάλυση εμπλουτισμού με βάση την Οντολογία Ασθενειών έδειξε ότι υπήρχαν 185 εμπλουτισμένοι όροι, οι οποίοι συνδέονταν στενά με τα διαφορικά εκφραζόμενα γονίδια. Οι πρώτοι 25 εμπλουτισμένοι όροι της Οντολογίας Ασθενειών παρουσιάζονται στην Εικόνα 38. Αξίζει να σημειωθεί ότι ο υψηλότερος εμπλουτισμένος όρος DO ήταν ο καρκίνος του ουροποιητικού συστήματος, μαζί με τα μη μικροκυτταρικά καρκινώματα του πνεύμονα, των νεφρών, του μαστού και του μυοσκελετικού συστήματος στις επόμενες θέσεις.



Εικόνα 38: Οι 25 σημαντικότεροι όροι της ανάλυσης εμπλουτισμού με βάση την Οντολογία Ασθενειών (DO).

3.7 Ανάλυση του Δικτύου Αλληλεπιδράσεων Πρωτεΐνης – Πρωτεΐνης

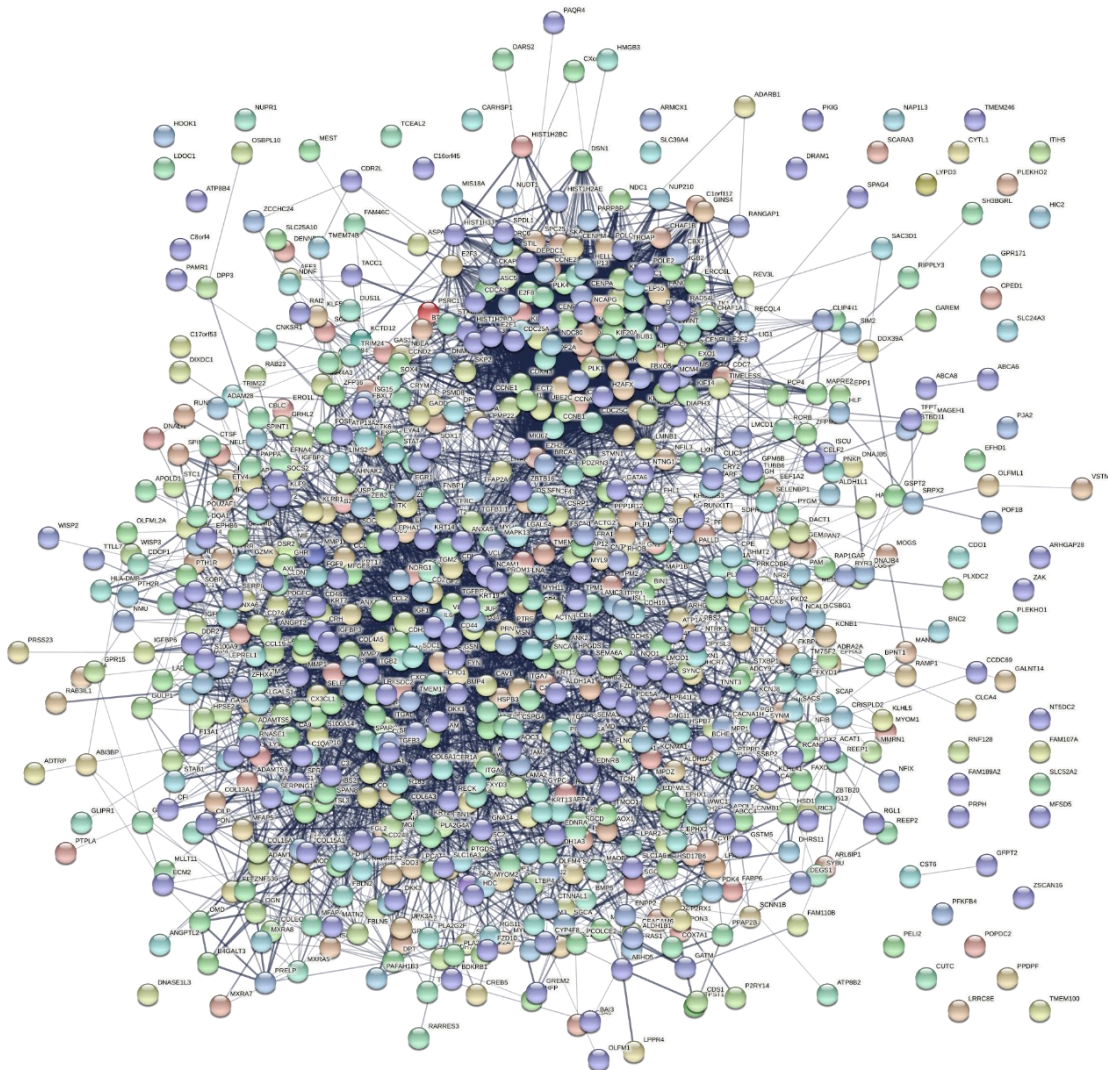
Το δίκτυο αλληλεπιδράσεων πρωτεΐνης – πρωτεΐνης (ΑΠΠ) των 815 διαφορικά εκφραζόμενων γονιδίων κατασκευάστηκε και απεικονίστηκε μέσω της βάσης δεδομένων STRING (Εικόνα 39). Το δίκτυο περιελάμβανε **813 κόμβους** και **8.260 ακμές**, με μέσο βαθμό κόμβου 20,3 και μέσο συντελεστή τοπικής ομαδοποίησης 0,382. Η τιμή-p εμπλουτισμού του δικτύου ΑΠΠ ήταν $< 10^{-16}$, υποδεικνύοντας ότι οι πρωτεΐνες ήταν βιολογικά συνδεδεμένες ως ομάδα.

Οι κόμβοι του δικτύου ΑΠΠ ταξινομήθηκαν εφαρμόζοντας τις 10 μεθόδους τοπολογικής ανάλυσης που αναφέρθηκαν στην Ενότητα 2.9 και, ως εκ τούτου, προέκυψαν 10 ταξινομημένες λίστες γονιδίων. Οι

μέθοδοι αυτές περιλάμβαναν τοπικούς καθώς και καθολικούς αλγορίθμους, όπως υλοποιούνται στο πρόσθετο *cytoHubba* του λογισμικού *Cytoscape*. Προκειμένου να προκύψει μία τελική λίστα ταξινομημένων γονιδίων, χρησιμοποιήθηκε η ανθεκτική συνθετική μέθοδος (Robust Rank Aggregation – RRA). Έτσι, η τελική ταξινομημένη λίστα γονιδίων, με βάση το πρόσθετο *cytoHubba*, περιελάμβανε 129 γονίδια – κόμβους, τα οποία παρουσίασαν τιμή- $p < 0,01$ (Πίνακας 4Α).

Επιπλέον, πραγματοποιήθηκε ανάλυση συστάδων (clusters) χρησιμοποιώντας το πρόσθετο *MCODE* του λογισμικού *Cytoscape* και διατηρήθηκαν μόνο οι συστάδες με βαθμολογία μεγαλύτερη από επτά. Σύμφωνα με αυτό, οι επιλεγμένες ενότητες περιλάμβαναν 133 γονίδια – κόμβους και 3.346 ακμές που ομαδοποιήθηκαν σε τρεις συστάδες. Η πρώτη συστάδα περιείχε 83 κόμβους και 3.045 ακμές, με βαθμολογία συστάδας 74,268. Η δεύτερη συστάδα περιείχε 23 κόμβους και 198 ακμές, με βαθμολογία συστάδας 18. Τέλος, η τρίτη συστάδα περιείχε 27 κόμβους και 103 ακμές, με βαθμολογία 7,923. Ολόκληρη η λίστα των γονιδίων – κόμβων για κάθε μία από τις τρεις συστάδες, που περιείχε συνολικά 133 γονίδια, παρουσιάζεται στον παρακάτω πίνακα (Πίνακας 4Β).

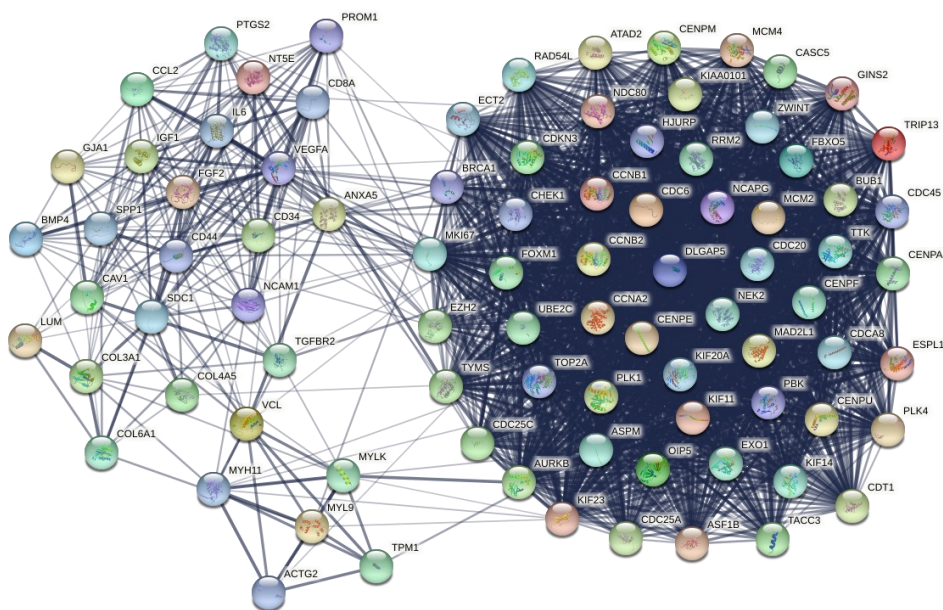
Ο τελικός κατάλογος των γονιδίων – κόμβων, με βάση τα πρόσθετα *cytoHubba* και *MCODE* του *Cytoscape*, περιείχε **87 κοινά γονίδια**, τα οποία θεωρήθηκαν ως οι σημαντικότεροι κόμβοι του δικτύου ΑΠΠ (Πίνακας 4Γ). Το δίκτυο ΑΠΠ αυτών των γονιδίων παρουσιάζεται στην Εικόνα 40.



Εικόνα 39: Το κατασκευασμένο δίκτυο ΑΠΠ, όπως οπτικοποιήθηκε από τη βάση δεδομένων STRING. Οι κόμβοι αντιπροσωπεύουν πρωτεΐνες και οι ακμές αντιπροσωπεύουν συσχετίσεις μεταξύ πρωτεΐνης-πρωτεΐνης. Το πάχος της γραμμής αντιπροσωπεύει το σκορ εμπιστοσύνης της εκάστοτε λειτουργικής συσχέτισης.

Πίνακας 4: Τα γονίδια – κόμβοι με βάση τις μεθόδους ανάλυσης των προσθέτων *cytoHubba* και *MCODE* του λογισμικού *Cytoscape*.

Α. Γονίδια της τελικής ταξινομημένης λίστα με βάση τις 10 τοπολογικές μεθόδους του προσθέτου <i>cytoHubba</i>			
<i>IL6, VEGFA, CCNB1, BRCA1, CCNA2, CD44, TYMS, CDH1, LMNB1, AURKB, EZH2, MKI67, KIF23, ECT2, MCM4, CDC6, PLK1, CDC25C, CDKN3, CENPA, MMP2, TOP2A, CENPE, PBK, NDC80, FOXM1, SPP1, IGF1, UBE2C, RRM2, KIF11, CHEK1, CD8A, CCNB2, ASPM, NCAM1, FLNA, LGALS4, ITPR1, DLGAP5, CDCA8, COL5A1, TIMELESS, CDC20, DMD, PPARGC1A, WNT5A, BUB1, KIF20A, EXO1, CDC25A, VCL, LUM, CCND2, CD34, MCM2, MAD2L1, HPGDS, ISL1, ESRP1, SKP2, NCAPG, CENPU, HJURP, CCL2, TPM1, CDH11, PLK4, FABP4, H2AFX, GJA1, DHCR7, PTGS2, MSN, ANXA5, COL6A1, TRIP13, OIP5, MYH11, KRT20, TTK, MYL9, CAV1, FBXO5, PROM1, BMP4, CDT1, KIAA0101, CCNE1, ANXA1, FGFR3, SNCA, ATAD2, ESPL1, FASN, NT5E, ZWINT, SDC1, FGF2, NEK2, ACTG2, KIF14, COL3A1, EPCAM, ASF1B, IGFBP5, RAD54L, CYP11B1, STMN1, COL4A5, ATF3, CASC5, CENPM, ERBB3, DNMT3B, ITGB2, ISG15, ANK2, CDC45, PLAT, TACC3, EGRI, MYLK, CTSG, GINS2, ITGA8, CENPF, TGFBR2, OGN</i>			
Β. Γονίδια που συμπεριλαμβάνονται στις 3 πρώτες συστάδες με βάση την ανάλυση του προσθέτου <i>MCODE</i>			
Συστάδα	Τιμή	Κόμβοι	Συστάδα γονιδίων
1	74,268	83	<i>PLK4, TRIP13, CDC45, PBK, RRM2, ERCC6L, CHAF1A, DEPDC1, DLGAP5, ASPM, E2F8, MAD2L1, CDCA8, CCNB1, BRCA1, FANCI, FBXO5, CENPA, KIAA0101, TK1, TACC3, DTL, CDCA3, HJURP, CENPE, ZWINT, ESPL1, POLQ, OIP5, CDC25C, ASF1B, CDKN3, POLE2, CCNB2, CHAF1B, EZH2, UBE2C, RAD54L, CDT1, MCM5, CDC20, TROAP, CKS2, NEK2, SPC25, MKI67, CHEK1, TTK, CDC6, GINS2, BUB1, CENPU, CCNE2, STIL, KIF14, TYMS, CDC7, MCM2, KIF23, KNTC1, SKA1, CASC5, CENPF, HELLS, NUSAP1, ATAD2, CEP55, NCAPG, MCM4, NDC80, ECT2, TOP2A, CENPM, CDC25A, MCM10, ORC1, KIF20A, AURKB, CCNA2, PLK1, EXO1, FOXM1, KIF11</i>
2	18	23	<i>CXCL12, PTGS2, BMP4, IL6, GJA1, CD34, FGF2, NES, PROM1, CD8A, VEGFA, CD44, SDC1, SPP1, ANXA5, NCAM1, SELP, CCL2, CCL5, IGF1, CSF1R, NT5E, SELE</i>
3	7,923	27	<i>TGFBI, COL6A2, THBS2, TPM1, MYH11, ACTG2, COL6A1, COL13A1, COL3A1, TGFBR2, VCL, FBLN2, COL4A5, CTSK, LYVE1, CLDN5, ANGPT2, LUM, MYL9, LEPREL1, TPM2, SPARC, MYLK, CAV1, ADAMTS5, TAGLN, FMOD</i>
Γ. Κοινά γονίδια μεταξύ των μεθόδων ανάλυσης των προσθέτων <i>cytoHubba</i> και <i>MCODE</i>			
<i>IL6, VEGFA, CCNB1, BRCA1, CCNA2, CD44, TYMS, AURKB, EZH2, MKI67, KIF23, ECT2, MCM4, CDC6, PLK1, CDC25C, CDKN3, CENPA, TOP2A, CENPE, PBK, NDC80, FOXM1, SPP1, IGF1, UBE2C, RRM2, KIF11, CHEK1, CD8A, CCNB2, ASPM, NCAM1, DLGAP5, CDCA8, CDC20, BUB1, KIF20A, EXO1, CDC25A, VCL, LUM, CD34, MCM2, MAD2L1, NCAPG, CENPU, HJURP, CCL2, TPM1, PLK4, GJA1, PTGS2, ANXA5, COL6A1, TRIP13, OIP5, MYH11, TTK, MYL9, CAV1, FBXO5, PROM1, BMP4, CDT1, KIAA0101, ATAD2, ESPL1, NT5E, ZWINT, SDC1, FGF2, NEK2, ACTG2, KIF14, COL3A1, ASF1B, RAD54L, COL4A5, CASC5, CENPM, CDC45, TACC3, MYLK, GINS2, CENPF, TGFBR2</i>			



Εικόνα 40: Το δίκτυο ΑΠΠ των τελικών 87 βασικών γονιδίων – κόμβων. Οι κόμβοι αντιπροσωπεύουν τις πρωτεΐνες και οι ακμές αντιπροσωπεύουν τις συσχετίσεις μεταξύ πρωτεΐνης-πρωτεΐνης. Το πάχος της γραμμής αντιπροσωπεύει το σκορ εμπιστοσύνης μιας λειτουργικής συσχέτισης.

3.8 Ανάλυση Δικτύων Συνέκφρασης Γονιδίων

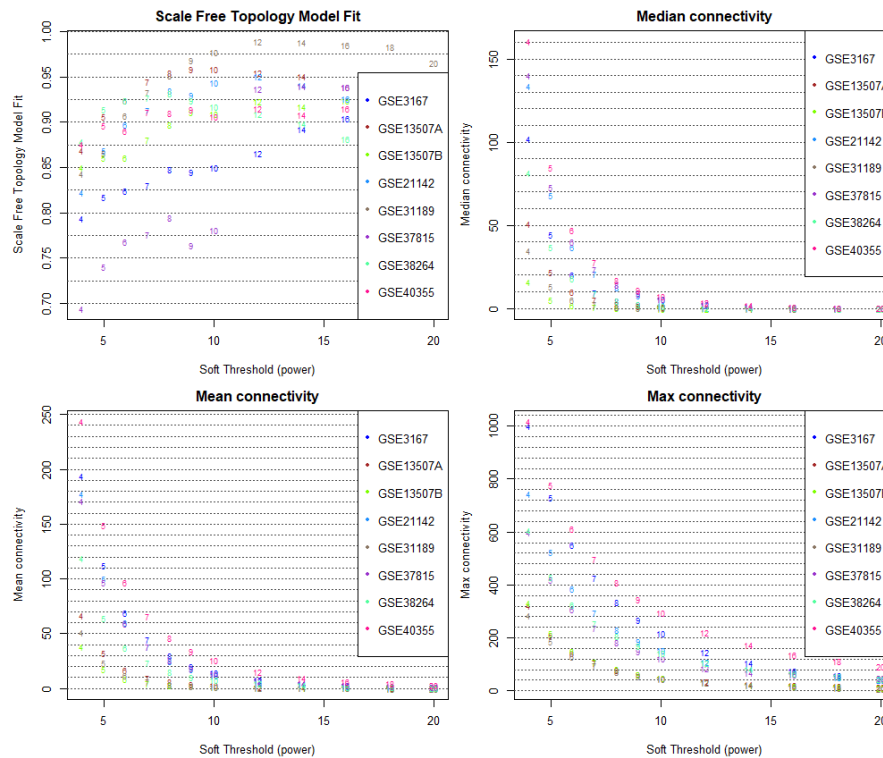
Προκειμένου να εντοπιστούν τα γονίδια, από το σύνολο των 8.201 κοινών γονιδίων σε όλες τις πλατφόρμες, που συνδέονται άμεσα με τον KOK, διεξήχθη μια ανάλυση δικτύων σταθμισμένης γονιδιακής συνέκφρασης (WGCNA), η οποία περιλάμβανε μόνο τα σύνολα δεδομένων που περιείχαν περισσότερα από 20 δείγματα. Με βάση τα ιεραρχικά δέντρα ομαδοποίησης (hierarchical clustering trees), δεν εντοπίστηκαν ακραίες τιμές, γεγονός το οποίο ήταν αναμενόμενο καθώς αυτές είχαν αφαιρεθεί στα προηγούμενα βήματα (Ενότητα 3.2). Μετά τη διόρθωση των επιδράσεων των φαινομένων παρτίδας, μέσω της συνάρτησης *sna* του ομώνυμου πακέτου R/Bioconductor, χρησιμοποιήθηκαν συνολικά 482 δείγματα, καταναμημένα σε οκτώ σύνολα δεδομένων, στην παρούσα συναινετική ανάλυση τοπολογίας δικτύου.

Επιλέχθηκε το επτά ως η κατάλληλη συναινετική δύναμη ελαφριού κατωφλίου (soft-thresholding power), καθώς αυτή είναι η χαμηλότερη δύναμη όπου πληρούνται δύο προϋποθέσεις: ο δείκτης προσαρμογής της τοπολογίας χωρίς κλίμακα (scale-free topology) να φτάνει το 0,77 και οι μέσες μετρήσεις συνδεσιμότητας να βρίσκονται κάτω από 30 (Εικόνα 41). Μαζί με τη δύναμη κατωφλίου, ορίστηκε το 30 ως το ελάχιστο μέγεθος της ομάδας γονιδίων και το 0,25 ως το ύψος, για τον αλγόριθμο δυναμικής κοπής δέντρων. Βάσει των ανωτέρω κριτηρίων ελήφθησαν οκτώ συναινετικές ομάδες συνέκφρασης γονιδίων. Ο αριθμός των γονιδίων που συμπεριλήφθηκαν σε κάθε ομάδα (module) κυμαινόταν από 34 έως 1.901 (Εικόνα 42). Η γκρι ομάδα περιείχε 4.898 γονίδια που δεν μπορούσαν να αντιστοιχιστούν σε καμία άλλη ομάδα.

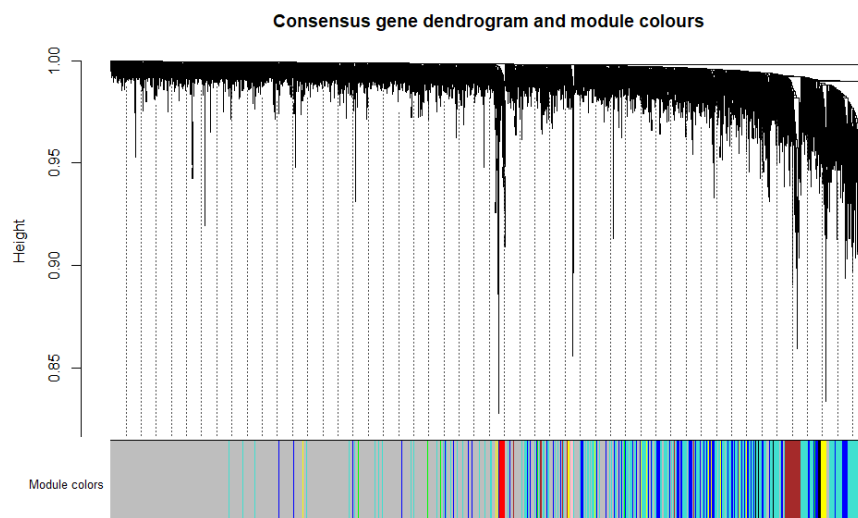
Μετά τον προσδιορισμό των γονιδιακών ομάδων, υπολογίστηκαν οι συντελεστές συσχέτισης κάθε ομάδας και για τα οκτώ σύνολα δεδομένων (Εικόνα 43). Με σκοπό να προκύψει ένας συναινετικός θερμικός χάρτης συσχέτισης ομάδας – χαρακτηριστικού, διατηρήθηκαν μόνο οι ομάδες οι οποίες είχαν σταθερό πρόσημο του συντελεστή συσχέτισης σε όλα τα σύνολα δεδομένων. Για τις ομάδες αυτές, η χαμηλότερη απόλυτη τιμή του συντελεστή συσχέτισης σε όλα τα σύνολα δεδομένων και η υψηλότερη τιμή-*p* αποδόθηκαν ως συναινετικός συντελεστής συσχέτισης και σημαντικότητα της εκάστοτε ομάδας, αντίστοιχα. Στις υπόλοιπες ομάδες αποδόθηκε η μηδενική τιμή ως η συναινετική συσχέτιση (Εικόνα 44). Το πλεονέκτημα του συναινετικού θερμικού χάρτη συσχέτισης είναι ότι διατηρεί τις σχέσεις μεταξύ ομάδας και χαρακτηριστικών που είναι παρούσες σε όλα τα σύνολα δεδομένων και, ως εκ τούτου, μπορεί να θεωρηθεί κατά μία έννοια ως ένα είδος επικύρωσης. Πρέπει να σημειωθεί ότι όλοι οι συντελεστές συσχέτισης, καθώς και οι τιμές σημαντικότητας μεταξύ ομάδων και χαρακτηριστικών, ήταν πολύ χαμηλές για το σύνολο δεδομένων GSE31189, πιθανώς λόγω των ισχυρών επιδράσεων φαινομένων παρτίδας που παρέμειναν παρόλη τη διόρθωση. Ως εκ τούτου, το εν λόγω σύνολο δεδομένων δε συμπεριλήφθηκε στη συναινετική συσχέτιση ομάδων – χαρακτηριστικών. Με βάση τον τελικό θερμικό χάρτη των συσχετίσεων ομάδων – χαρακτηριστικών, προσδιορίστηκε ότι η τιρκουάζ ($\text{cor} = -0,68$, $\text{τιμή-p} = 3\text{e-}08$), η καφέ ($\text{cor} = -0,65$, $\text{τιμή-p} = 1\text{e-}07$), η μαύρη ($\text{cor} = -0,54$, $\text{τιμή-p} = 2\text{e-}04$), η μπλε ($\text{cor} = 0,71$, $\text{τιμή-p} = 9\text{e-}11$), η πράσινη ($\text{cor} = 0,66$, $\text{τιμή-p} = 4\text{e-}08$) και η κίτρινη ($\text{cor} = 0,64$, $\text{τιμή-p} = 3\text{e-}08$) ομάδα συσχετίστηκαν περισσότερο με τον KOK ($\text{τιμή-p} < 0,01$) και χαρακτηρίστηκαν ως **βασικές ομάδες – κλειδιά** (Εικόνα 44). Η τιρκουάζ ομάδα περιείχε 1.901 γονίδια, η καφέ ομάδα περιελάμβανε 764 γονίδια, η μαύρη 231 γονίδια, η μπλε ομάδα περιελάμβανε 139 γονίδια, η πράσινη ομάδα περιελάμβανε 115 γονίδια και, τέλος, η κίτρινη ομάδα ενσωμάτωνε 58 γονίδια.

Στη συνέχεια, υπολογίστηκαν για κάθε ένα γονίδιο των βασικών ομάδων οι γονιδιακές σημαντικότητες (Gene Significances – GS) και οι συνδρομές στους στα μέλη της ομάδας (Module Memberships – MM). Τα κριτήρια για τον προσδιορισμό των γονιδίων – κόμβων που σχετίζονται σε μεγάλο βαθμό με τον KOK ορίστηκαν ως εξής: μετα-*Z*-τιμή της συνδρομής στην ομάδα (MM) και της γονιδιακής σημαντικότητας (GS) στο ανώτερο ή κατώτερο τεταρτημόριο κάθε ομάδας. Εντοπίστηκαν έτσι 815 γονίδια – κόμβοι από την τιρκουάζ ομάδα, 345 γονίδια – κόμβοι από την καφέ ομάδα, 91 γονίδια από

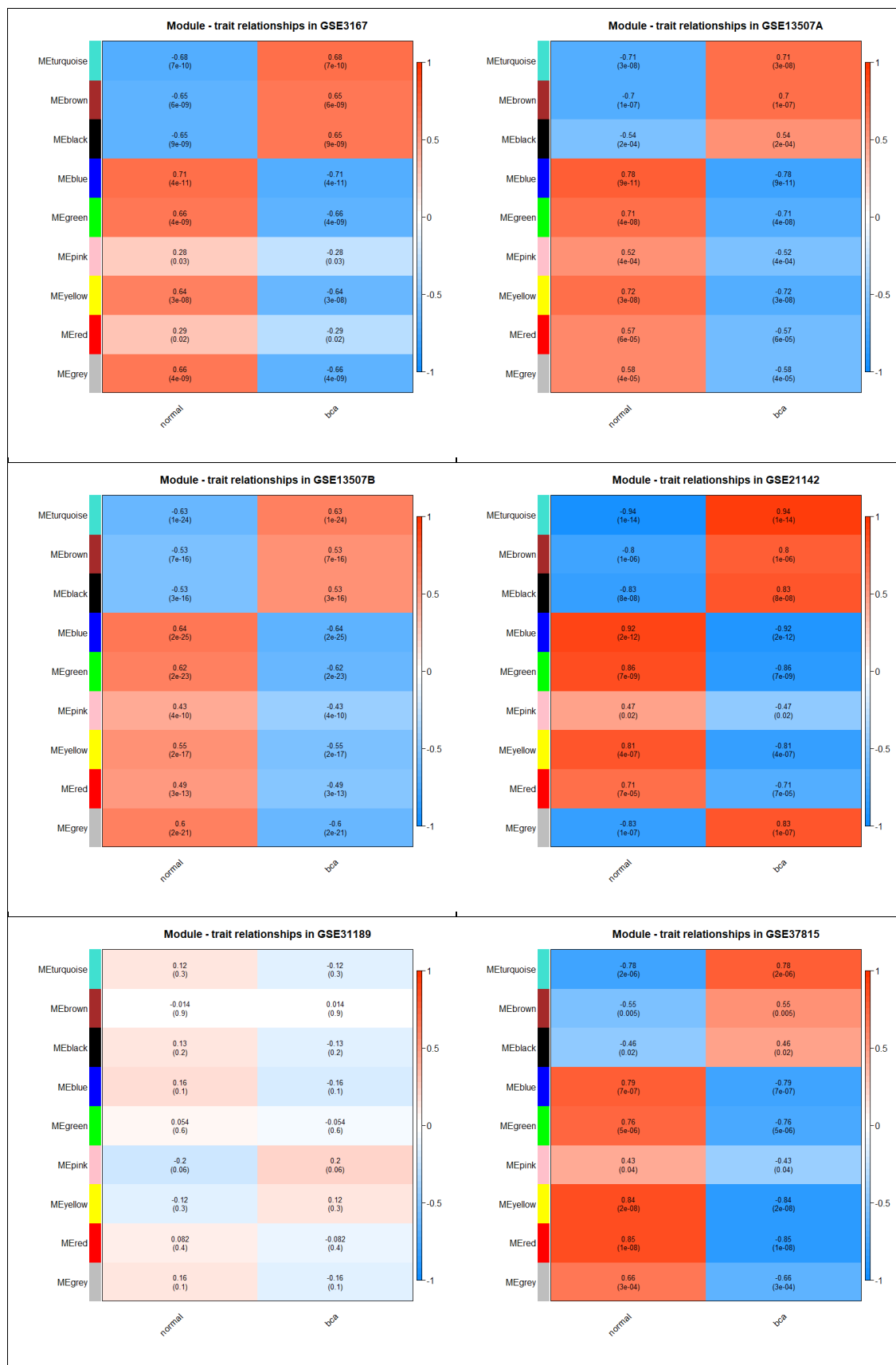
τη μαύρη ομάδα, 49 γονίδια από τη μπλε ομάδα, 50 γονίδια από την πράσινη ομάδα και 21 γονίδια από την κίτρινη ομάδα, τα οποία πληρούσαν τα κριτήρια ένταξης (Εικόνα 45). Τέλος, συνδυάστηκαν τα γονίδια – κόμβοι κάθε βασικής ενότητας με τα γονίδια – κόμβους που προέκυψαν από την ανάλυση δικτύου ΑΠΠ και προσδιορίστηκαν τα **σημαντικά γονίδια – κόμβοι** (ή «κλειδιά») (**key hub genes**) της ανάλυσής μας. Αυτά τα γονίδια – κόμβοι – κλειδιά παρατίθενται στον παρακάτω πίνακα (Πίνακας 5).

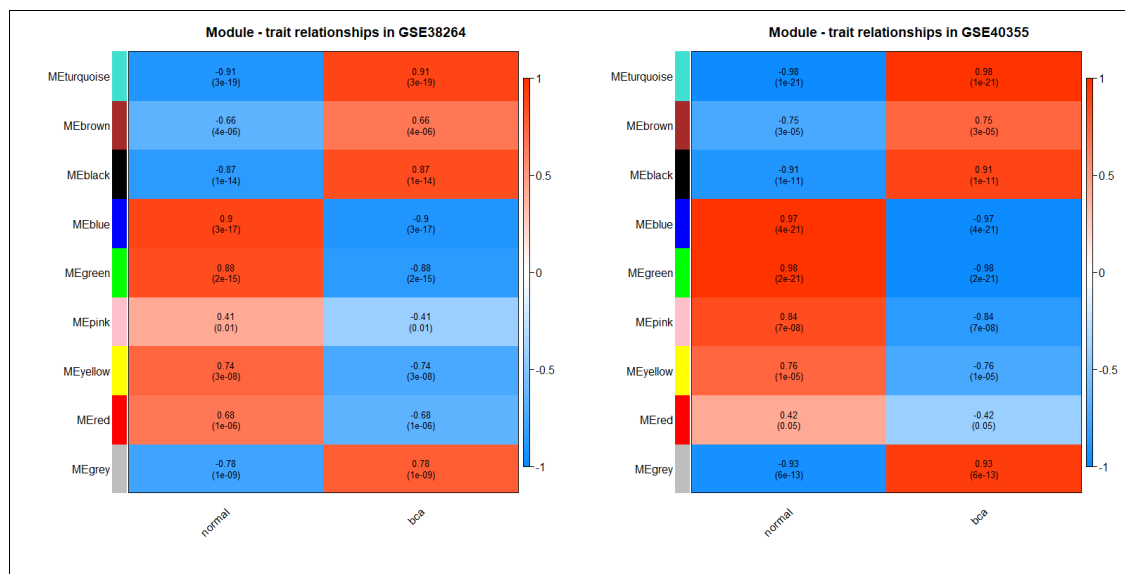


Εικόνα 41: Σύνοψη των βασικών δεικτών του δικτύου σε συνάρτηση με τη δύναμη ελαφριού κατωφλίου για κάθε ένα από τα οκτώ επιλεγμένα σύνολα δεδομένων τα οποία παρουσιάζονται με διαφορετικά χρώματα.

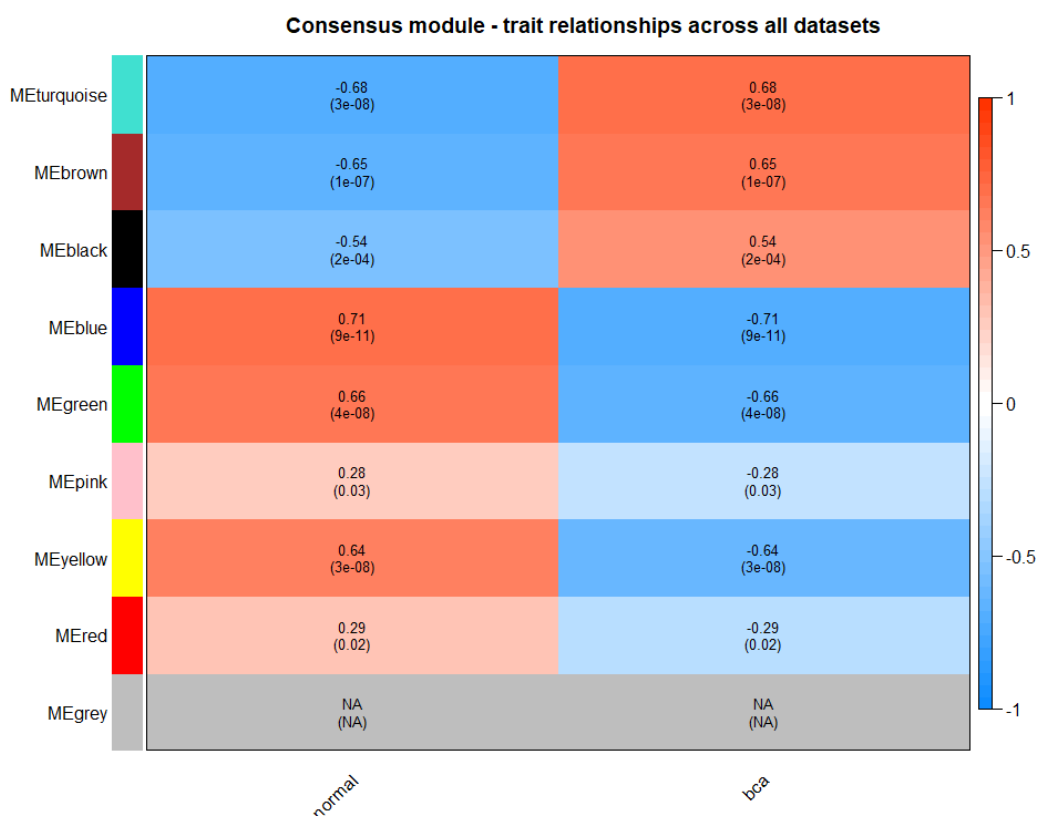


Εικόνα 42: Το δενδρόγραμμα της ιεραρχικής ομαδοποίησης των 8.201 γονιδίων με σκοπό τον προσδιορισμό συναινετικών ομάδων με βάση τη συναινετική τοπολογική επικάλυψη. Τα γονίδια μιας κοινής ομάδας λαμβάνουν το ίδιο χρώμα, όπως παρουσιάζεται στη χρωματική ζώνη κάτω από το δενδρόγραμμα. Τα γονίδια που δεν αντιστοιχούν σε καμία από τις ομάδες έχουν γκριζό χρώμα.

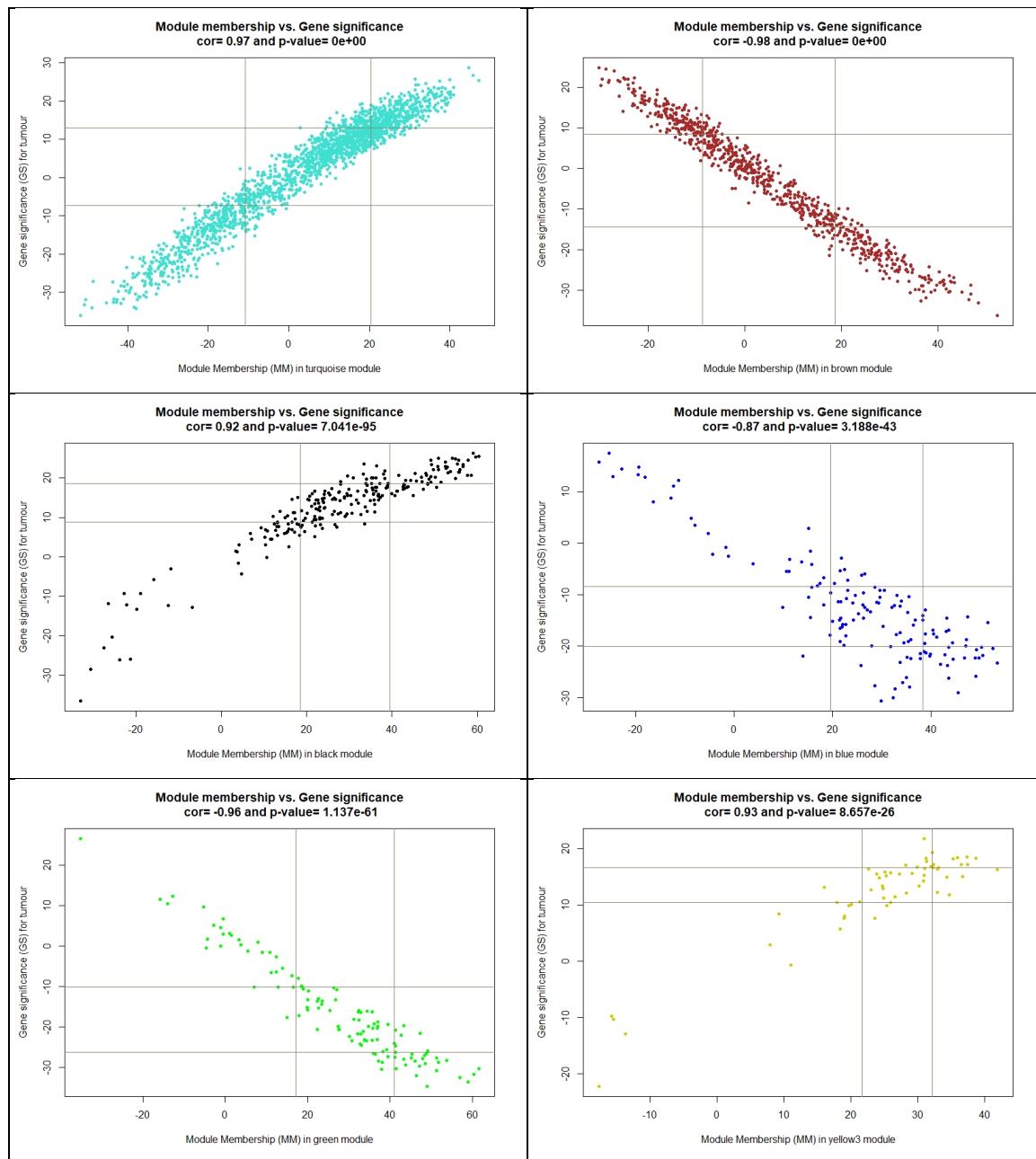




Εικόνα 43: Οι θερμικοί χάρτες των σχέσεων των ιδιοδιανυσμάτων των συναινετικών ομάδων και των φαινοτυπικών χαρακτηριστικών για καθένα από τα οκτώ σύνολα δεδομένων. Κάθε γραμμή αντιστοιχεί στο ιδιοδιάνυσμα της συναινετικής ομάδας και κάθε στήλη αντιστοιχεί σε ένα φαινοτυπικό χαρακτηριστικό. Κάθε κελί περιέχει την αντίστοιχη συσχέτιση (που κυμαίνεται από μπλε σε κόκκινο) και την τιμή-p.



Εικόνα 44: Ο συναινετικός θερμικός χάρτης των σχέσεων των ιδιοδιανυσμάτων των συναινετικών ομάδων και των φαινοτυπικών χαρακτηριστικών. Κάθε γραμμή αντιστοιχεί στο ιδιοδιάνυσμα της συναινετικής ομάδας και κάθε στήλη αντιστοιχεί σε ένα φαινοτυπικό χαρακτηριστικό. Κάθε κελί περιέχει την αντίστοιχη συσχέτιση (που κυμαίνεται από μπλε σε κόκκινο) και την τιμή-p.



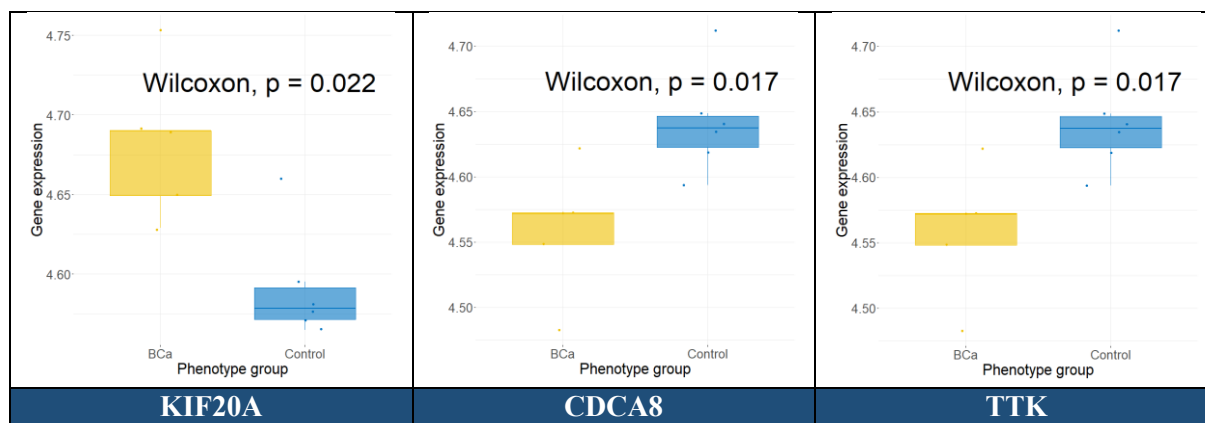
Εικόνα 45: Διαγράμματα διασποράς της γονιδιακής σημαντικότητας (GS) για τον φαινότυπο του καρκίνου και της συμμετοχής στην ομάδα (MM) για τις βασικές ομάδες. Οι γραμμές υποδεικνύουν το ανώτερο και το κατώτερο τεταρτημόριο.

Πίνακας 5: Τα γονίδια – κόμβοι – κλειδιά (key hub genes) της παρούσας μελέτης, τα οποία αποτελούν την τομή μεταξύ των γονιδίων – κόμβων που προέκυψαν από την ανάλυση του δικτύου ΑΙΠΠ και των γονιδίων – κόμβων που προέκυψαν από την ανάλυση των δικτύων συνέκφρασης των γονιδίων.

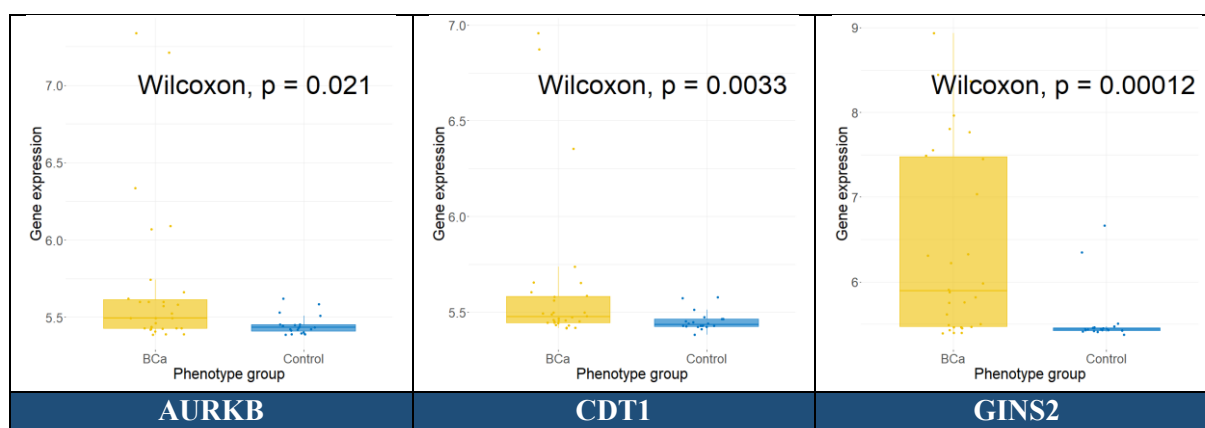
Ομάδα	Κοινά γονίδια - κόμβοι
Τυρκουάζ	<i>ACTG2, ANXA5, AURKB, BUB1, CD34, CD44, CDC25A, CDT1, CENPM, ESPL1, EXO1, FGF2, GINS2, KIF20A, NCAM1</i>
Καφέ	<i>CAV1, COL3A1, COL4A5, IGF1, LUM, MYLK, PROM1, SDC1, SPPI, TPM1, VCL, VEGFA</i>
Μαύρη	<i>ASPM, CCNA2, CCNB1, CCNB2, CDC20, CDC45, CDCA8, CDKN3, CENPA, CENPF, CENPU, DLGAP5, ECT2, EZH2, FOXM1, HJURP, KIF11, KIF14, KIF23, MCM2, MCM4, MKI67, NCAPG, NDC80, NEK2, PBK, PLK4, RAD54L, TOP2A, TTK, UBE2C, ZWINT</i>
Μπλε	--
Πράσινη	<i>COL6A1, MYH11</i>
Κίτρινη	--

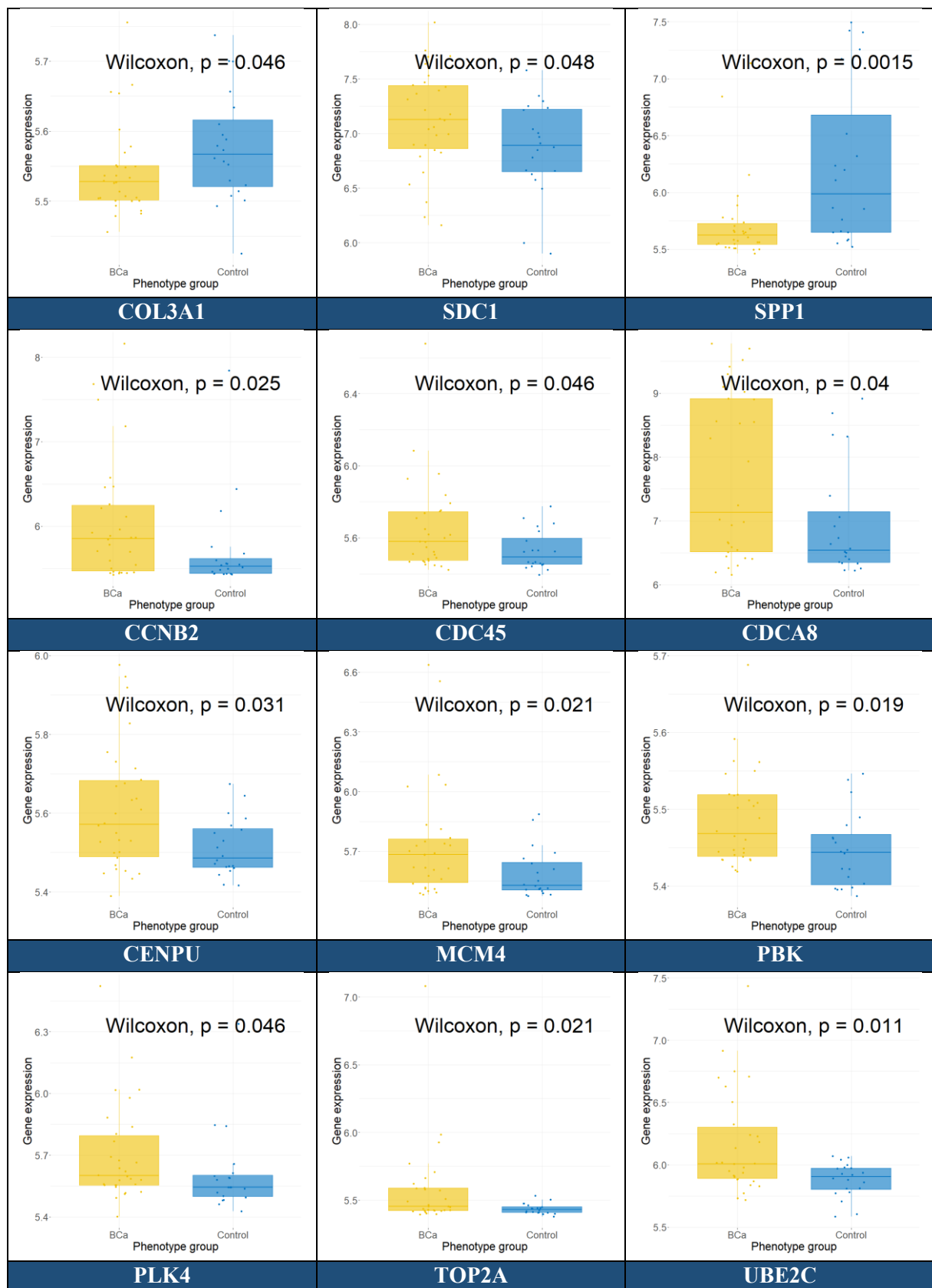
3.9 Διαφορική Έκφραση Γονιδίων στα Ούρα και στο Πλάσμα Αίματος

Τα πρωτογενή δεδομένα γονιδιακής έκφρασης των δειγμάτων ούρων, από τις σειρές GSE51843 και GSE68020, που χαρακτηρίστηκαν από την πλατφόρμα Illumina Human HT12 v4 BeadChip (GPL10558), ελήφθησαν και υποβλήθηκαν σε κατάλληλη προεπεξεργασία. Τα πρώτο σύνολο δεδομένων αποτελείται από πέντε δείγματα ούρων από ασθενείς με ΚΟΚ και έξι δείγματα ούρων ελέγχου από υγιή άτομα. Το δεύτερο σύνολο αποτελείται από 30 δείγματα ούρων από ασθενείς με ΚΟΚ και 20 δείγματα ούρων ελέγχου από υγιή άτομα, αντίστοιχα. Σε κάθε σύνολο δεδομένων εφαρμόστηκε το μη παραμετρικό τεστ του Wilcoxon για την αξιολόγηση της στατιστικής σημαντικότητας μεταξύ των ομάδων δειγμάτων ΚΟΚ και ελέγχου για όλα τα γονίδια – κλειδιά. Τα γονίδια – κλειδιά με τιμή- $p < 0,05$ προσδιορίστηκαν ως σημαντικά διαφοροποιημένα μεταξύ των δειγμάτων ούρων ΚΟΚ και ελέγχου. Αυτά τα γονίδια περιλάμβαναν τα *KIF20A*, *CDCA8* και *TTK* για το σύνολο δεδομένων GSE51843 και τα *AURKB*, *CDT1*, *GINS2*, *COL3A1*, *SDC1*, *SPPI*, *CCNB2*, *CDC45*, *CDCA8*, *CENPU*, *MCM4*, *PBK*, *PLK4*, *TOP2A* και *UBE2C* για το σύνολο δεδομένων GSE68020. Τα αποτελέσματα του μη παραμετρικού τεστ Wilcoxon για αυτά τα γονίδια παρουσιάζονται με τη μορφή θηκογραμμάτων (boxplots) στην Εικόνα 46 και στην Εικόνα 47.



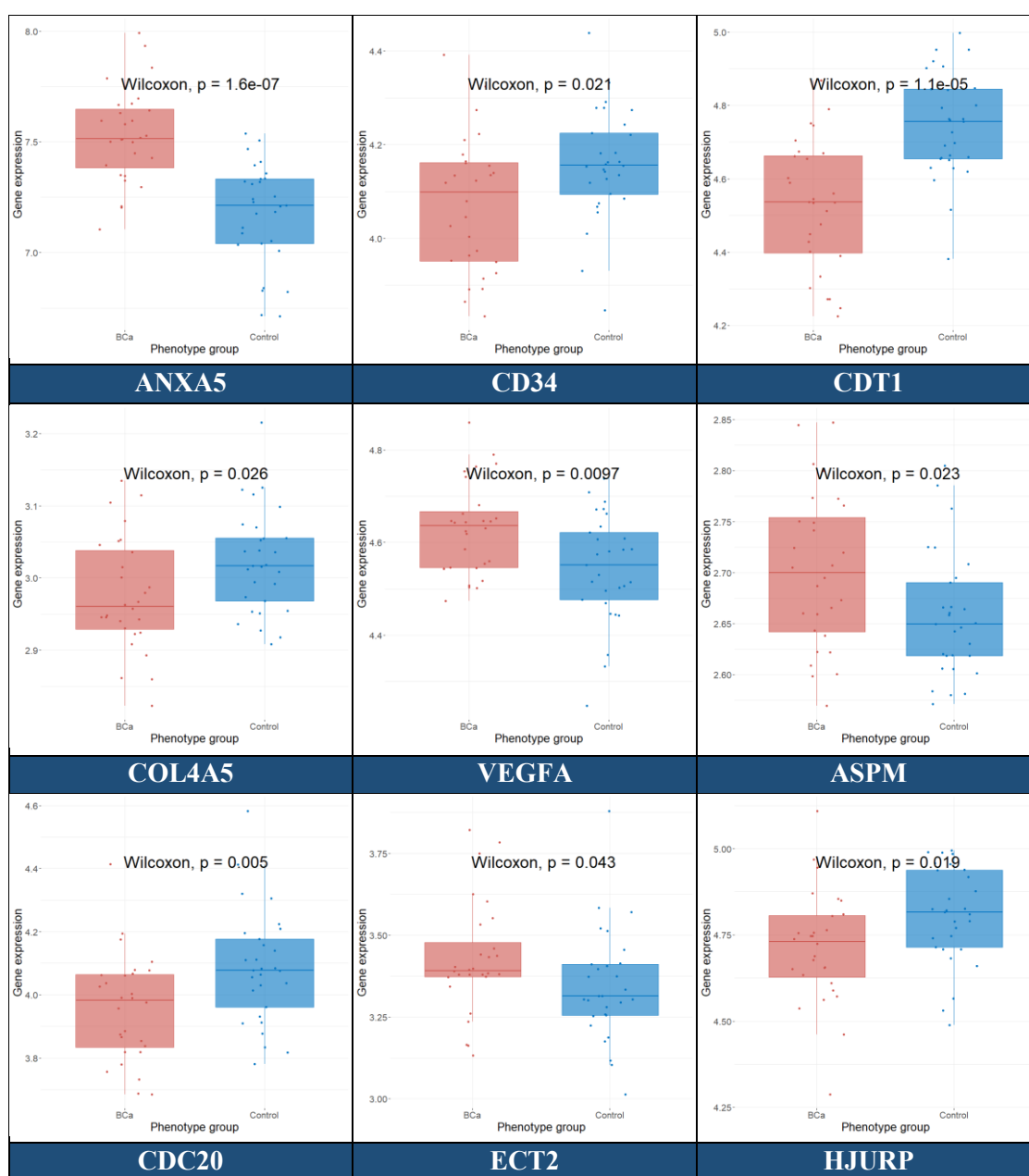
Εικόνα 46: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα ούρων με τη χρήση του μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE51843 (n = 11).

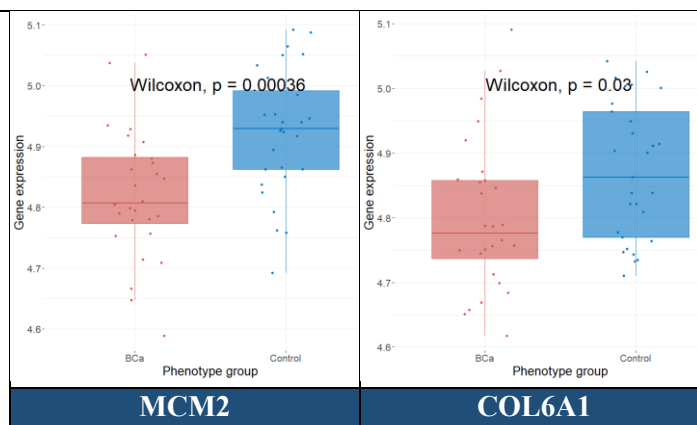




Εικόνα 47: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα ούρων με τη χρήση μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE68020 (n = 50).

Τα πρωτογενή δεδομένα γονιδιακής έκφρασης των δειγμάτων πλάσματος αίματος, από τη σειρά GSE138118, που χαρακτηρίστηκαν από την πλατφόρμα Affymetrix Human Gene 2.1 ST Array (GPL17692), ελήφθησαν και υποβλήθηκαν σε κατάλληλη προεπεξεργασία. Το συγκεκριμένο σύνολο δεδομένων αποτελείται από 28 δείγματα πλάσματος αίματος από ασθενείς με KOK και 29 δείγματα πλάσματος αίματος από υγιή άτομα ελέγχου. Το σύνολο δεδομένων διορθώθηκε από την επίδραση των φαινομένων παρτίδας. Το μη παραμετρικού τεστ Wilcoxon διεξήχθη για να αξιολογηθεί η στατιστική σημαντικότητα μεταξύ των δειγμάτων αίματος με KOK και ελέγχου για όλα τα γονίδια – κλειδιά. Τα γονίδια – κλειδιά με τιμή- $p < 0,05$ θεωρήθηκαν ως σημαντικά διαφοροποιημένα μεταξύ των δειγμάτων αίματος με KOK και των δειγμάτων αίματος ελέγχου. Τα γονίδια αυτά περιλάμβαναν τα *ANXA5*, *CD34*, *CDT1*, *COL4A5*, *VEGFA*, *ASPM*, *CDC20*, *ECT2*, *HJURP*, *MCM2* και *COL6A1*. Τα αποτελέσματα του μη παραμετρικού τεστ Wilcoxon για αυτά τα γονίδια παρουσιάζονται σε ηχογράμματα στην Εικόνα 48.





Εικόνα 48: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά σε δείγματα πλάσματος αίματος με τη χρήση του μη παραμετρικού τεστ Wilcoxon στο σύνολο δεδομένων GSE138118 (n = 57).

3.10 Προγνωστικά Γονίδια για τον Καρκίνο της Ουροδόχου Κύστης

Με σκοπό να προσδιοριστεί ποια από τα 61 γονίδια – κλειδιά σχετίζονται άμεσα με την κλινική έκβαση – πρόγνωση των ασθενών με ΚΟΚ, αξιολογήθηκαν περαιτέρω τα γονίδια αυτά εφαρμόζοντας μονο-μεταβλητή ανάλυση παλινδρόμησης Cox σε δεδομένα επιβίωσης 165 ασθενών με ΚΟΚ διαφόρων σταδίων (GSE13507). Η μονο-μεταβλητή ανάλυση παλινδρόμησης Cox έδειξε ότι 46 γονίδια είχαν στατιστικά σημαντική συσχέτιση με τη συνολική επιβίωση (Πίνακας 6). Εν συνεχεία, πραγματοποιήθηκε ανάλυση παλινδρόμησης LASSO Cox με 10-πλη διασταυρούμενη επικύρωση, για να μειωθεί περαιτέρω ο αριθμός των σημαντικών γονιδίων και να εντοπιστούν εκείνα που είχαν υψηλή συσχέτιση με τη συνολική επιβίωση των ασθενών ΚΟΚ. Εντοπίστηκαν έτσι εννέα γονίδια, συγκεκριμένα τα *ACTG2*, *ASPM*, *CDC48*, *COL3A1*, *COL4A5*, *FOXM1*, *MKI67*, *PLK4* και *SPPI*. Η πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox έδειξε ότι η έκφραση των ***COL3A1***, ***FOXM1*** και ***PLK4*** συνδέονταν σε μεγάλο βαθμό και ανεξάρτητα με την πρόγνωση των ασθενών με ΚΟΚ (Πίνακας 6) και χρησιμοποιήθηκε για τον υπολογισμό του συντελεστή κάθε γονιδίου στο προγνωστικό μοντέλο. Έτσι, κατασκευάστηκε ένα προγνωστικό μοντέλο υπογραφής τριών γονιδίων. Βάσει αυτού, υπολογίστηκε ένα προγνωστικό σκορ κινδύνου για κάθε ασθενή του συνόλου εκπαίδευσης (GSE13507) με βάση τα διαφορετικά επίπεδα έκφρασης των τριών γονιδίων, χρησιμοποιώντας τον προγνωστικό δείκτη (ΠΔ):

$$\text{ΠΔ} = 0.5405 \cdot \text{εκφ}_{\text{COL3A1}} + 1.6748 \cdot \text{εκφ}_{\text{FOXM1}} - 0.9583 \cdot \text{εκφ}_{\text{PLK4}}$$

Το διάγραμμα δάσους (forest plot) του προγνωστικού μοντέλου απεικονίζεται στην Εικόνα 49C. Η ανάλυση συσχέτισης της συν-έκφρασης για καθένα από τα τρία ζεύγη γονιδίων του παραπάνω μοντέλου πραγματοποιήθηκε μέσω της πλατφόρμας GEPIA2 και έδειξε ότι κανένα ζεύγος γονιδίων δεν είχε συντελεστή συσχέτισης Pearson μεγαλύτερο από 0,6 στον ΚΟΚ (Εικόνα 50).

Υπολογίζοντας τον προγνωστικό δείκτη για κάθε έναν από τους ασθενείς του συνόλου εκπαίδευσης και ομαδοποιώντας τους με βάση τη διάμεση τιμή, ο χρόνος επιβίωσης για τους ασθενείς υψηλού κινδύνου (κακή πρόγνωση) (n = 82) ήταν σημαντικά χειρότερος (τιμή-p < 0,0001) από εκείνον των ασθενών χαμηλού κινδύνου (καλή πρόγνωση) (n = 83), όπως φαίνεται από την ανάλυση των καμπύλων Kaplan – Meier (Εικόνα 49A). Επιπλέον, η προγνωστική υπογραφή των τριών γονιδίων αξιολογήθηκε και ως προς την προγνωστική της ακρίβεια πραγματοποιώντας χρονοεξαρτώμενη ανάλυση ROC, σε συγκεκριμένους χρόνους έπειτα από τη διάγνωση, και συγκεκριμένα στο πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση. Η τιμή AUC για τους διάφορους αυτούς χρόνους αποκοπής ήταν 0,796, 0,779, 0,846 και 0,8, αντίστοιχα (Εικόνα 49B).

Επίσης, το προγνωστικό μοντέλο εξετάστηκε ως προς την απόδοση σε δύο ανεξάρτητα σύνολα δεδομένων. Στο πρώτο σύνολο δεδομένων (GSE32894), η ομάδα ασθενών χαμηλού κινδύνου (n = 112),

όπως υπολογίστηκε από το προγνωστικό μοντέλο, επέδειξε σημαντικά μεγαλύτερη συνολική επιβίωση (τιμή- $p < 0,0001$) σε αντίθεση με την ομάδα ασθενών υψηλού κινδύνου ($n = 112$) (Εικόνα 51A). Οι χρονοεξαρτώμενες καμπύλες ROC σχεδιάστηκαν και οι τιμές AUC για το πρώτο, το τρίτο, το πέμπτο και το δέκατο έτος ήταν 0,819, 0,86, 0,871 και 0,789, αντίστοιχα (Εικόνα 51B). Στο δεύτερο σύνολο δεδομένων (GSE32548), οι ομάδες ασθενών χαμηλού κινδύνου ($n = 65$) και υψηλού κινδύνου ($n = 66$) εμφάνισαν επίσης σημαντικά διαφορετικούς χρόνους επιβίωσης (τιμή- $p < 0,0001$) (Εικόνα 52A). Ομοίως, σχεδιάστηκαν οι χρονοεξαρτώμενες καμπύλες ROC και οι αντίστοιχες τιμές AUC ήταν 0,82, 0,806, 0,774 και 0,724 (Εικόνα 52B).

Τέλος, διερευνήθηκε κατά πόσον η γονιδιακή υπογραφή, η οποία αποτελούνταν από τα *COL3A1*, *FOXMI* και *PLK4*, θα μπορούσε να είναι προγνωστική για την συνολική επιβίωση ή την ελεύθερη νόσου επιβίωση των ασθενών με KOK της κοόρτης TCGA. Για τον σκοπό αυτό, εξετάστηκε η υπογραφή των τριών γονιδίων μέσω της ανάλυσης επιβίωσης Kaplan – Meier χρησιμοποιώντας την πλατφόρμα GEPIA2. Οι ασθενείς διαχωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου ($n = 201$ σε κάθε ομάδα) με βάση τη διάμεση τιμή έκφρασης για αυτά τα γονίδια και οι δύο ομάδες εμφάνισαν στατιστικά σημαντικά διαφορετικό χρόνο συνολικής επιβίωσης (τιμή- $p = 0,02$) (Εικόνα 53A). Όσον αφορά τον χρόνο ελεύθερης νόσου επιβίωσης, το κατώτερο 15% και το ανώτερο 85% των ταξινομημένων τιμών έκφρασης χρησιμοποιήθηκαν για τη διάκριση των ομάδων ασθενών σε χαμηλής και υψηλής έκφρασης, αντίστοιχα ($n = 61$ σε κάθε ομάδα). Η διαφορά μεταξύ των καμπυλών ελεύθερης νόσου επιβίωσης αυτών των ομάδων έκφρασης βρέθηκε κι αυτή στατιστικά σημαντική (τιμή- $p = 0,019$) (Εικόνα 53B).

Συμπερασματικά, η ανάλυση επιβίωσης Kaplan – Meir και οι τιμές AUC στις διάφορες χρονικές στιγμές αποκοπής έδειξαν ότι το μοντέλο των τριών γονιδίων κατέχει πολύ καλή προγνωστική ακρίβεια όσον αφορά την διακριτοποίηση της επιβίωσης των ασθενών με KOK. Έτσι, τα αποτελέσματα αυτά επικυρώνουν την προγνωστική αξία της συγκεκριμένης γονιδιακής υπογραφής.

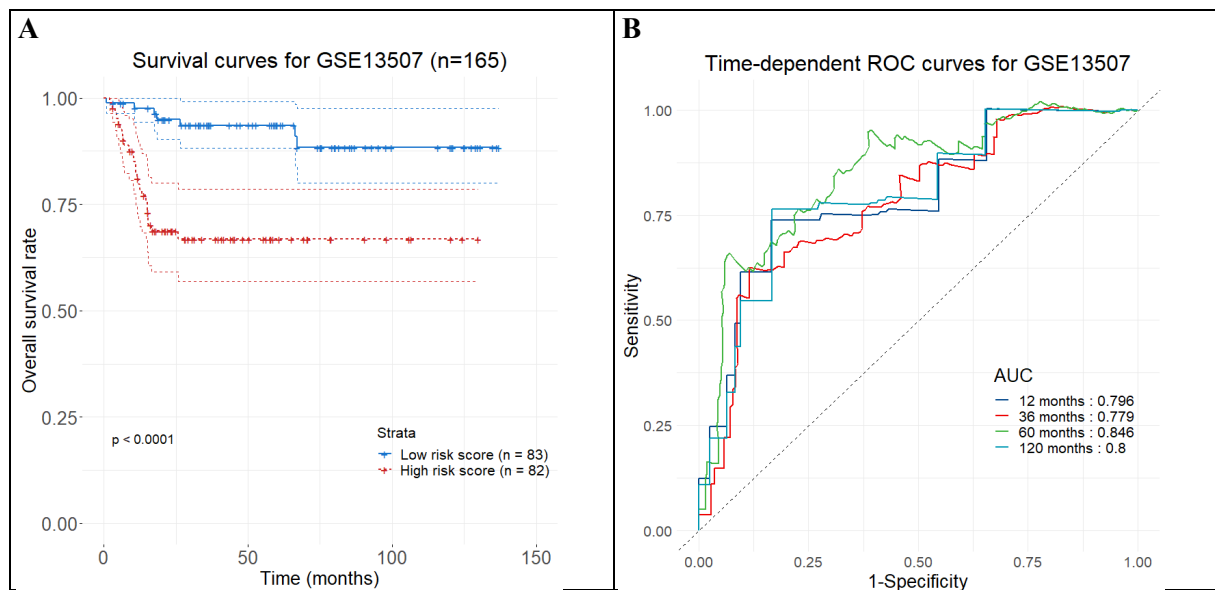
Πίνακας 6: Μονο-μεταβλητή και πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox των γονιδίων – κλειδιών που σχετίζονται με την επιβίωση των ασθενών με KOK

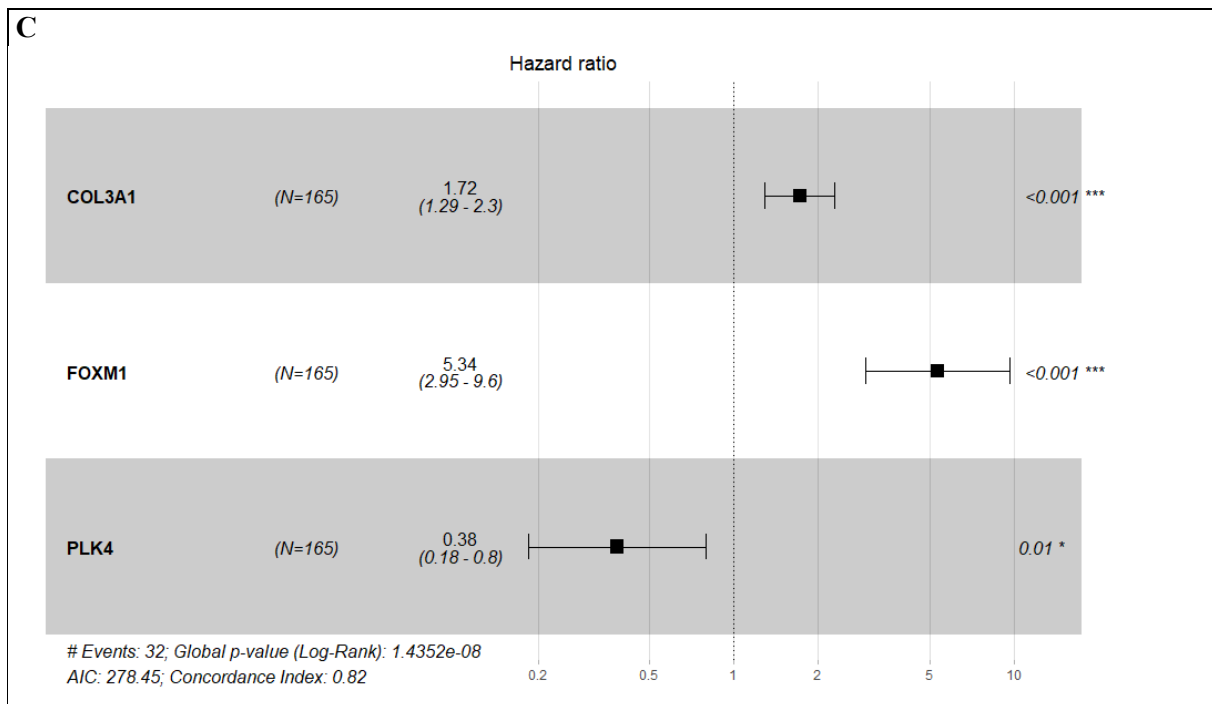
HR: λόγος κινδύνου (Hazard ratio), CI: Διάστημα εμπιστοσύνης (Confidence interval)

*τιμή- $p < 0.05$, **τιμή- $p < 0.01$, ***τιμή- $p < 0.001$, ****τιμή- $p < 0.0001$

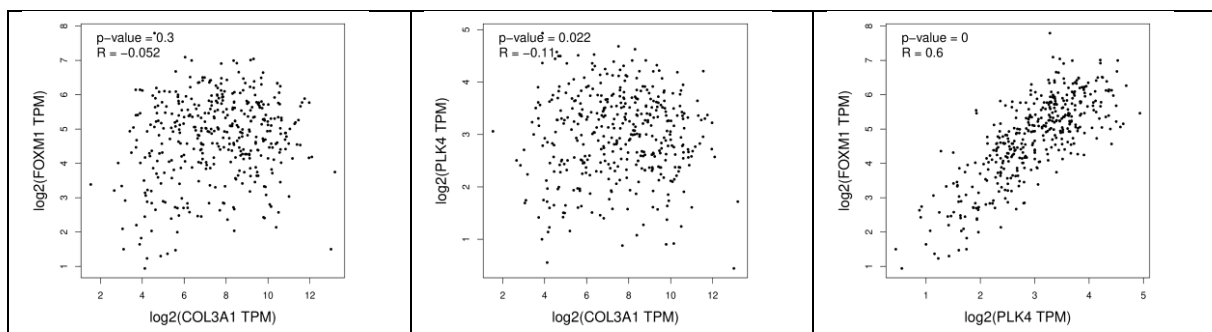
Γονίδιο	Μονο-μεταβλητή Ανάλυση		Πολυ-μεταβλητή Ανάλυση	
	HR (95% CI)	Τιμή- p	HR (95% CI)	Τιμή- p
ACTG2	1.2 (1 - 1.4)	1.40e-02 *	--	--
AURKB	2.1 (1.5 - 3)	3.30e-05 ****	--	--
BUB1	1.8 (1.2 - 2.6)	1.90e-03 **	--	--
CDC25A	3.1 (1.6 - 5.9)	9.80e-04 ***	--	--
CDT1	1.9 (1.4 - 2.6)	1.00e-04 ***	--	--
CENPM	2.1 (1.4 - 3.1)	1.90e-04 ***	--	--
ESPL1	2.8 (1.7 - 4.6)	5.30e-05 ****	--	--
EXO1	3.8 (1.9 - 7.7)	2.50e-04 ***	--	--
GINS2	1.8 (1.2 - 2.5)	1.40e-03 **	--	--
KIF20A	1.8 (1.3 - 2.5)	6.00e-04 ***	--	--
COL3A1	1.5 (1.1 - 1.9)	3.50e-03 **	1.72 (1.29 - 2.29)	0.000223 ***
COL4A5	0.66 (0.51 - 0.85)	1.70e-03 **	--	--
LUM	1.3 (1 - 1.6)	1.80e-02 *	--	--
SPP1	1.4 (1.1 - 1.7)	4.80e-03 **	--	--
ASPM	1.9 (1.4 - 2.6)	5.40e-05 ****	--	--
CCNA2	1.6 (1.2 - 2.3)	2.50e-03 **	--	--
CCNB1	1.7 (1.1 - 2.6)	1.00e-02 *	--	--
CCNB2	1.8 (1.3 - 2.4)	9.90e-05 ****	--	--
CDC20	1.7 (1.3 - 2.3)	1.90e-04 ***	--	--
CDC45	2.7 (1.7 - 4.2)	1.30e-05 ****	--	--
CDC48	2.2 (1.5 - 3.2)	2.40e-05 ****	--	--

CDKN3	2.1 (1.5 - 3)	3.40e-05 ****	--	--
CENPA	1.8 (1.3 - 2.5)	2.70e-04 ***	--	--
CENPF	2 (1.5 - 2.7)	7.50e-06 ****	--	--
CENPU	1.7 (1.1 - 2.6)	2.20e-02 *	--	--
DLGAP5	1.7 (1.2 - 2.3)	1.20e-03 **	--	--
ECT2	3 (1.5 - 6.2)	3.00e-03 **	--	--
EZH2	2.2 (1.4 - 3.4)	5.30e-04 ***	--	--
FOXM1	2.7 (1.8 - 4)	3.20e-06 ****	5.34 (2.95 - 9.64)	2.87e-08 ****
HJURP	2.1 (1.5 - 2.9)	4.30e-05 ****	--	--
KIF11	1.9 (1.3 - 2.8)	6.00e-04 ***	--	--
KIF14	2.4 (1.5 - 3.7)	1.40e-04 ***	--	--
KIF23	3 (1.5 - 5.7)	1.20e-03 **	--	--
MCM2	1.7 (1.2 - 2.3)	1.70e-03 **	--	--
MCM4	1.4 (1 - 1.9)	3.80e-02 *	--	--
MKI67	9.8 (4 - 24)	8.60e-07 ****	--	--
NCAPG	2.1 (1.5 - 3)	6.40e-05 ****	--	--
NDC80	1.5 (1.1 - 2.1)	1.30e-02 *	--	--
NEK2	2.5 (1.4 - 4.5)	2.40e-03 **	--	--
PBK	1.7 (1.2 - 2.4)	2.50e-03 **	--	--
PLK4	1.6 (1 - 2.5)	3.90e-02 *	0.38 (0.19 - 0.80)	0.010188 *
RAD54L	2.3 (1.5 - 3.4)	4.40e-05 ****	--	--
TOP2A	1.5 (1.2 - 1.9)	1.50e-03 **	--	--
TTK	1.8 (1.3 - 2.4)	5.50e-04 ***	--	--
UBE2C	2.3 (1.5-3.6)	2.50e-04 ***	--	--
ZWINT	3 (1.4-6)	3.00e-03 **	--	--

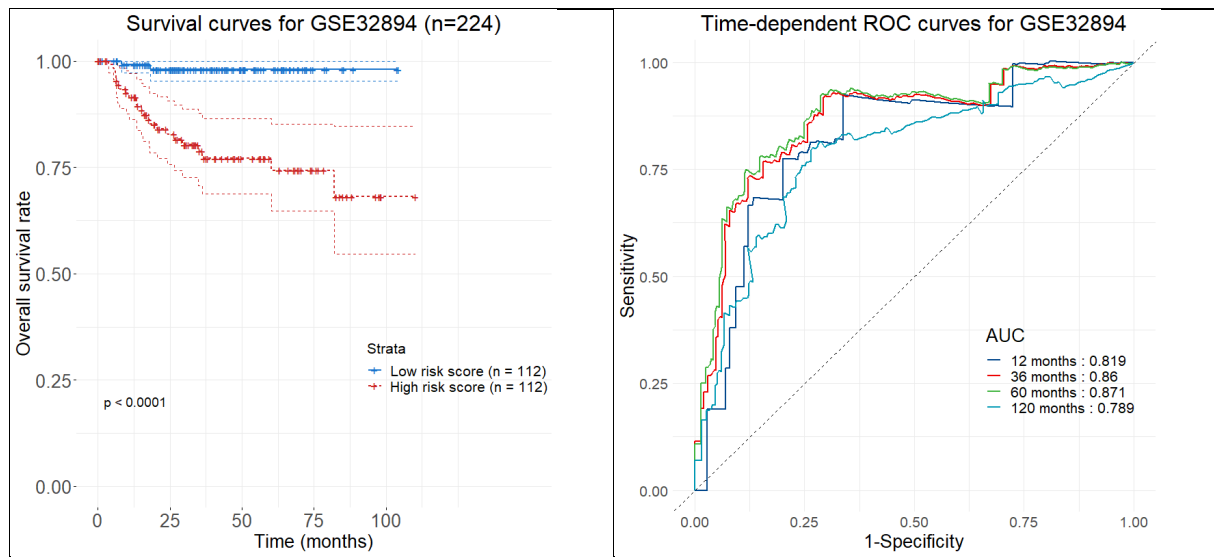




Εικόνα 49: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE13507 (n = 165) (σύνολο εκπαίδευσης)
(Α) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με ΚΟΚ, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με τον διάμεσο προγνωστικό δείκτη.
(Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, το τρίτο, το πέμπτο και το δέκατο έτος μετά τη διάγνωση) με σκοπό την αξιολόγηση της απόδοσης του προγνωστικού μοντέλου των τριών γονιδίων.
(Γ) Διάγραμμα δάσους για την πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox στο σύνολο δεδομένων GSE13507. Η εικόνα ενσωματώνει την τιμή του λόγου κινδύνου (e^{coef}) μαζί με το 95% διάστημα εμπιστοσύνης (CI) και την τιμή-p για κάθε γονίδιο.



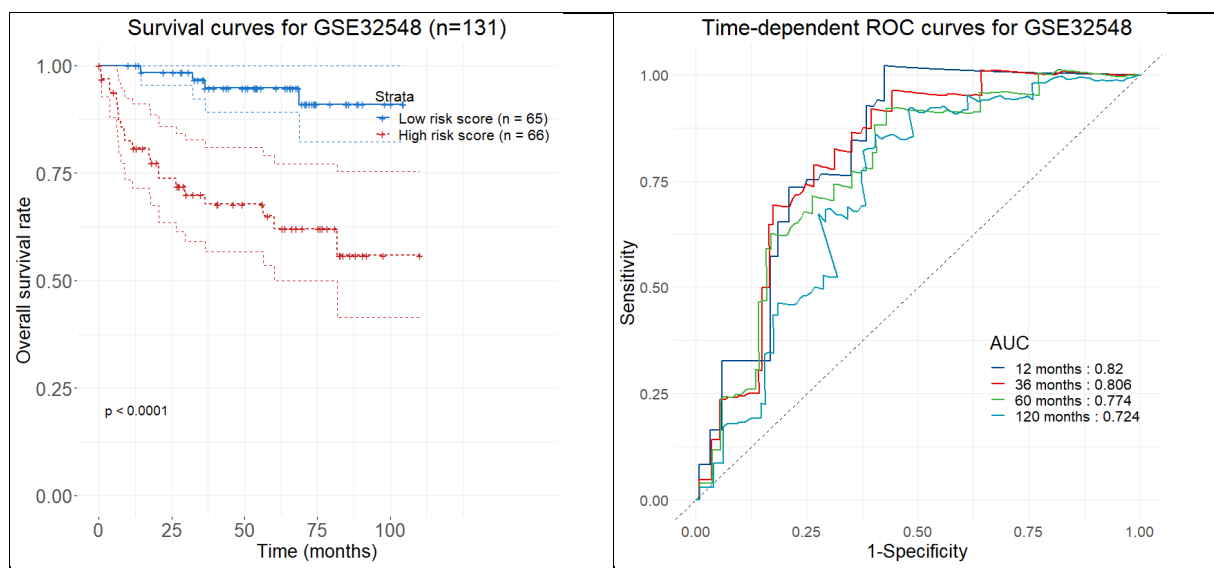
Εικόνα 50: Ανάλυση συσχέτισης της γονιδιακής έκφρασης των τριών γονιδίων του προγνωστικού μοντέλου, που προέκυψαν από την πλατφόρμα GEPIA2. Χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson και η ανάλυση έδειξε ότι δεν υπάρχει στατιστικά σημαντικός συντελεστής συσχέτισης μεταξύ αυτών των γονιδίων για τον ΚΟΚ (μέγιστη τιμή 0,6).



Εικόνα 51: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE32894 (n = 224) (πρώτο σύνολο δοκιμής)

(A) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με KOK, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προγνωστικό δείκτη.

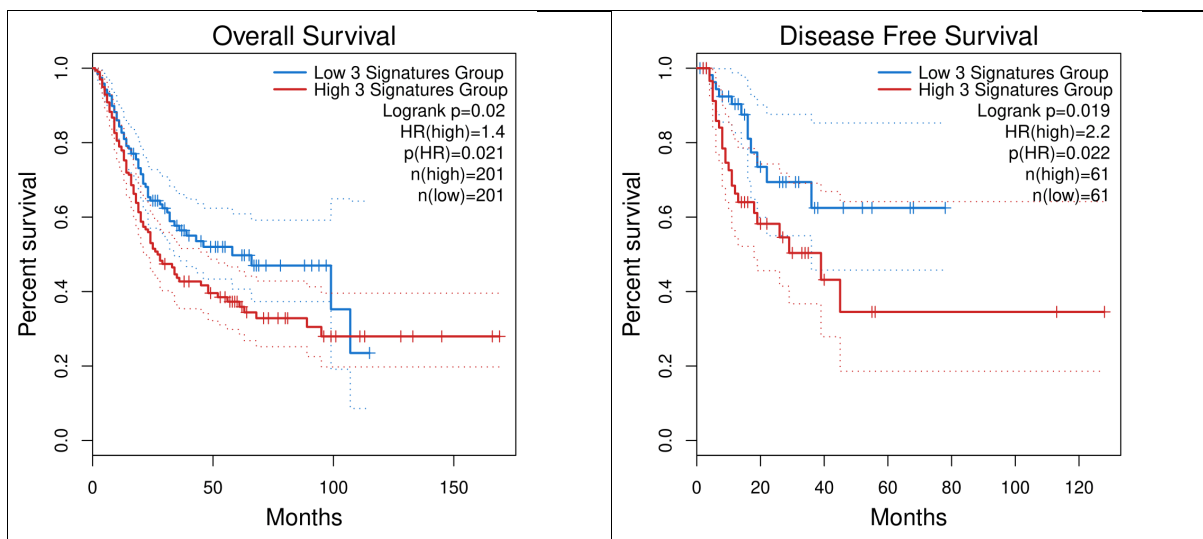
(B) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προγνωστικής απόδοσης του μοντέλου των τριών γονιδίων.



Εικόνα 52: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE32548 (n = 131) (δεύτερο σύνολο δοκιμής)

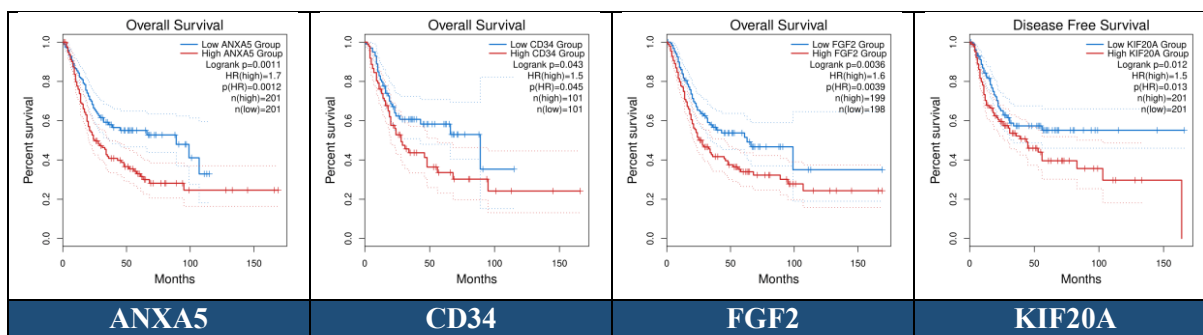
(A) Καμπύλες Kaplan – Meier για τη συνολική επιβίωση των ασθενών με KOK, όπως διαστρωματώνονται με βάση τον προγνωστικό δείκτη των τριών γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προγνωστικό δείκτη.

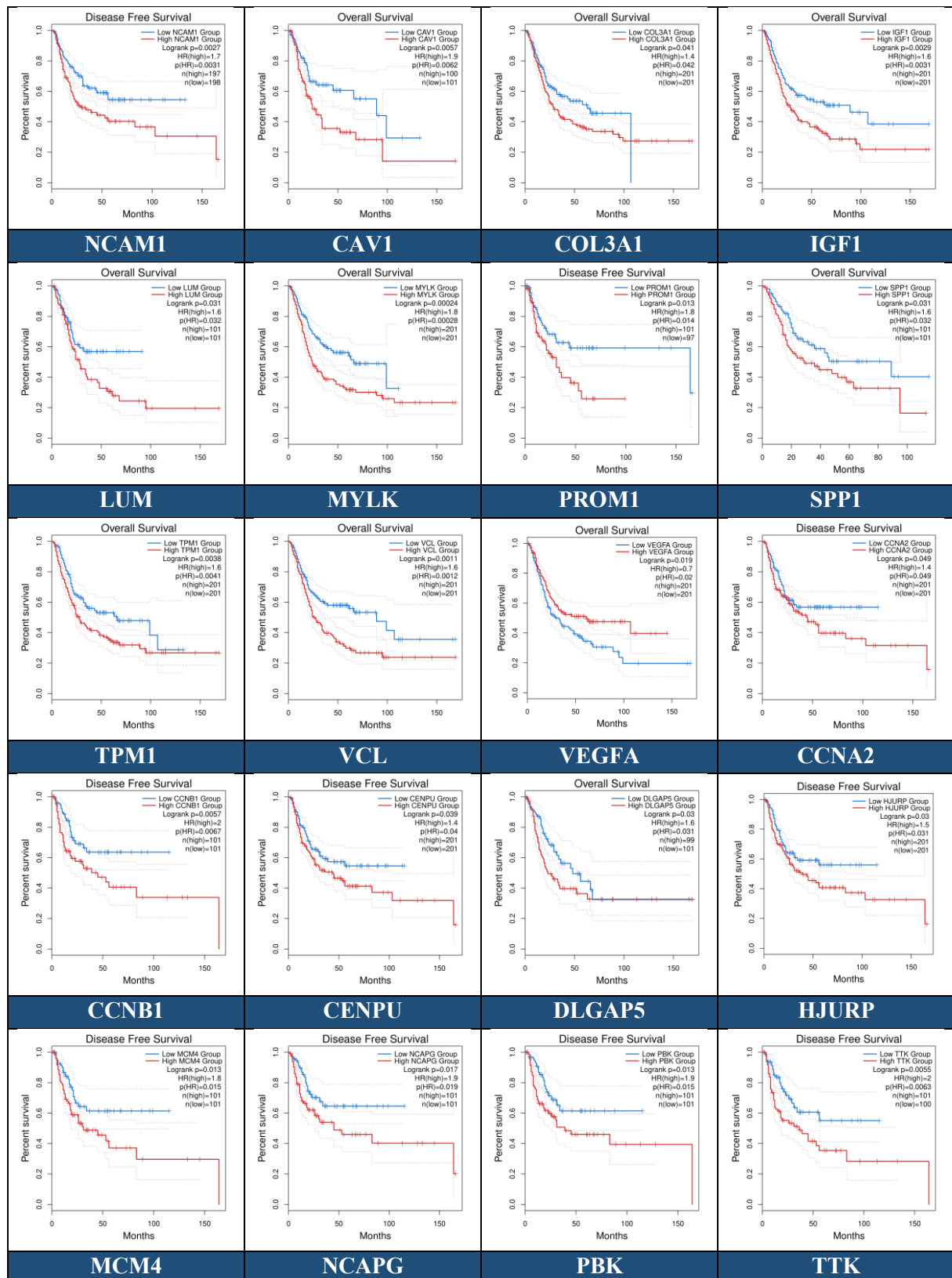
(B) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προγνωστικής απόδοσης του μοντέλου των τριών γονιδίων.

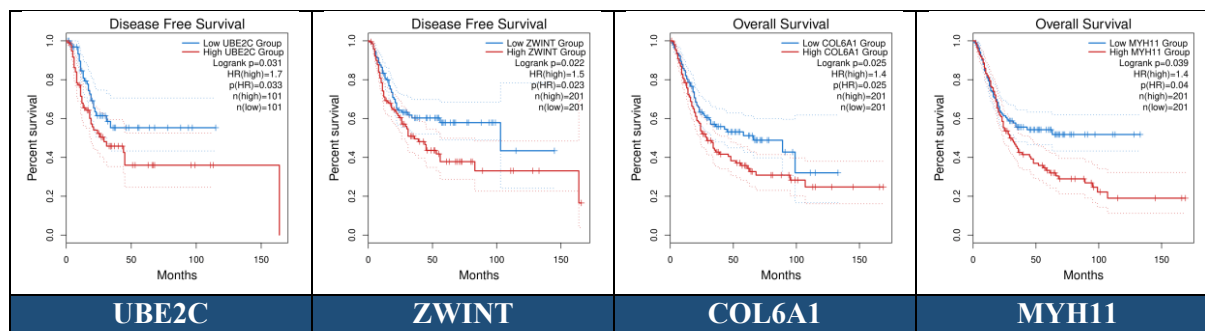


Εικόνα 53: Διαγράμματα επιβίωσης Kaplan – Meier για την προγνωστική υπογραφή των τριών γονιδίων, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών υψηλού και χαμηλού κινδύνου, αντίστοιχα. Οι ασθενείς ομαδοποιήθηκαν σύμφωνα με τη διάμεση τιμή έκφρασης των τριών γονιδίων, όσον αφορά τη συνολική επιβίωση, και σύμφωνα με μια προσαρμοσμένη υψηλή και χαμηλή τιμή αποκοπής 85% και 15%, αντίστοιχα, όσον αφορά την επιβίωση ελεύθερης νόσου.

Τα διαγράμματα επιβίωσης για κάθε ένα από τα μεμονωμένα γονίδια – κλειδιά δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2 και χρησιμοποιήθηκαν για την αξιολόγηση της προγνωστικής αξίας των 61 γονιδίων – κλειδίων ως προς τη συνολική επιβίωση, καθώς και την ελεύθερη νόσου επιβίωση στον KOK (Εικόνα 54). Στα διαγράμματα της συνολικής και της ελεύθερης νόσου επιβίωσης συγκρίθηκαν οι ομάδες υψηλής και χαμηλής έκφρασης στους ιστούς με KOK και η τιμή- $p < 0,05$ θεωρήθηκε στατιστικά σημαντική. Τα αυξημένα επίπεδα έκφρασης των γονιδίων **ANXA5**, **CD34**, **FGF2**, **CAV1**, **COL3A1**, **IGF1**, **LUM**, **MYLK**, **SPPI**, **TPM1**, **VCL**, **DLGAP5**, **COL6A1** και **MYH11** βρέθηκαν να συσχετίζονται με χειρότερη συνολική επιβίωση των ασθενών, ενώ τα επίπεδα έκφρασης του **VEGFA** βρέθηκαν να συσχετίζονται αντιστρόφως ανάλογα με τη συνολική επιβίωση (Εικόνα 54). Επιπλέον, τα υψηλά επίπεδα έκφρασης των γονιδίων **KIF20A**, **NCAM1**, **PROM1**, **CCNA2**, **CCNB1**, **CENPU**, **HJURP**, **MCM4**, **NCAPG**, **PBK**, **TTK**, **UBE2C** και **ZWINT** βρέθηκαν να συσχετίζονται με χειρότερη ελεύθερης νόσου επιβίωση. Για τα υπόλοιπα γονίδια – κλειδιά δεν παρατηρήθηκαν στατιστικά σημαντικές διαφορές ως προς την επιβίωση των ασθενών (τα δεδομένα δεν παρουσιάζονται).





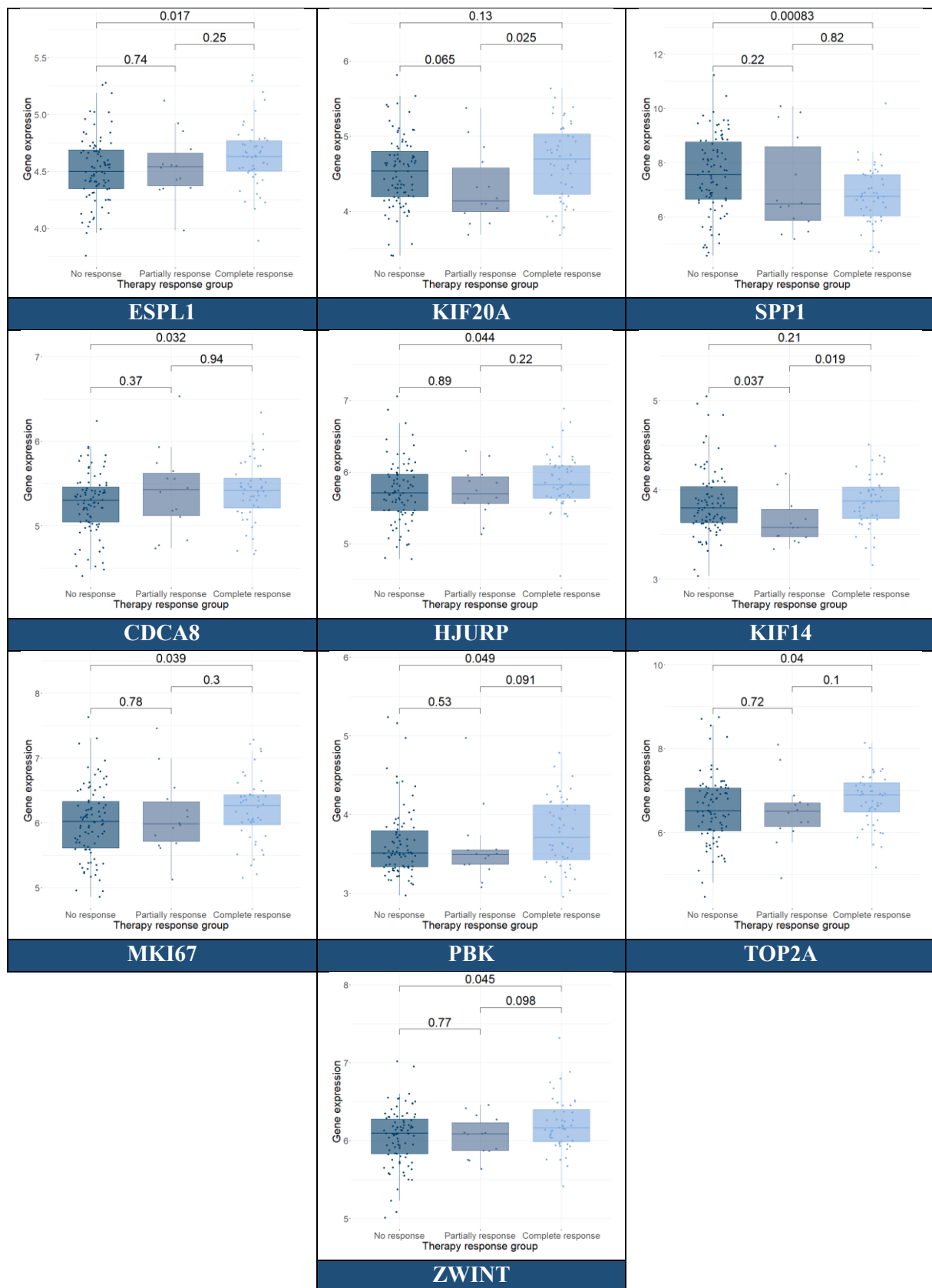


Εικόνα 54: Διαγράμματα επιβίωσης Kaplan – Meier για τα βασικά γονίδια – κόμβους, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών υψηλού και χαμηλού κινδύνου, αντίστοιχα. Οι ασθενείς ομαδοποιήθηκαν σύμφωνα με τις διάμεσες τιμές αποκοπής ή τις τιμές τεταρτημόριου.

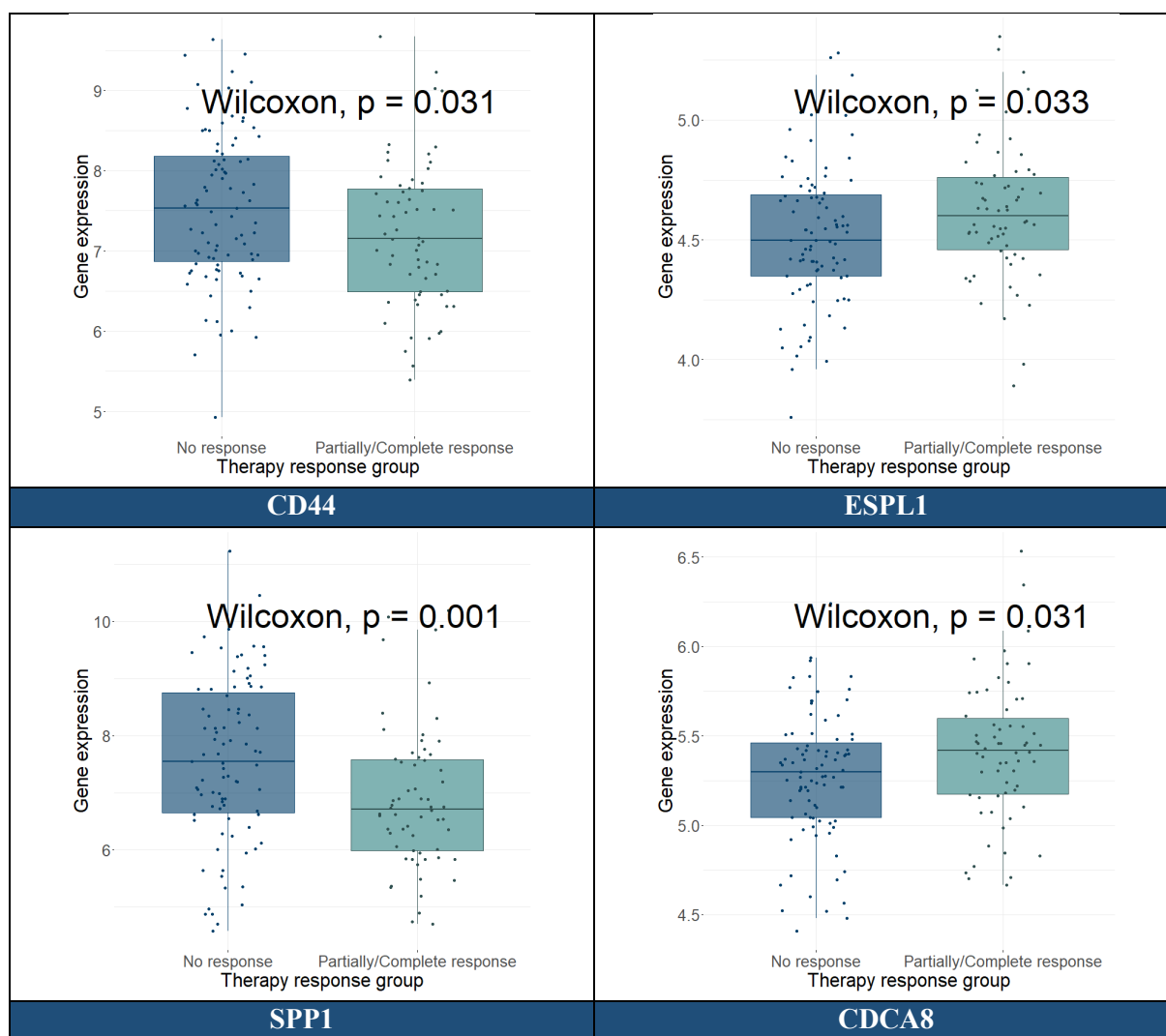
3.11 Προβλεπτικά Γονίδια για τον Καρκίνο της Ουροδόχου Κύστης

Τα πρωτογενή δεδομένα της γονιδιακής έκφρασης των δειγμάτων μυοδιηθητικού ΚΟΚ, από τη σειρά GSE169455, που χαρακτηρίστηκαν από την πλατφόρμα Affymetrix Human Gene 2.1 ST Array (GPL17692), ελήφθησαν και υποβλήθηκαν σε κατάλληλη προεπεξεργασία. Το σύνολο δεδομένων αποτελείται από 149 ασθενείς με μυοδιηθητικό ΚΟΚ που έλαβαν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη. Τα δεδομένα διορθώθηκαν ώστε να εξομαλυνθούν τα φαινόμενα παρτίδας. Εν συνεχεία, εφαρμόστηκε ο έλεγχος αθροίσματος διατάξεων του Wilcoxon για την αξιολόγηση της στατιστικής σημαντικότητας κάθε γονιδίου – κλειδιού μεταξύ των ομάδων «Καμία ανταπόκριση», «Μερική ανταπόκριση» και «Πλήρης ανταπόκριση». Επιπλέον, διερευνήθηκε η στατιστική σημαντικότητα μεταξύ των ομάδων «Καμία ανταπόκριση» και «Μερική ή πλήρης ανταπόκριση» για όλα τα γονίδια – κλειδιά.

Τα γονίδια – κλειδιά με τιμή- $p < 0,05$ προσδιορίστηκαν ως σημαντικά διαφοροποιημένα μεταξύ των ομάδων. Αυτά τα γονίδια αποτελούνταν από τα *ESPL1*, *SPPI*, *CDCA8*, *HJURP*, *MKI67*, *PBK*, *TOP2A* και *ZWINT* μεταξύ των ασθενών που δεν ανταποκρίθηκαν στη θεραπεία και εκείνων που ανταποκρίθηκαν πλήρως, το γονίδιο *KIF14* μεταξύ των ασθενών που δεν ανταποκρίθηκαν στη θεραπεία και εκείνων που ανταποκρίθηκαν μερικώς, και τα γονίδια *KIF20A* και *KIF14* μεταξύ των ασθενών που ανταποκρίθηκαν μερικώς και πλήρως. Τα αποτελέσματα του ελέγχου αθροίσματος διατάξεων του Wilcoxon για αυτά τα γονίδια παρουσιάζονται στην Εικόνα 55. Όσον αφορά τη σύγκριση των δύο κατηγοριών, μεταξύ των ομάδων «Καμία ανταπόκριση» και «Μερική ή πλήρης ανταπόκριση», τα γονίδια *CD44*, *ESPL1*, *SPPI* και *CDCA8* εκφράστηκαν σημαντικά διαφορετικά. Τα αποτελέσματα του ελέγχου αθροίσματος διατάξεων του Wilcoxon για τα γονίδια αυτά παρουσιάζονται στην Εικόνα 56.



Εικόνα 55: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά μεταξύ των ομάδων «Καμία ανταπόκριση», «Μερική ανταπόκριση» και «Πλήρης ανταπόκριση» όσον αφορά την προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη των ασθενών με μυοδιηθητικό ΚΟΚ.



Εικόνα 56: Σημαντικά διαφορικά εκφραζόμενα γονίδια – κλειδιά μεταξύ των ομάδων «Καμία ανταπόκριση» και «Μερική ή πλήρης ανταπόκριση» όσον αφορά την προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη των ασθενών με μυοδιηθητικό ΚΟΚ.

Για τη διερεύνηση της προβλεπτικής αξίας των γονιδίων – κλειδιών όσον αφορά την ανταπόκριση των ασθενών στη χημειοθεραπεία με βάση τη σισπλατίνη, χρησιμοποιήθηκε η μονο-μεταβλητή ανάλυση παλινδρόμησης Cox για καθένα από τα γονίδια αυτά σε συνολικά 149 ασθενείς με μυοδιηθητικό ΚΟΚ (GSE169455). Τα κλινικά χαρακτηριστικά των ασθενών για αυτό το σύνολο δεδομένων περιέχουν πληροφορίες σχετικά με την ελεύθερη υποτροπής επιβίωση (RFS), την ειδική της νόσου επιβίωση (CSS) και τη συνολική επιβίωση (OS). Ως εκ τούτου, πραγματοποιήθηκε μονο-μεταβλητή ανάλυση παλινδρόμησης Cox για κάθε ένα από αυτά τα συμβάντα και τα γονίδια – κλειδιά. Μόνο δύο γονίδια, και συγκεκριμένα το *SPPI* και το *CDCA8*, βρέθηκαν να είναι στατιστικά σημαντικά συγκριτικά με την ελεύθερη υποτροπής επιβίωση, την ειδική της νόσου επιβίωση και τη συνολική επιβίωση ταυτόχρονα, με το γονίδιο *SPPI* να παρουσιάζει ισχυρή στατιστική σημαντικότητα (Πίνακας 7). Επιπλέον, πραγματοποιήθηκε ανάλυση παλινδρόμησης LASSO Cox με 10-πλη διασταυρούμενη επικύρωση με σκοπό να επιλεγθούν τα γονίδια – κλειδιά τα οποία είχαν υψηλή συσχέτιση με την ελεύθερη υποτροπής επιβίωση στον ΚΟΚ. Εντοπίστηκαν συνολικά 19 γονίδια, συγκεκριμένα τα *ACTG2*, *ANXA5*, *AURKB*, *CCNA2*, *CCNB2*, *CD44*, *CDC45*, *CDCA8*, *CDT1*, *CENPA*, *COL6A1*, *DLGAP5*, *IGF1*, *KIF14*, *NCAM1*, *NEK2*, *SPPI*, *VCL* και *ZWINT*. Η πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox σε αυτά τα γονίδια έδειξε ότι η έκφραση των *ANXA5*, *CD44*, *NCAM1*, *SPPI*, *CDCA8* και *KIF14* σχετιζόταν σε μεγάλο

βαθμό και ανεξάρτητα με την ελεύθερη υποτροπής επιβίωση των ασθενών με KOK (Πίνακας 7), και χρησιμοποιήθηκε για τον υπολογισμό του συντελεστή κάθε γονιδίου στο προβλεπτικό μοντέλο. Τέλος, κατασκευάστηκε το προβλεπτικό μοντέλο των έξι γονιδίων και ο τύπος του προβλεπτικού δείκτη (ΠΒΔ) για το μοντέλο αυτό ήταν:

$$ΠΒΔ = |-0.87492 \cdot εκφ_{ANXA5} + 0.50317 \cdot εκφ_{CD44} + 0.46781 \cdot εκφ_{NCAM1} + 0.54406 \cdot εκφ_{SPPI} - 1.70391 \cdot εκφ_{CDCA8} + 1.54315 \cdot εκφ_{KIF14}|$$

Το διάγραμμα δάσους του προβλεπτικού μοντέλου παρουσιάζεται στην Εικόνα 57C. Η ανάλυση συσχέτισης της συν-έκφρασης για καθένα από τα 15 ζεύγη γονιδίων του παραπάνω μοντέλου πραγματοποιήθηκε μέσω της πλατφόρμας GEPIA2 και έδειξε ότι κανένα ζευγάρι αυτών των γονιδίων δεν είχε συντελεστή συσχέτισης Pearson μεγαλύτερο από 0,44 στον KOK (Εικόνα 58).

Στη συνέχεια, υπολογίστηκε ο προβλεπτικός δείκτης για κάθε έναν από τους ασθενείς του συνόλου εκπαίδευσης (GSE169455), και παρατηρήθηκε πως ο χρόνος της ελεύθερης υποτροπής επιβίωσης για τους ασθενείς υψηλού κινδύνου ($n = 74$) ήταν σημαντικά χειρότερος (τιμή- $p < 0,0001$) από εκείνον των ασθενών χαμηλού κινδύνου ($n = 74$), όπως φαίνεται από την ανάλυση των καμπύλων Kaplan – Meier (Εικόνα 57A). Επιπλέον, αξιολογήθηκε η προβλεπτική απόδοση της υπογραφής των έξι γονιδίων με τη χρήση της χρονοεξαρτώμενης ανάλυσης ROC σε συγκεκριμένους χρόνους μετά τη διάγνωση, και συγκεκριμένα στο πρώτο, τρίτο, πέμπτο και δέκατο έτος. Η τιμή AUC στους διάφορους χρόνους αποκοπής ήταν 0,6, 0,69, 0,707 και 0,6, αντίστοιχα (Εικόνα 57B).

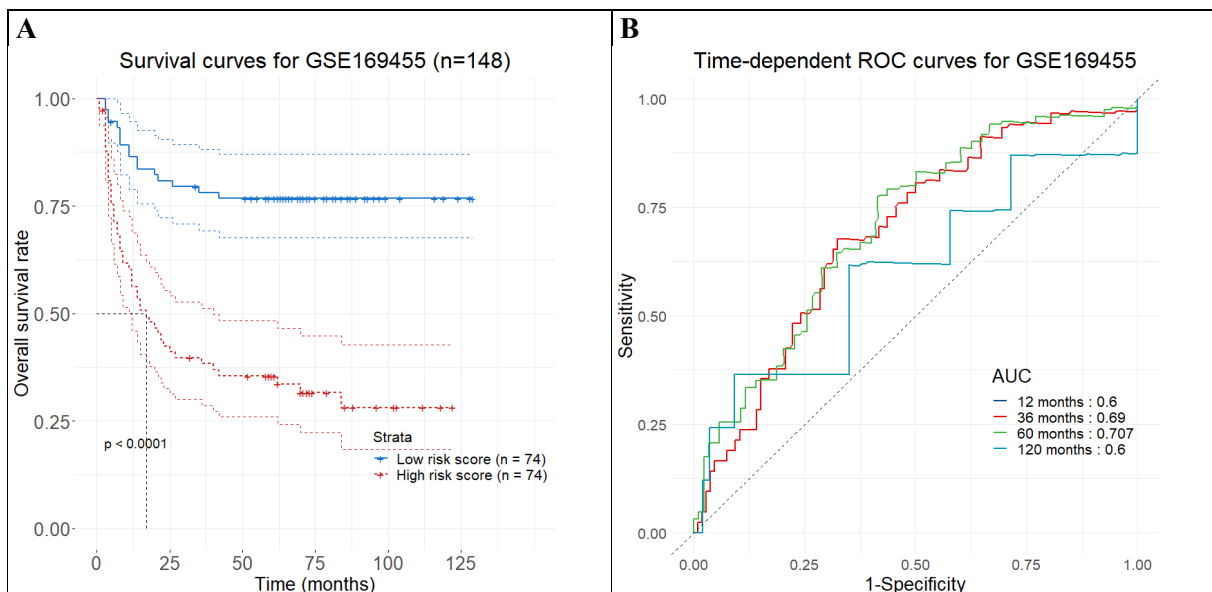
Το προβλεπτικό μοντέλο αξιολογήθηκε σε δύο ανεξάρτητα σύνολα δεδομένων. Στο πρώτο σύνολο δεδομένων (GSE87304), τα δεδομένα επιβίωσης ήταν διαθέσιμα για τους 258 από τους 305 ασθενείς. Έτσι, η ομάδα ασθενών χαμηλού κινδύνου ($n = 129$), όπως υπολογίστηκε από το προβλεπτικό μοντέλο, επέδειξε στατιστικά σημαντικά μεγαλύτερη ελεύθερης υποτροπής επιβίωση (DFS) (τιμή- $p < 0,012$) σε σύγκριση με την ομάδα ασθενών υψηλού κινδύνου ($n = 129$) (Εικόνα 59A). Επίσης, σχεδιάστηκαν οι χρονοεξαρτώμενες καμπύλες ROC και οι τιμές AUC για το τρίτο, το πέμπτο και το έκτο έτος ήταν 0,699, 0,572 και 0,582, αντίστοιχα (Εικόνα 59B). Στο δεύτερο σύνολο δεδομένων (GSE69795), η ομάδα ασθενών χαμηλού κινδύνου ($n = 19$), όπως υπολογίστηκε από το προβλεπτικό μοντέλο, επέδειξε στατιστικά σημαντικά μεγαλύτερη ελεύθερης υποτροπής επιβίωση (DFS) (τιμή- $p < 0,038$) σε σύγκριση με την ομάδα ασθενών υψηλού κινδύνου ($n = 19$) (Εικόνα 60A). Ακόμη, σχεδιάστηκαν οι χρονοεξαρτώμενες καμπύλες ROC και οι τιμές AUC για το πρώτο, τρίτο, πέμπτο και έβδομο έτος ήταν 0,619, 0,673, 0,611 και 0,611, αντίστοιχα (Εικόνα 60B).

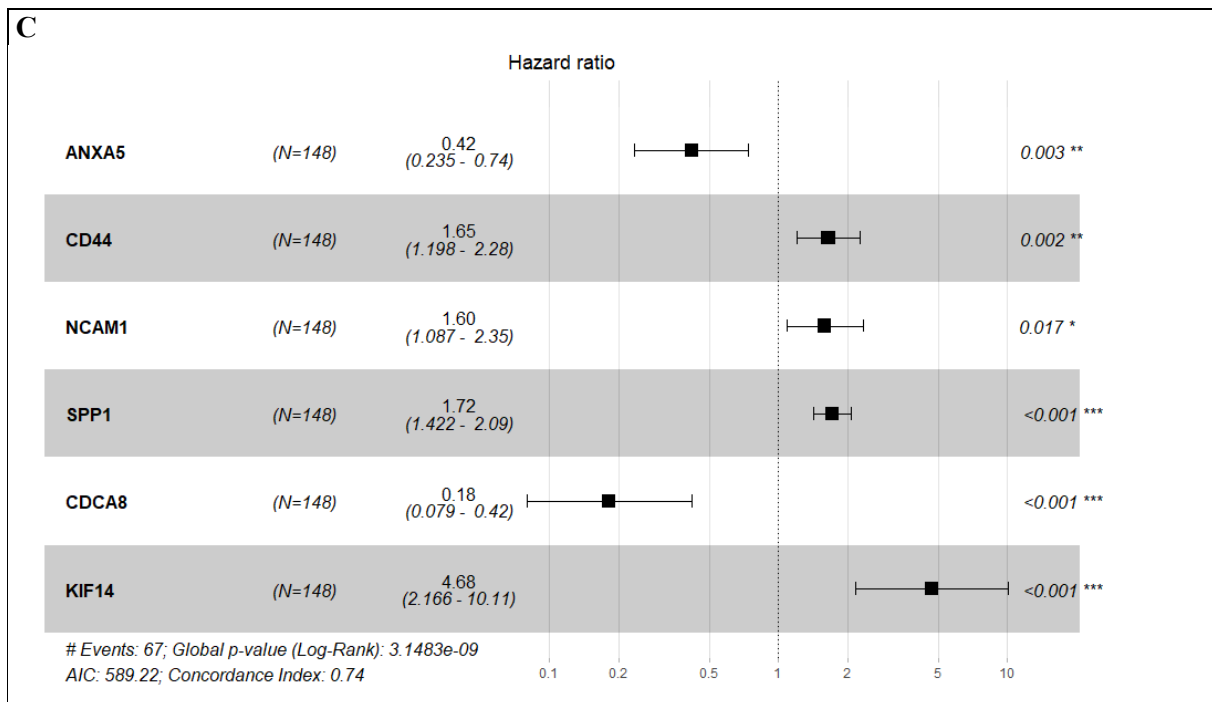
Στη συνέχεια, διερευνήθηκε κατά πόσον η υπογραφή των έξι γονιδίων που αποτελείται από τα *ANXA5*, *CD44*, *NCAM1*, *SPPI*, *CDCA8* και *KIF14* θα μπορούσε να είναι προγνωστική για τη συνολική επιβίωση ή την ελεύθερης υποτροπής επιβίωση των ασθενών με KOK της κοόρτης TCGA. Έτσι, ελέγχθηκε αυτή η υπογραφή με την ανάλυση επιβίωσης Kaplan – Meier, χρησιμοποιώντας την πλατφόρμα GEPIA2. Οι ασθενείς διαχωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου ($n = 121$ σε κάθε ομάδα) με βάση προσαρμοσμένες τιμές αποκοπής κάτω και άνω ορίου, 30% και 70%, αντίστοιχα. Οι δύο ομάδες εμφάνισαν στατιστικά σημαντικά διαφορετικό χρόνο συνολικής επιβίωσης (τιμή- $p = 0,0011$) αλλά και χρόνο ελεύθερης υποτροπής επιβίωσης (τιμή- $p = 0,0014$) (Εικόνα 61). Αυτό έδειξε ότι η υπογραφή των έξι γονιδίων είχε επίσης προγνωστική αξία όσον αφορά τη συνολική επιβίωση αλλά και την ελεύθερης υποτροπής επιβίωση των ασθενών με KOK.

Συμπερασματικά, η ανάλυση επιβίωσης Kaplan – Meier και οι τιμές AUC για τους διάφορους χρόνους αποκοπής έδειξαν ότι το προβλεπτικό μοντέλο των έξι γονιδίων παρουσιάζει αρκετά καλή προβλεπτική ικανότητα όσον αφορά τον χρόνο ελεύθερης υποτροπής επιβίωσης των ασθενών με μυοδιηθητικό KOK που έλαβαν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη και, συνεπώς, θα μπορούσε να αξιολογηθεί περαιτέρω κλινικά για το κατά πόσον μπορεί να προβλέψει την ανταπόκριση των ασθενών με μυοδιηθητικό KOK στην προεγχειρητική χημειοθεραπεία.

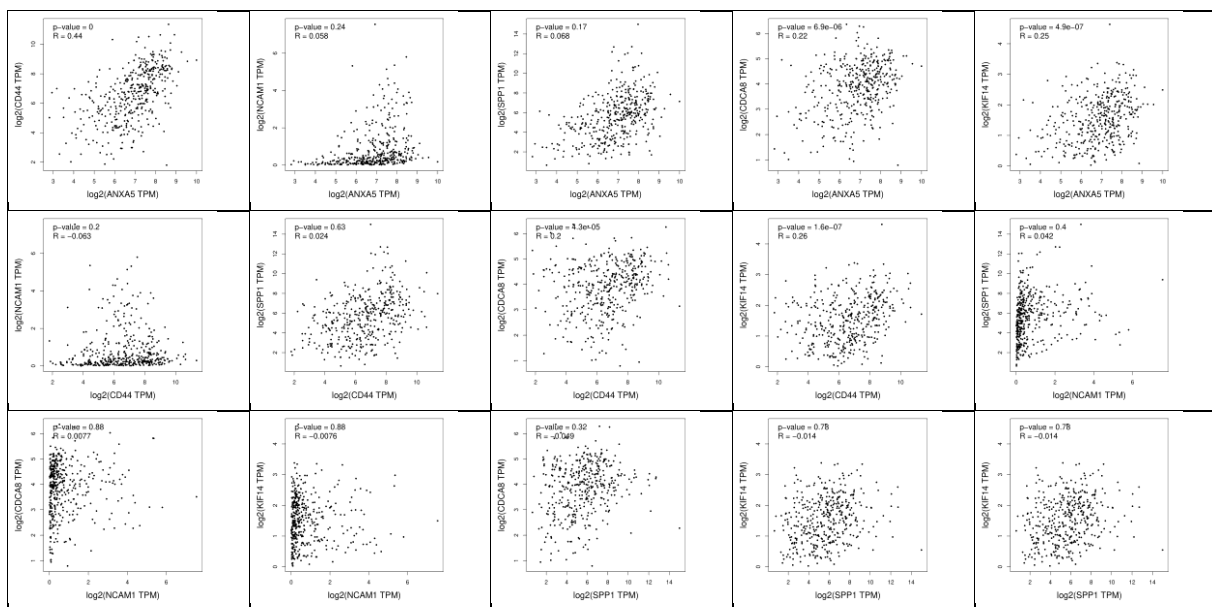
Πίνακας 7: Μονο-μεταβλητή και πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox των γονιδίων – κλειδιών που σχετίζονται με την επιβίωση των ασθενών με μυοδιθητικό ΚΟΚ που λάμβαναν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη
HR: λόγος κινδύνου (Hazard ratio), CI: Διάστημα εμπιστοσύνης (Confidence interval)
*τιμή-p < 0.05, **τιμή-p < 0.01, ***τιμή-p < 0.001, ****τιμή-p < 0.0001

Γονίδια σχετικά με την ελεύθερη υποτροπής επιβίωση (RFS)	Μονο-μεταβλητή ανάλυση		Πολυ-μεταβλητή ανάλυση	
	HR (95% CI)	Τιμή-p	HR (95% CI)	Τιμή-p
ANXA5	1.1 (0.70 - 1.70)	0.7	0.42 (0.24 - 0.74)	0.00268 **
CD44	1.3 (0.99 - 1.70)	0.055	1.65 (1.20 - 2.28)	0.00220 **
NCAM1	1.1 (0.81 - 1.60)	0.48	1.60 (1.09 - 2.35)	0.01718 *
IGF1	0.48 (0.24 - 0.98)	0.043 *	--	--
SPP1	1.5 (1.30 - 1.80)	3.3e-06 ****	1.72 (1.42 - 2.09)	2.93e-08 ****
CDC48	0.5 (0.26 - 0.96)	0.037 *	0.18 (0.08 - 0.42)	5.72e-05 ****
KIF14	1.8 (0.87 - 3.70)	0.11	4.68 (2.17 - 10.11)	8.59e-05 ****
Γονίδια σχετικά με την ειδική της νόσου επιβίωση (CSS)				
ANXA5	1.1 (0.68 - 1.70)	0.74	0.44 (0.24 - 0.82)	0.009708 **
CD44	1.3 (0.95 - 1.60)	0.11	1.60 (1.12 - 2.29)	0.009785 **
NCAM1	1.1 (0.82 - 1.60)	0.47	1.43 (0.99 - 2.05)	0.049989 *
SPP1	1.4 (1.20 - 1.70)	5.8e-05 ****	1.64 (1.35 - 2.01)	1.15e-06 ****
CDC48	0.48 (0.24 - 0.94)	0.034 *	0.17 (0.07 - 0.41)	6.49e-05 ****
KIF14	1.7 (0.84 - 3.60)	0.14	4.82 (2.12 - 10.96)	0.000171 ***
Γονίδια σχετικά με τη συνολική επιβίωση (OS)				
ACTG2	1.3 (1.00 - 1.60)	0.038 *	--	--
ANXA5	0.97 (0.63 - 1.50)	0.9	0.41 (0.23 - 0.72)	0.002139 **
CD44	1.2 (0.96 - 1.60)	0.095	1.63 (1.17 - 2.28)	0.003812 **
NCAM1	1.1 (0.82 - 1.50)	0.48	1.42 (1.00 - 2.01)	0.047272 *
SPP1	1.4 (1.20 - 1.60)	0.00012 ***	1.60 (1.32 - 1.92)	1.1e-06 ****
CDC48	0.48 (0.25 - 0.91)	0.024 *	0.19 (0.09 - 0.42)	5.0e-05 ****
KIF14	1.6 (0.78 - 3.20)	0.21	4.45 (2.01 - 9.85)	0.000225 ***

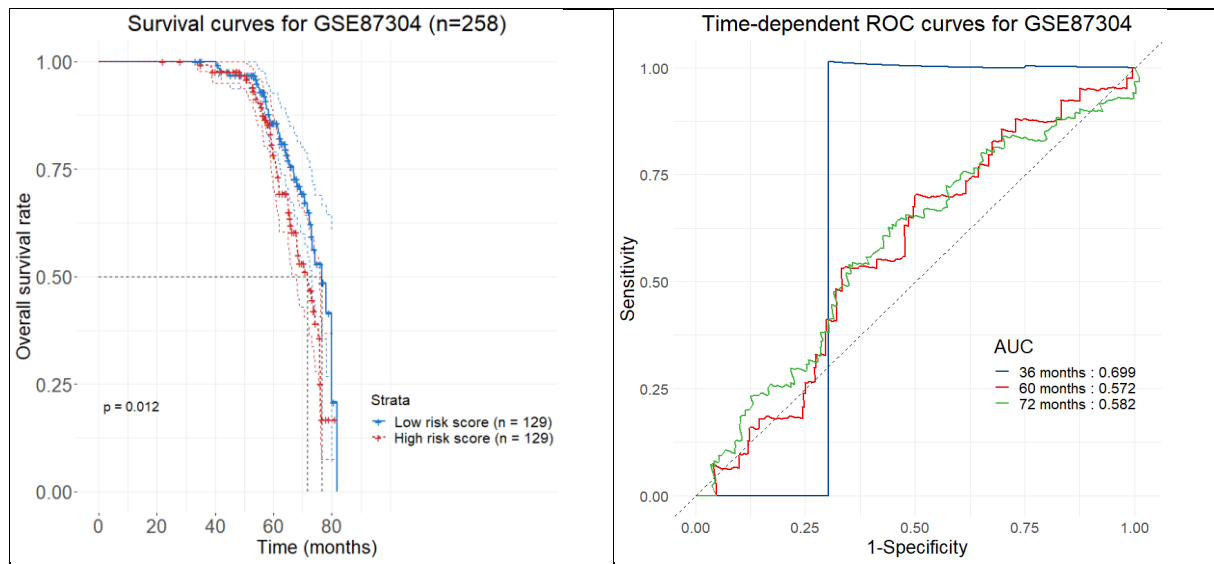




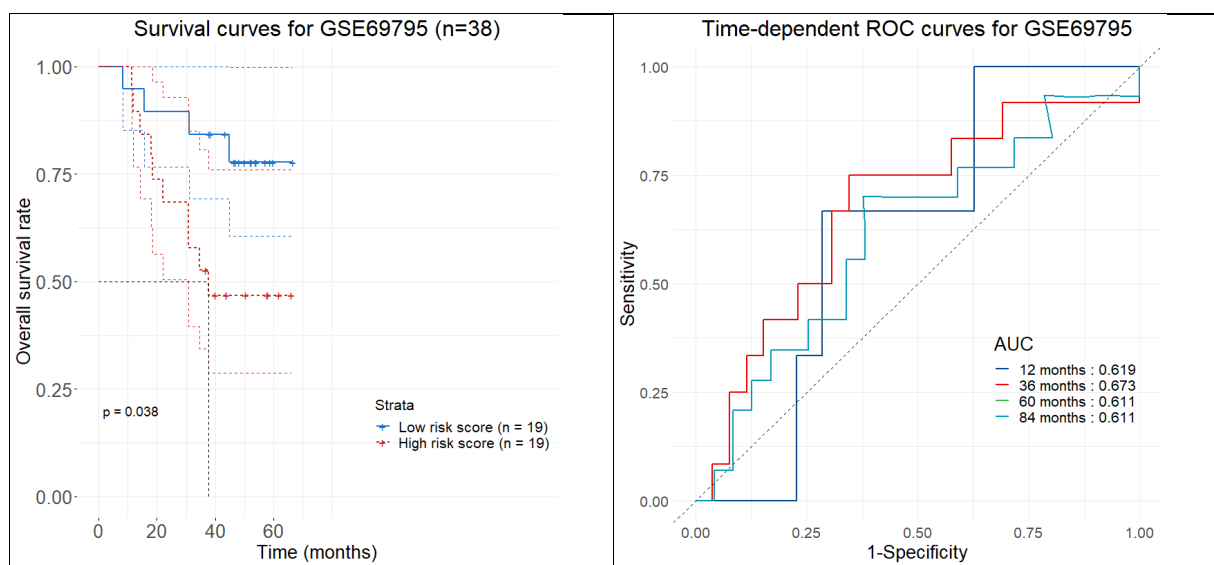
Εικόνα 57: Ανάλυση επιβίωσης για το σύνολο δεδομένων GSE169455 (n = 148) (σύνολο εκπαίδευσης)
(Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη.
(Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και δέκατο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων.
(Γ) Διάγραμμα δάσους για την πολυμεταβλητή ανάλυση παλινδρόμησης Cox στο σύνολο δεδομένων GSE169455. Η εικόνα ενσωματώνει την τιμή του λόγου κινδύνου (e^{coef}) μαζί με το 95% διάστημα εμπιστοσύνης (CI) και την τιμή-p για κάθε γονίδιο.



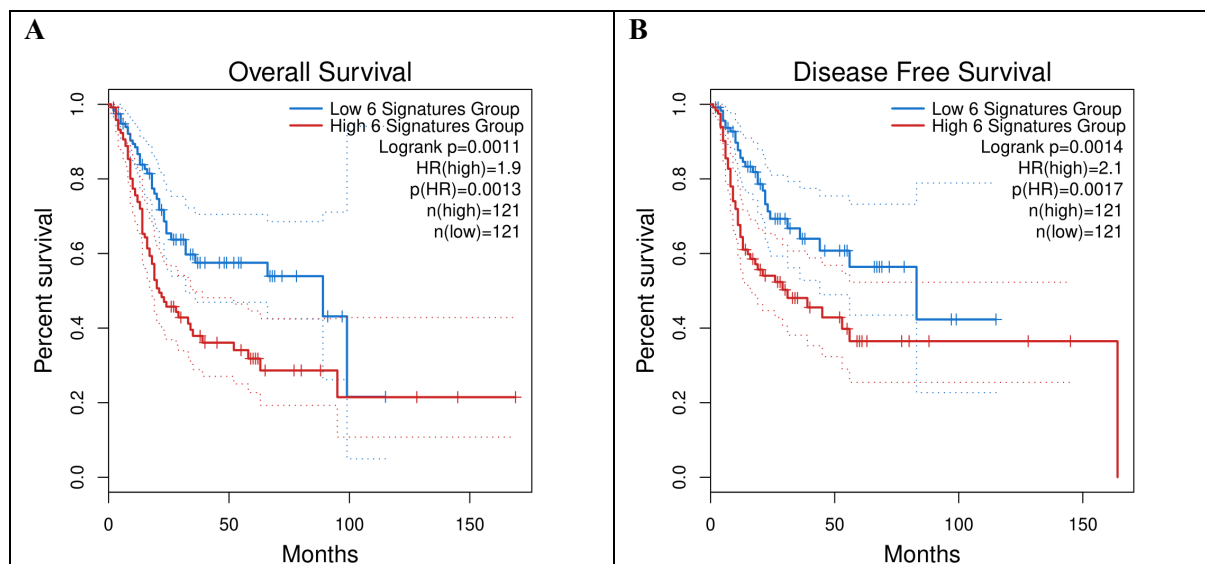
Εικόνα 58: Ανάλυση συσχέτισης της γονιδιακής έκφρασης των έξι γονιδίων του προβλεπτικού μοντέλου, που προέκυψαν από την πλατφόρμα GEPIA2. Χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson και έδειξε ότι δεν υπάρχει στατιστικά σημαντικός συντελεστής συσχέτισης μεταξύ αυτών των γονιδίων για τον ΚΟΚ (μέγιστη τιμή 0,44).



Εικόνα 59: Ανάλυση επιβίωσης με χρήση του συνόλου δεδομένων GSE87304 (n = 258) (πρώτο σύνολο δοκιμής)
(Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη.
(Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το τρίτο, πέμπτο και έκτο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων.



Εικόνα 60: Ανάλυση επιβίωσης με χρήση του συνόλου δεδομένων GSE69795 (n = 19) (δεύτερο σύνολο δοκιμής)
(Α) Καμπύλες Kaplan – Meier για την επιβίωση ελεύθερης νόσου των ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη, όπως διαστρωματώθηκαν με βάση τον προβλεπτικό δείκτη των έξι γονιδίων. Οι ασθενείς χωρίστηκαν σε ομάδες χαμηλού και υψηλού κινδύνου σύμφωνα με το διάμεσο προβλεπτικό δείκτη.
(Β) Χρονοεξαρτώμενες καμπύλες ROC (προβλέψεις για το πρώτο, τρίτο, πέμπτο και έβδομο έτος μετά τη διάγνωση) για την αξιολόγηση της προβλεπτικής ακρίβειας του μοντέλου των έξι γονιδίων.

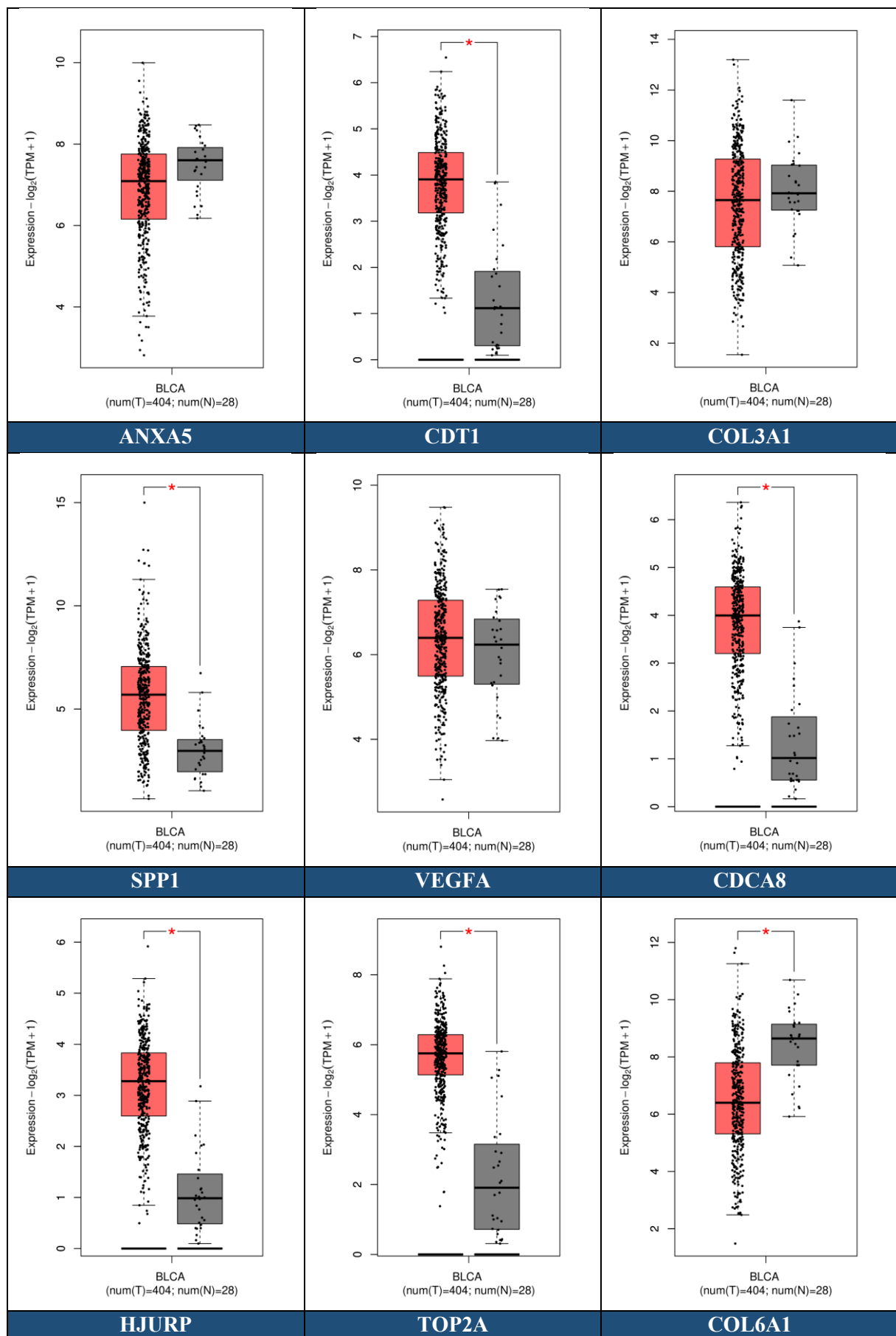


Εικόνα 61: Διαγράμματα επιβίωσης Kaplan – Meier της προβλεπτικής υπογραφής των έξι γονιδίων, που δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι κόκκινες και μπλε γραμμές υποδεικνύουν τις ομάδες ασθενών με υψηλή και χαμηλή έκφραση για τα έξι αυτά γονίδια, αντίστοιχα. Οι ασθενείς ομαδοποιήθηκαν (n = 121 σε κάθε ομάδα) σύμφωνα με τα προσαρμοσμένα χαμηλά και υψηλά κατώφλια αποκοπής, 30% και 70%, αντίστοιχα, όσον αφορά τη σύγκριση της συνολικής τους επιβίωσης και της ελεύθερης νόσου επιβίωσής τους.

3.12 Ταυτοποίηση της Έκφρασης των Βιοδεικτών και Ανοσοϊστοχημία

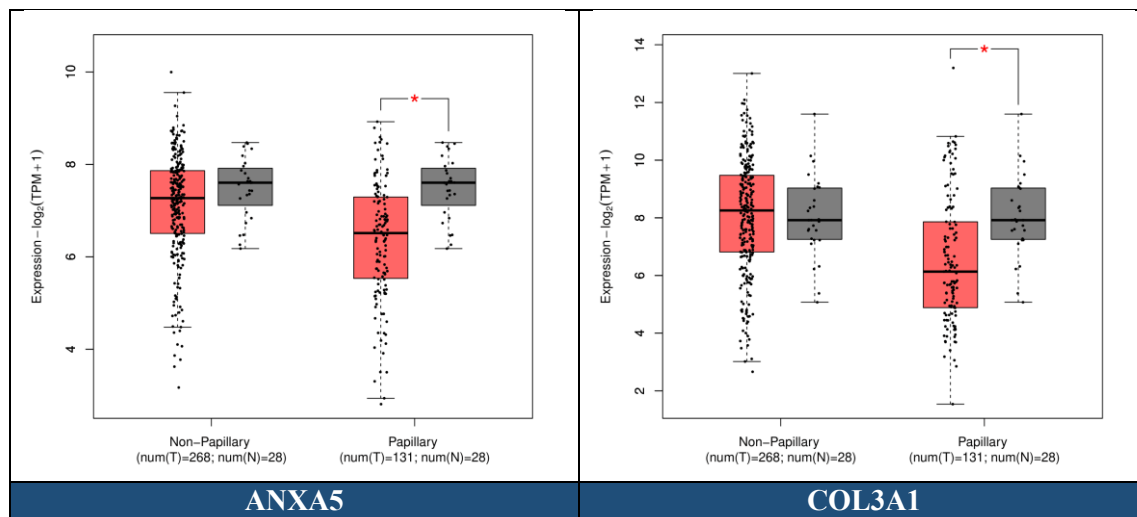
Τα σημαντικά γονίδια – κόμβοι που παρατηρήθηκε ότι εκφράζονται διαφορεικά μεταξύ των καρκινικών και των υγιών δειγμάτων είτε στα ούρα είτε στο πλάσμα αίματος των ασθενών με ΚΟΚ (τιμή- $p < 0,05$) και ταυτόχρονα αποδείχθηκε να έχουν κάποια προγνωστική ή προβλεπτική αξία, θεωρήθηκαν ως οι **πιθανοί σημαντικοί βιοδείκτες** της παρούσας ενοποιητικής μετα-ανάλυσης. Αυτοί οι πιθανοί βιοδείκτες περιλαμβάνουν τα γονίδια *ANXA5*, *CDT1*, *COL3A1*, *SPPI*, *VEGFA*, *CDCA8*, *HJURP*, *TOP2A* και *COL6A1*.

Με σκοπό την επιβεβαίωση των διαφορικών επιπέδων έκφρασης mRNA των προτεινόμενων βιοδεικτών – γονιδίων μεταξύ των δειγμάτων ΚΟΚ και των υγιών δειγμάτων, τα σύνολα δεδομένων TCGA και GTEx αναλύθηκαν με τη χρήση της πλατφόρμας GEPIA2. Οι επιλεγμένες τιμές αποκοπής ορίστηκαν οι εξής: $|\log_2FC| = 1$ και τιμή- $p = 0,05$. Τα αντίστοιχα θερμογράμματα δημιουργήθηκαν και μεταφορτώθηκαν από την πλατφόρμα GEPIA2 (Εικόνα 62). Τα διαγράμματα επικύρωσης έδειξαν ότι τα αποτελέσματα της ανάλυσης διαφορικής έκφρασης της παρούσας έρευνας για όλα τα γονίδια επικυρώθηκαν, όσον αφορά την παρουσία υπο- ή υπερ-έκφρασης. Ωστόσο, για τρία γονίδια, και συγκεκριμένα τα *ANXA5*, *COL3A1* και *VEGFA*, οι διαφορές μεταξύ των μέσων όρων της ομάδας των δειγμάτων ΚΟΚ και της ομάδας των δειγμάτων ελέγχου είχαν χαμηλότερη τιμή $|\log_2FC|$ από την επιλεγμένη και, ως εκ τούτου, δεν χαρακτηρίστηκαν ως στατιστικά σημαντικές. Για να καταστεί δυνατή η λεπτομερέστερη διερεύνηση της έκφρασης αυτών των γονιδίων για τους κύριους υπότυπους του ΚΟΚ (δηλ. τον μη θηλώδες ΚΟΚ και τον θηλώδες ΚΟΚ), αναλύθηκε περαιτέρω η έκφρασή τους και διαπιστώθηκε ότι τα γονίδια *ANXA5* και *COL3A1* εκφράζονταν σημαντικά διαφορετικά στον θηλώδη υπότυπο ΚΟΚ, χρησιμοποιώντας τις προαναφερθείσες τιμές αποκοπής (Εικόνα 63).



Εικόνα 62: Η ανάλυση των επιπέδων της γονιδιακής έκφρασης των εννέα προτεινόμενων βιοδεικτών σε ασθενείς με ΚΟΚ, όπως δημιουργήθηκε με τη χρήση της πλατφόρμας GEPIA2. Τα κόκκινα πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης mRNA σε ιστούς ΚΟΚ και τα γκρι πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης σε υγιείς ιστούς ελέγχου της ουροδόχου κύστης από ασθενείς των κοορτών TCGA και GTEx.

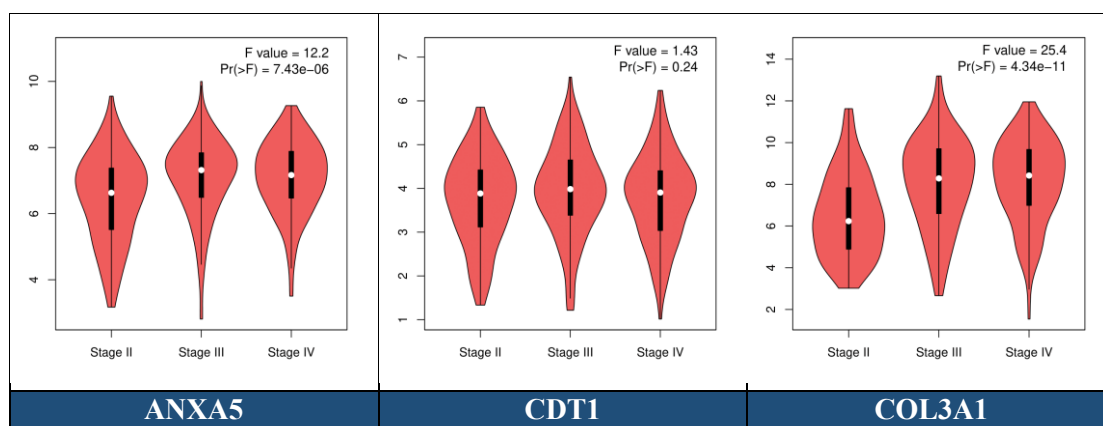
Ως κριτήρια αποκοπής για τη στατιστική σημαντικότητα εφαρμόστηκαν τα: τιμή- $p < 0,05$ και $|\log_2FC| < 1$.
BLCA: καρκίνος της ουροδόχου κύστης, TPM: Αριθμός μεταγραφών ανά εκατομμύριο.

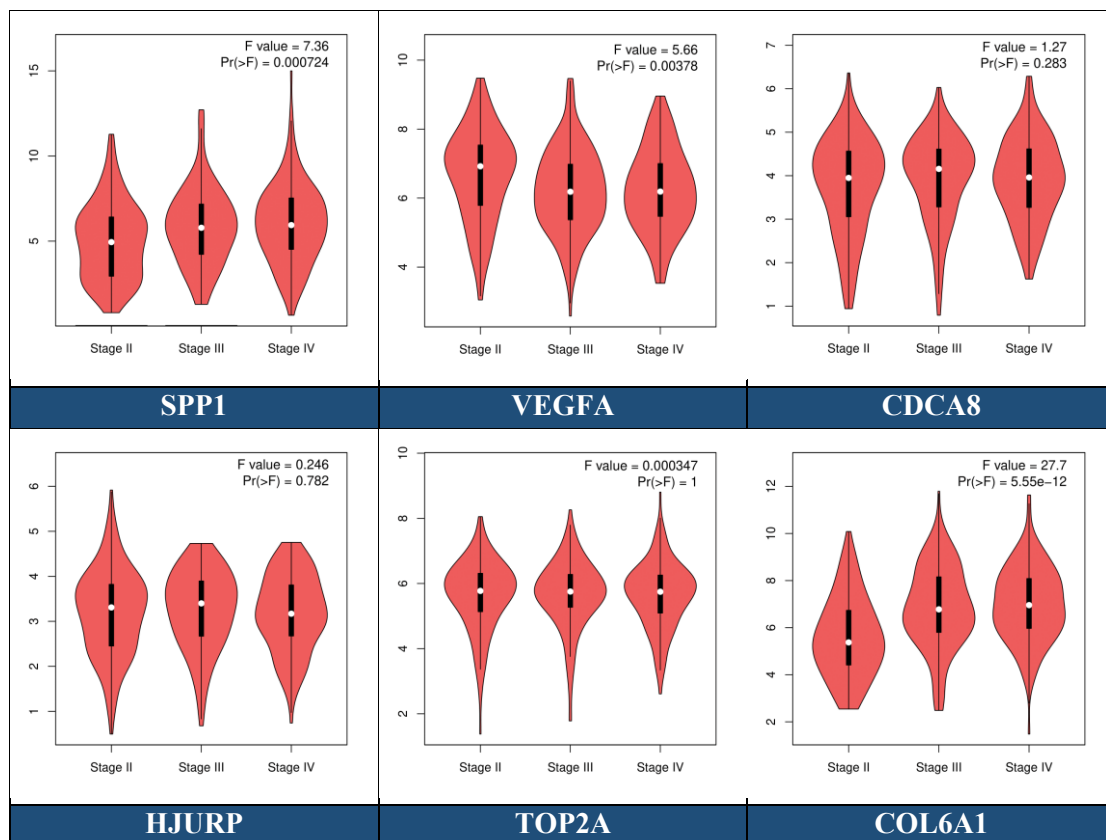


Εικόνα 63: Η ανάλυση του επιπέδου της γονιδιακής έκφρασης των γονιδίων *ANXA5* και *COL3A1* σε ασθενείς με ΚΟΚ για τους μη-θηλώδεις και θηλώδεις υπότυπους, που δημιουργήθηκε με τη χρήση της πλατφόρμας GEPIA2. Τα κόκκινα πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης mRNA σε ιστούς υποτύπων ΚΟΚ και τα γκρι πλαίσια αντιπροσωπεύουν τα επίπεδα έκφρασης σε υγιείς ιστούς ελέγχου της ουροδόχου κύστης από ασθενείς των κοορτών TCGA και GTEx.

Ως κριτήρια αποκοπής για τη στατιστική σημαντικότητα εφαρμόστηκαν τα: τιμή- $p < 0,05$ και $|\log_2FC| < 1$.
BLCA: καρκίνος της ουροδόχου κύστης, TPM: Αριθμός μεταγραφών ανά εκατομμύριο.

Σε μια προσπάθεια να διερευνηθεί η συσχέτιση των προτεινόμενων βιοδεικτών με τα διάφορα παθολογικά στάδια του ΚΟΚ, χρησιμοποιήθηκαν τα δεδομένα από τους ασθενείς με ΚΟΚ της κοόρτης TCGA και η αντίστοιχη λειτουργία από την πλατφόρμα GEPIA2. Η ανάλυση έδειξε ότι πέντε από τα εννέα γονίδια συσχετίζονται έντονα με τα **παθολογικά στάδια του ΚΟΚ**, αναδεικνύοντας έτσι και την προγνωστική τους αξία όσον αφορά τον ΚΟΚ. Συγκεκριμένα, τα γονίδια *ANXA5*, *COL3A1*, *SPPI1*, *VEGFA* και *COL6A1* εντοπίστηκαν να συσχετίζονται σε μεγάλο βαθμό με τα στάδια του ΚΟΚ, ενώ για τα υπόλοιπα δεν βρέθηκε σημαντική συσχέτιση (Εικόνα 64).

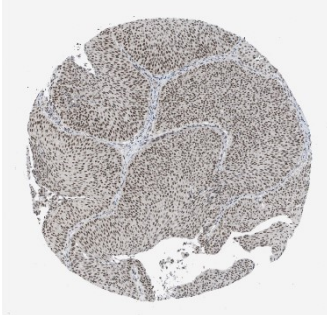
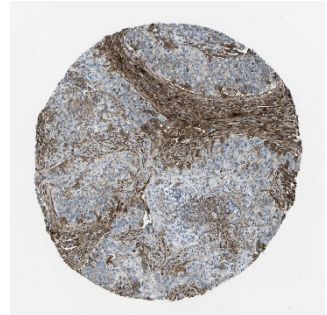
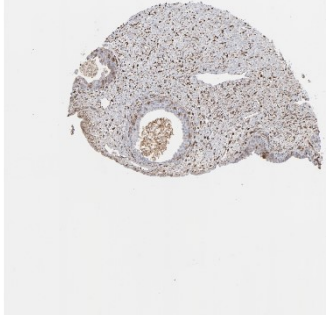
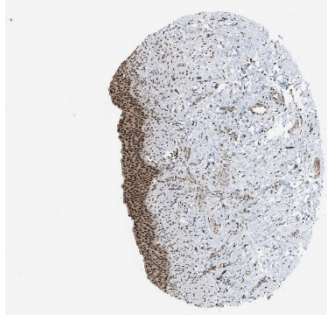
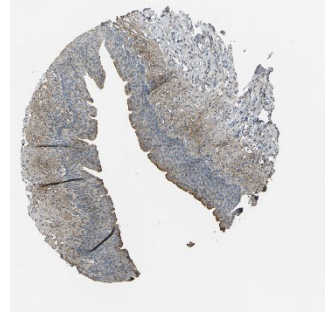
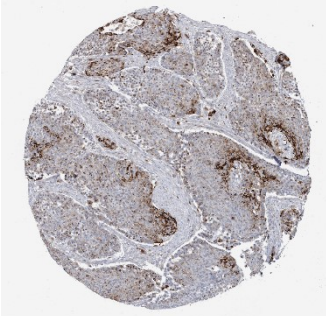
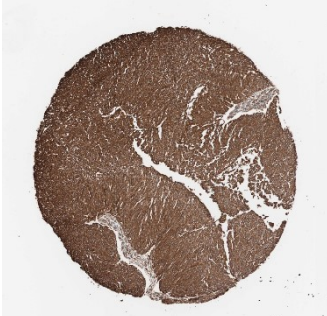
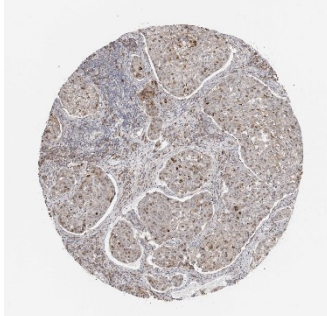


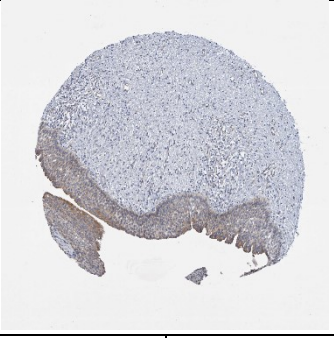
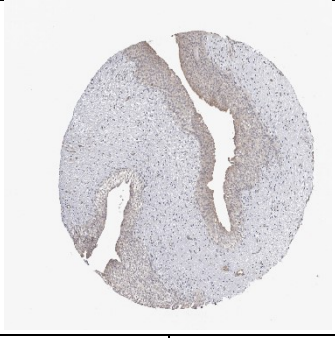
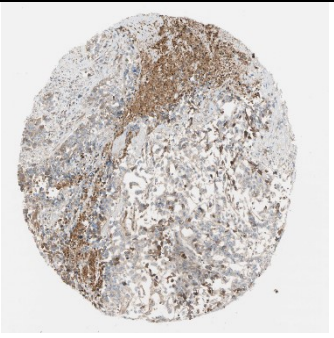
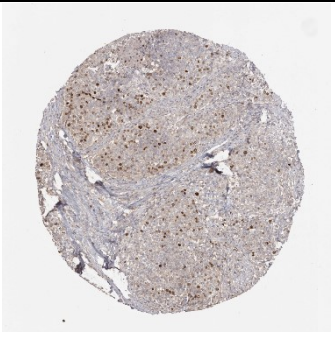
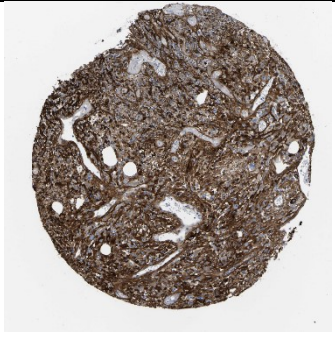
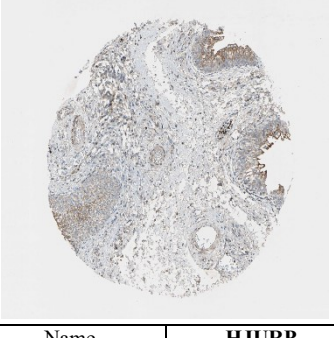
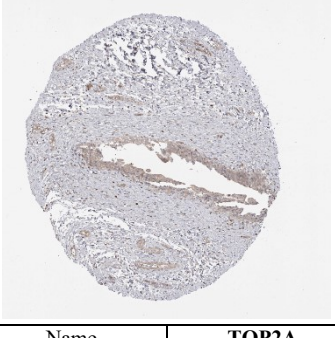
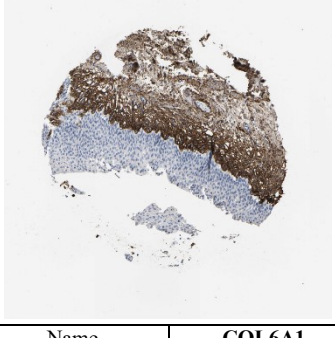


Εικόνα 64: Διαγράμματα βιολιού που αντιπροσωπεύουν τη συσχέτιση μεταξύ των επιπέδων έκφρασης των εννέα προτεινόμενων βιοδεικτών – γονιδίων και των τριών παθολογικών σταδίων του ΚΟΚ σε ασθενείς με ΚΟΚ, τα οποία δημιουργήθηκαν με τη χρήση της πλατφόρμας GEPIA2 και με βάση την κοόρτη TCGA.

Οι τιμές F και Pr(>F) υποδηλώνουν την τιμή του στατιστικού F-τεστ και την τιμή-p του τεστ, αντίστοιχα.

Η παρούσα έρευνα δεν συμπεριελάμβανε κάποια κλινική επικύρωση των αποτελεσμάτων. Για το λόγο αυτό, αξιοποιήθηκε ο Άτλαντας των Ανθρώπινων Πρωτεϊνών (HPA) για να εξεταστούν τα επίπεδα έκφρασης των πρωτεϊνών που κωδικοποιούνται από τους προτεινόμενους βιοδείκτες στον ιστό της ουροδόχου κύστης, τόσο για τις παθολογικές όσο και για τις φυσιολογικές καταστάσεις. Η ανάλυση ανοσοϊστοχημικών (IHC) εικόνων από τον HPA αποκάλυψε ότι τα γονίδια **SPPI**, **CDCA8** και **TOP2A** παρουσίασαν υψηλή ένταση χρώσης αντισώματος στους ιστούς με ΚΟΚ και χαμηλή ένταση χρώσης στους φυσιολογικούς ιστούς. Επιπλέον, τα γονίδια **ANXA5**, **COL3A1**, **HJURP** είχαν μεσαία ένταση χρώσης στους καρκινικούς ιστούς, σε αντίθεση με την χαμηλή ένταση στη φυσιολογική ουροδόχο κύστη. Τα γονίδια **CDT1** και **VEGFA** παρουσίασαν υψηλή ένταση χρώσης τόσο στους ιστούς με ΚΟΚ όσο και στους φυσιολογικούς ιστούς της ουροδόχου κύστης. Τέλος, η ένταση του αντισώματος για το **COL6A1** ήταν υψηλότερη στον ιστό με ΚΟΚ σε σύγκριση με τον αντίστοιχο φυσιολογικό ιστό, στον οποίο δεν ανιχνεύθηκε καμία χρώση. Η ανοσοϊστοχημική ανάλυση ανέδειξε ότι τα επίπεδα έκφρασης αυτών των πρωτεϊνών ήταν γενικά αυξημένα σε επίπεδο πρωτεϊνικής έκφρασης στον ΚΟΚ γεγονός που συνηγορεί στην επικύρωση των αποτελεσμάτων της έρευνας (Εικόνα 65).

ΚΟΚ						
	Name	ANXA5	Name	CDT1	Name	COL3A1
	Antibody	CAB003677	Antibody	HPA017224	Antibody	HPA007583
	Staining	Medium	Staining	High	Staining	Medium
	Intensity	Moderate	Intensity	Strong	Intensity	Moderate
	Quantity	>75%	Quantity	>75%	Quantity	75%-25%
	Patient ID	1906	Patient ID	1774	Patient ID	3517
Φυσιολογικά						
	Name	ANXA5	Name	CDT1	Name	COL3A1
	Staining	Low	Staining	High	Staining	Low
	Intensity	Moderate	Intensity	Strong	Intensity	Moderate
	Quantity	<25%	Quantity	>75%	Quantity	<25%
	Patient ID	2472	Patient ID	3265	Patient ID	3265
ΚΟΚ						
	Name	SPP1	Name	VEGFA	Name	CDCA8
	Antibody	HPA027541	Antibody	CAB005429	Antibody	HPA028258
	Staining	High	Staining	High	Staining	High
	Intensity	Strong	Intensity	Strong	Intensity	Strong
	Quantity	75%-25%	Quantity	>75%	Quantity	75%-25%
	Patient ID	3516	Patient ID	3265	Patient ID	2053

Φυσιολογικά						
	Name	SPPI	Name	VEGFA	Name	CDCA8
	Staining	Low		High		Low
	Intensity	Weak		Strong		Weak
	Quantity	>75%		>75%		75%-25%
	Patient ID	2198		1921		2198
ΚΟΚ						
	Name	HJURP	Name	TOP2A	Name	COL6A1
	Antibody	HPA008436		HPA026773		HPA019142
	Staining	Medium		High		Medium
	Intensity	Moderate		Strong		Moderate
	Quantity	75%-25%		75%-25%		>75%
	Patient ID	2031		2031		3184
Φυσιολογικά						
	Name	HJURP	Name	TOP2A	Name	COL6A1
	Staining	Low		Low		Not detected
	Intensity	Moderate		Weak		Negative
	Quantity	<25%		75%-25%		None
	Patient ID	3316		3534		3517

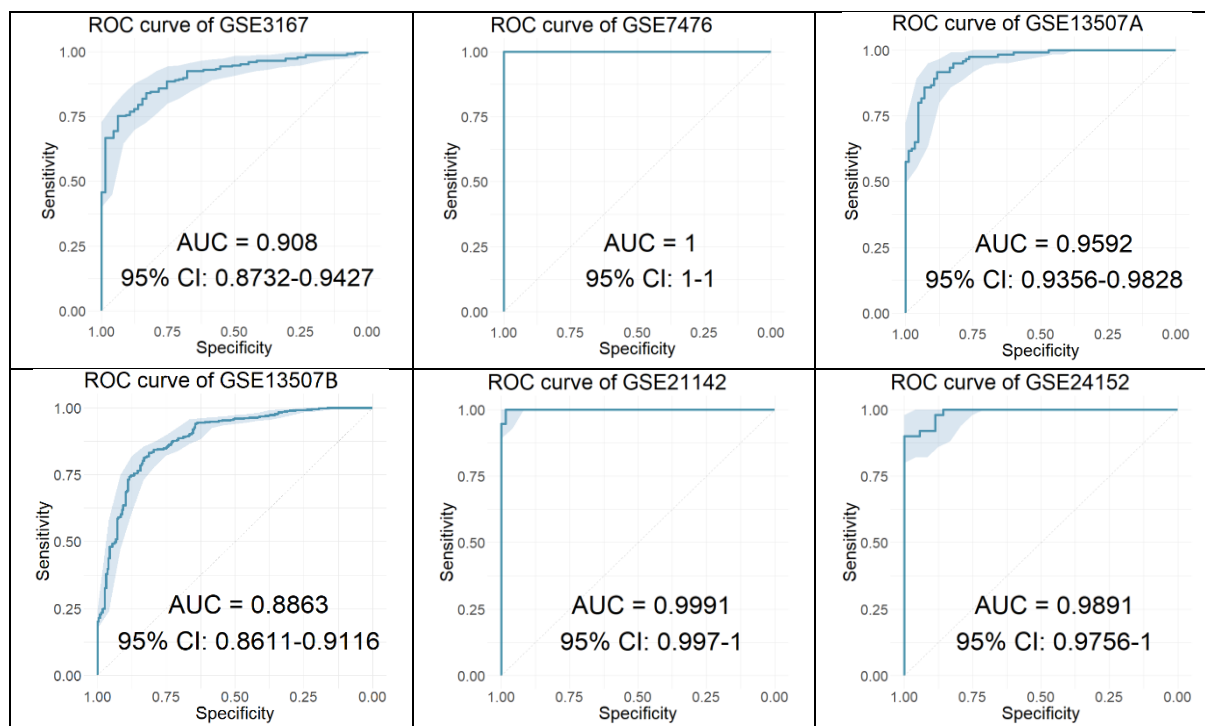
Εικόνα 65: Ανοσοϊστοχημική (IHC) επικύρωση των εννέα προτεινόμενων βιοδεικτών – γονιδίων σε δείγματα ανθρώπινου καρκινικού και φυσιολογικού ιστού της ουροδόχου κύστης, που προέρχονται από το Human Protein Atlas (www.proteinatlas.org). Η χρώση έδειξε ότι η πρωτεϊνική έκφραση των εννέα βιοδεικτών – γονιδίων ήταν υψηλότερη στους ιστούς με ΚΟΚ σε σύγκριση με δείγματα φυσιολογικών ιστών ουροδόχου κύστης.

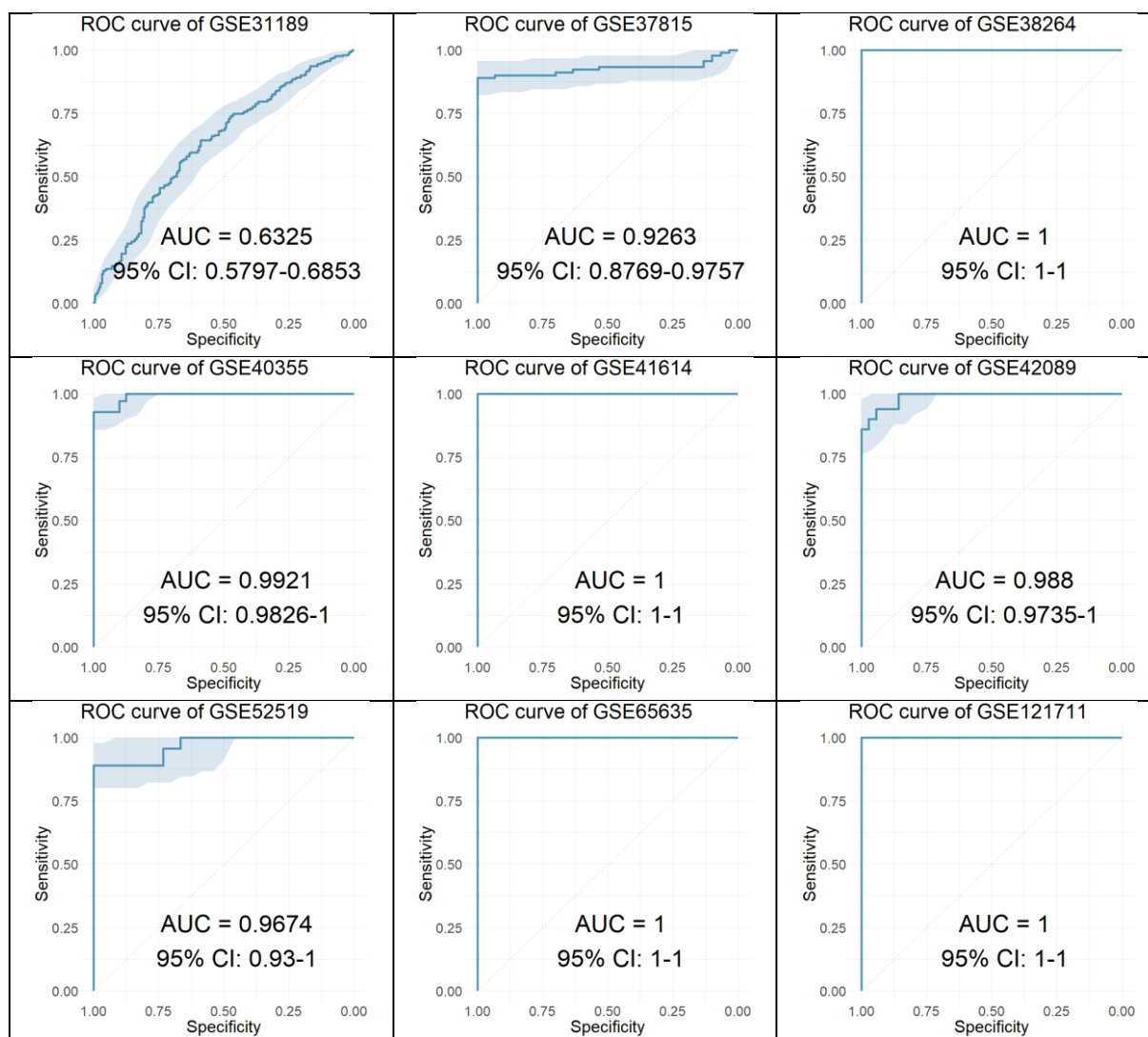
(Πηγή: <https://www.proteinatlas.org/>)

3.13 Διαγνωστική Απόδοση των Γονιδίων / Βιοδεικτών για τον Καρκίνο της Ουροδόχου Κύστης

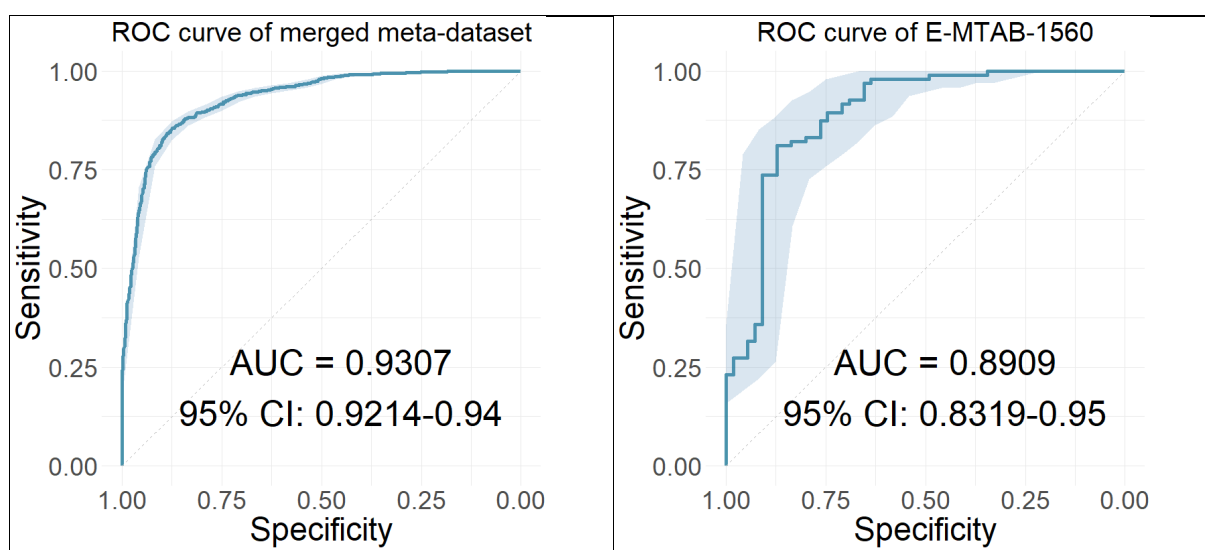
Με σκοπό να προσδιοριστεί η διαγνωστική αξία της προτεινόμενης γονιδιακής υπογραφής των εννέα προτεινόμενων βιοδεικτών – γονιδίων, χρησιμοποιήθηκαν τα γονίδια αυτά ως χαρακτηριστικά και κατασκευάστηκαν διάφορα μοντέλα ταξινόμησης, αξιοποιώντας όλα τα σύνολα δεδομένων (με περισσότερα από 10 δείγματα) που χρησιμοποιήθηκαν σε αυτή την ενοποιητική μετα-ανάλυση (Πίνακας 2), καθώς και το ενοποιημένο μετα-σύνολο δεδομένων αλλά και ένα εξωτερικό ανεξάρτητο σύνολο δεδομένων (ArrayExpress E-MTAB-1560).

Για τα μεμονωμένα σύνολα δεδομένων εφαρμόστηκε η 5-πλη διασταυρούμενη επικύρωση, ενώ για το τελικό ενοποιημένο μετα-σύνολο δεδομένων πραγματοποιήθηκε 10-πλη διασταυρούμενη επικύρωση, τα οποία επαναλήφθηκαν όλα 10 φορές. Οι προκύπτουσες καμπύλες ROC για όλα τα μοντέλα που κατασκευάστηκαν, μαζί με τις αντίστοιχες τιμές AUC και τα 95% διαστήματα εμπιστοσύνης, απεικονίστηκαν γραφικά (Εικόνα 66 και Εικόνα 67). Τα αποτελέσματα έδειξαν πολύ υψηλές διαγνωστικές επιδόσεις για τα διάφορα μοντέλα, με τις τιμές AUC να κυμαίνονται από 0,8863 έως 1,00 για τα μεμονωμένα σύνολα δεδομένων και να φτάνουν το 0,9307 και το 0,8909 για το ενοποιημένο μετα-σύνολο δεδομένων και το εξωτερικό σύνολο δεδομένων, αντίστοιχα. Το μοντέλο ταξινόμησης που κατασκευάστηκε από τη σειρά GSE31189 εμφάνισε τιμή AUC = 0,6325, καθώς το εν λόγω σύνολο δεδομένων παρουσίαζε σημαντικές επιδράσεις από φαινόμενα παρτίδας (όπως αναφέρεται στην Ενότητα 3.4) και δεν λήφθηκε υπόψη στη συνολική αξιολόγηση.





Εικόνα 66: Επικύρωση των διαγνωστικών μοντέλων, χρησιμοποιώντας τα εννέα γονίδια – βιοδείκτες ως χαρακτηριστικά, με ανάλυση καμπύλης ROC σε κάθε ένα από τα σύνολα δεδομένων που συμπεριλήφθηκαν στην ενοποιητική μετα-ανάλυση και περιείχαν τουλάχιστον 10 δείγματα (Πίνακας 2).



Εικόνα 67: Επικύρωση των διαγνωστικών μοντέλων, χρησιμοποιώντας τα εννέα γονίδια – βιοδείκτες ως χαρακτηριστικά, με ανάλυση καμπύλης ROC στο ενοποιημένο μετα-σύνολο δεδομένων ($n = 606$) και στο εξωτερικό σύνολο δεδομένων επικύρωσης E-MTAB-1650 ($n = 30$).

4. Συζήτηση

Ο ΚΟΚ αποτελεί έναν από τους πιο κοινούς τύπους καρκίνου παγκοσμίως (ακόμα κι αν αυτό δεν είναι ευρέως γνωστό), παρουσιάζει υψηλή επίπτωση, επιπολασμό, θνησιμότητα και ποσοστό υποτροπής, και εξακολουθεί να παραμένει ένα ανοιχτό κλινικό και κοινωνικό πρόβλημα. Η κατανόηση της παθοφυσιολογίας του έχει σαφώς βελτιωθεί τα τελευταία χρόνια, αλλά οι υποκείμενοι μοριακοί μηχανισμοί και το γενετικό του υπόβαθρο επιβάλλεται να αποσαφηνιστούν περαιτέρω. Σημαντικό εμπόδιο προς την κατεύθυνση αυτή αποτελεί το γεγονός ότι η ανίχνευσή του παραμένει σχετικά δύσκολη, λόγω της έλλειψης ειδικών και ευαίσθητων καρκινικών βιοδεικτών και της απουσίας νέων συμπτωμάτων κατά την εμφάνιση και εξέλιξη της νόσου. Το ζήτημα της ανεύρεσης νέων βιοδεικτών έχει καταστεί ακόμη πιο επιτακτικό κατά την εποχή της πανδημίας COVID-19, όπου παρατηρούνται σημαντικές καθυστερήσεις στην έγκαιρη διάγνωση και αντιμετώπιση των ασθενών [202]. Συνεπώς, αναδεικνύεται μία επιτακτική ανάγκη για την ανάπτυξη αποτελεσματικότερων διαγνωστικών, προγνωστικών και προβλεπτικών βιοδεικτών, προκειμένου να αντιμετωπιστεί καλύτερα η εμφάνιση και η εξέλιξη του ΚΟΚ.

Στην παρούσα μελέτη χρησιμοποιήθηκαν δεδομένα πειραμάτων μικροσυστοιχιών για να διερευνηθεί το προφίλ γονιδιακής έκφρασης στον ΚΟΚ. Στη μελέτη αυτή εντοπίστηκε, ενοποιήθηκε και αναλύθηκε μεγάλος αριθμός δειγμάτων από πολλαπλά πειράματα, στοχεύοντας στην κατάληξη σε πιο αξιόπιστα αποτελέσματα, έγκυρα στατιστικά συμπεράσματα και πληροφοριακές υπογραφές γονιδιακής έκφρασης. Προκειμένου να προσδιοριστεί μία έγκυρη λίστα διαφορά εκφραζόμενων γονιδίων για τον ΚΟΚ, πραγματοποιήθηκε συστηματική ανασκόπηση σε πολλαπλές προγενέστερες μελέτες, στο αποθετήριο GEO, χρησιμοποιώντας τις κατευθυντήριες γραμμές του PRISMA (Εικόνα 2). Βάσει των κριτηρίων ένταξης, επιλέχθηκαν τα σύνολα δεδομένων που πληρούσαν τα κριτήρια αυτά, λήφθηκαν όλα τα πρωτογενή δεδομένα από το GEO και το κάθε σύνολο δεδομένων προεπεξεργάστηκε σύμφωνα με την πλατφόρμα με την οποία είχε διενεργηθεί η ανάλυση. Έπειτα, ελέγχθηκε η ποιότητα όλων των δειγμάτων, αφαιρέθηκαν οι ακραίες τιμές και, τέλος, ορίστηκε ένα κοινό σύνολο γονιδιακών συμβόλων για όλα τα σύνολα δεδομένων. Τέλος, δημιουργήθηκε ένα ενοποιημένο μετα-σύνολο δεδομένων, που περιελάμβανε 410 δείγματα ιστού ΚΟΚ και 196 υγιή δείγματα ιστού ουροδόχου κύστης, από 18 ανεξάρτητα σύνολα δεδομένων, υιοθετώντας μια προσέγγιση ενοποίησης «πρώιμου σταδίου» (early-stage) [128]. Έτσι, η συγκεκριμένη ενοποιητική μετα-ανάλυση, η οποία είναι από τις μεγαλύτερες του είδους της βάσει της βιβλιογραφίας, εντόπισε 815 διαφορετικά εκφραζόμενα γονίδια μεταξύ των ιστών του ΚΟΚ και των φυσιολογικών ιστών.

Έπειτα, διερευνήθηκαν οι οδοί που υπερ-εκπροσωπούνται σημαντικά στη λίστα των διαφορετικά εκφραζόμενων γονιδίων. Τα αποτελέσματα της ανάλυσης με βάση την Οντολογία Γονιδίων αποκάλυψαν βιολογικές διεργασίες που σχετίζονται με την εξωκυτταρική μήτρα, την αγγειογένεση, την ανάπτυξη των μυών, την κυτταρική διαίρεση, την οργάνωση των χρωμοσωμάτων και την αντιγραφή του DNA, οι οποίες είναι όλες θεμελιώδεις διεργασίες για την ανάπτυξη και την εξέλιξη του καρκίνου (Εικόνα 35) [203]. Η ανάλυση εμπλουτισμού μονοπατιών με βάση την KEGG αποκάλυψε μονοπάτια εμπλουτισμένα μεταξύ άλλων στη σηματοδότηση PI3K-PKB/Akt, στα microRNAs στον καρκίνο, στον κυτταρικό κύκλο, στην εστιακή προσκόλληση, στη ρύθμιση του κυτταροσκελετού της ακτίνης, στη σηματοδότηση ασβεστίου, στις πρωτεογλυκάνες στον καρκίνο, στην κυτταρική γήρανση, στη συστολή των λείων μυών των αγγείων και στον καρκίνο της ουροδόχου κύστης (Εικόνα 36). Η ανάλυση των μονοπατιών με βάση τη Reactome έδειξε μονοπάτια εμπλουτισμένα στην οργάνωση της εξωκυτταρικής μήτρας, στα σημεία ελέγχου του κυτταρικού κύκλου, στους τελεστές της Rho GTPάσης, στον έλεγχο

της μεταφοράς του ινσουλινόμορφου αυξητικού παράγοντα (IGF), στην αντιγραφή του DNA, στην αποδιάταξη των αιμοπεταλίων και στην αποικοδόμηση του κολλαγόνου (Εικόνα 37). Τέλος, η ανάλυση εμπλουτισμού με βάση την Οντολογία Ασθενειών εμφάνισε τον καρκίνο του ουροποιητικού συστήματος ως τον πιο σημαντικά εμπλουτισμένο όρο, ακολουθούμενο από το μη μικροκυτταρικό καρκίνωμα του πνεύμονα, καθώς και τον καρκίνο των νεφρών, του μαστού, του μυοσκελετικού συστήματος, του νεφρού και του προστάτη (Εικόνα 38), υποδεικνύοντας την υψηλή συσχέτιση των εντοπισμένων διαφορεικά εκφραζόμενων γονιδίων με τον ΚΟΚ. Επίσης, η υψηλή συσχέτιση των εμπλουτισμένων όρων που περιλαμβάνουν άλλες μορφές καρκίνου αποκαλύπτει τους κοινούς μηχανισμούς και τα κοινά χαρακτηριστικά των διαφόρων τύπων καρκίνου μεταξύ τους [204].

Στην παρούσα μελέτη, συνδυάστηκαν τα αποτελέσματα από τις μεθόδους ανάλυσης του δικτύου ΑΠΠ και των δικτύων συνέκφρασης γονιδίων προκειμένου να εντοπιστούν τα γονίδια – κλειδιά για την εμφάνιση και την ανάπτυξη του ΚΟΚ (Πίνακας 5). Η συναινετική Ανάλυση Δικτύων Συνέκφρασης Γονιδίων (WGCNA) δημιούργησε ένα δίκτυο που βασίζεται στη συσχέτιση μεταξύ των γονιδίων, καθώς πρόκειται για μη επιβλεπόμενη μέθοδο εκμάθησης, ενώ το δίκτυο ΑΠΠ δημιουργήθηκε με βάση τις γνωστές αλληλεπιδράσεις μεταξύ των ανθρώπινων πρωτεϊνών. Η ανάλυση του δικτύου ΑΠΠ οδήγησε σε ένα πυκνά συνδεδεμένο δίκτυο πρωτεϊνών, το οποίο υποδεικνύει υψηλή βιολογική σημασία (Εικόνα 39). Στην Ανάλυση Δικτύων Συνέκφρασης Γονιδίων, παρά την μετα-ανάλυση διαφορετικών και ετερογενών συνόλων δεδομένων, προσδιορίστηκαν υψηλοί συσχετισμοί μεταξύ των συναινετικών βασικών ομάδων και των φαινοτυπικών χαρακτηριστικών (Εικόνα 44). Αξίζει να σημειωθεί ότι τα γονίδια – κόμβοι που περιέχονταν στην καφέ ομάδα βρέθηκαν να έχουν ισχυρή συσχέτιση με τη συνολική επιβίωση και την ελεύθερης νόσου επιβίωση των ασθενών με ΚΟΚ, ενώ τα γονίδια της μαύρης ομάδας είχαν συνάφεια με την ελεύθερης νόσου επιβίωση των ασθενών. Ο συνδυασμός των γονιδίων – κόμβων που εντοπίστηκαν με αυτές τις δύο μεθόδους οδήγησε σε 61 κοινά γονίδια που χαρακτηρίστηκαν ως γονίδια – κόμβοι – κλειδιά για την παρούσα ανάλυση.

Ένα κρίσιμο τρέχον ζήτημα στην έρευνα του ΚΟΚ είναι η δυνατότητα εύκολης και έγκαιρης ανίχνευσής του με λιγότερο επεμβατικές μεθόδους και, ιδανικά, με δείκτες που παρουσιάζουν υψηλή ευαισθησία και ειδικότητα [205]. Οι μοριακοί δείκτες, όπως τα κυκλοφορούντα mRNAs, στα ούρα και στο αίμα θα μπορούσαν να προσφέρουν πολλά υποσχόμενες πηγές για την κατανόηση του ΚΟΚ, καθώς και του σχετικού μικρο- και μακρο-περιβάλλοντός του. Ως εκ τούτου, εξετάστηκε κατά πόσον καθένα από τα βασικά γονίδια – κόμβους εκφράζεται διαφορετικά στα ούρα ή στο πλάσμα αίματος ασθενών με ΚΟΚ. Στα δείγματα ούρων, 17 γονίδια, και συγκεκριμένα τα γονίδια *AURKB*, *CCNB2*, *CDC45*, *CDC48*, *CDT1*, *CENPU*, *COL3A1*, *GINS2*, *KIF20A*, *MCM4*, *PBK*, *PLK4*, *SDC1*, *SPP1*, *TOP2A*, *TTK*, *UBE2C*, παρουσίασαν στατιστικά σημαντική διαφορική έκφραση μεταξύ των ασθενών με ΚΟΚ και των υγιών ατόμων. Σε προηγούμενες μελέτες, ορισμένα από αυτά τα γονίδια έχουν εντοπιστεί σε διάφορες καταστάσεις και ασθένειες. Πιο συγκεκριμένα, η οστεοποντίνη (Secreted Phosphoprotein 1 – *SPP1*) διερευνήθηκε στα ούρα ασθενών με νεφρολιθίαση [206], ασθενών με τη νόσο Alzheimer [207] καθώς και σε καρκινοπαθείς που παρουσίαζαν νεφροτοξικότητα οφειλόμενη στη σισπλατίνη [208], και βρέθηκε ότι παρέχει διαγνωστική αξία. Η συνδεκάνη 1 (Syndecan 1 – *SDC1*) μετρήθηκε με ανοσοπροσδιορισμό πολλαπλών αναλύσεων μαζί με εννέα άλλους πρωτεϊνικούς βιοδείκτες για τη διάγνωση του ΚΟΚ [209] και για την ανίχνευση του υποτροπιάζοντος ΚΟΚ [210]. Ακόμη, το ένζυμο σύζευξης της ουβικιτίνης E2C (Ubiquitin Conjugating Enzyme E2 C – *UBE2C*) αναλύθηκε σε δείγματα ούρων ως δυναμικός διαγνωστικός δείκτης για τον ΚΟΚ [211]. Επίσης, διερευνήθηκε η δημιουργία προφίλ πεπτιδίων στα ούρα, χρησιμοποιώντας ένα πάνελ 22 δεικτών, συμπεριλαμβανομένης της άλφα μονής αλυσίδας κολλαγόνου τύπου III (Collagen Type III Alpha 1 Chain – *COL3A1*), για την κλινική διάγνωση της προεκλαμψίας [212]. Τέλος, η συντήρηση του μινιχρωμοσώματος 5 (Minichromosome Maintenance Complex Component 5 – *MCM5*), μια πρωτεΐνη της ίδιας οικογένειας με την *MCM4*,

μετρήθηκε σε δείγματα ούρων με τη χρήση ανοσοφθορομετρικής ανάλυσης για τη διάγνωση του καρκίνου του ουροποιητικού συστήματος [213]. Πέραν από τις ανωτέρω πρωτεΐνες, η βιβλιογραφία για τους υπόλοιπους βιοδείκτες ούρων είναι εξαιρετικά περιορισμένη, εάν υπάρχει. Ως εκ τούτου, τα γονίδια που εντοπίστηκαν ως πιθανοί βιοδείκτες στα ούρα θα πρέπει να διερευνηθούν για την πιθανή διαγνωστική τους αξία σε ασθενείς με ΚΟΚ.

Όσον αφορά τα δείγματα πλάσματος αίματος, 11 γονίδια, και συγκεκριμένα τα *ANXA5*, *ASPM*, *CD34*, *CDC20*, *CDT1*, *COL4A5*, *COL6A1*, *ECT2*, *HJURP*, *MCM2* και *VEGFA*, παρουσίασαν στατιστικά σημαντική διαφορική έκφραση μεταξύ των ασθενών με ΚΟΚ και των υγιών ατόμων. Σε ευρήματα προγενέστερων μελετών, τα επίπεδα της αννεξίνης A5 (Annexin A5 – *ANXA5*) στο πλάσμα διερευνήθηκαν ως δυνητικός βιοδείκτης για τη διάγνωση του άσθματος [214], για τη διάκριση μεταξύ έγκυων και μη έγκυων γυναικών [215], καθώς και για τη διάγνωση της κίρρωσης του ήπατος και το ηπατοκυτταρικό καρκίνωμα [216]. Ακόμη, η ομόλογος πρωτεΐνη (Assembly Factor For Spindle Microtubules – *ASPM*) ανιχνεύθηκε σε κυκλοφορούντα καρκινικά κύτταρα μέσω γονιδιωματικού χαρακτηρισμού ενός κυττάρου (single-cell genomic) σε ασθενείς με καρκίνο [217]. Το γονίδιο *CD34* χρησιμεύει ως βασικός βιοδείκτης στην έρευνα διαφόρων ασθενειών, καθώς χρησιμοποιείται συνήθως για τον εντοπισμό και την απομόνωση ανθρώπινων αιμοποιητικών βλαστικών / προγονικών κυττάρων που εφαρμόζονται στη μεταμόσχευση μυελού των οστών. Λόγω της υψηλής ευαισθησίας του όσον αφορά τη διαφοροποίηση των ενδοθηλιακών κυττάρων, έχει επίσης μελετηθεί ως δείκτης για τη διάγνωση του καρκίνου [218]. Επιπλέον, βρέθηκε ότι τα εξωκυτταρικά mRNA της πρωτεΐνης αναγνώρισης διακλάδωσης Holliday (Holliday Junction Recognition Protein – *HJURP*) εκφράζονται σε σημαντικά επίπεδα στο πλάσμα αίματος ασθενών με καρκίνο του πνεύμονα [219]. Η πρωτεΐνη του αγγειακού ενδοθηλιακού αυξητικού παράγοντα A (Vascular Endothelial Growth Factor A – *VEGFA*) διαδραματίζει σημαντικό ρόλο στην ανάπτυξη των αιμοφόρων αγγείων και, ως εκ τούτου, στις ασθένειες που τα αφορούν. Στις ασθένειες αυτές περιλαμβάνονται οι καρδιακές παθήσεις [220], η ασθένεια COVID-19 [221] και διάφοροι τύποι καρκίνου [222], όπως ο καρκίνος των ωοθηκών [223] αλλά και ο καρκίνος του μαστού [224]. Συμπερασματικά, φαίνεται ότι η αξία των περισσότερων από αυτούς τους δυνητικούς μοριακούς αιματολογικούς βιοδείκτες παραμένει ασαφής για τον ΚΟΚ. Ωστόσο, η περαιτέρω κλινική διερεύνησή τους μπορεί να προσφέρει σημαντικές προοπτικές για τη βελτίωση της κατανόησης του ΚΟΚ, καθώς και να βοηθήσει στην έγκαιρη αναγνώρισή του, τη διαστρωμάτωση των ασθενών και την αποδοτικότερη πρόβλεψη της έκβασης.

Με σκοπό να αξιολογηθεί η προγνωστική αξία των βασικών γονιδίων – κόμβων, πραγματοποιήθηκε μονο-μεταβλητή ανάλυση παλινδρόμησης Cox, ανάλυση παλινδρόμησης LASSO και πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox. Με βάση τα γονίδια και τους συντελεστές που προέκυψαν από την πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox, δημιουργήθηκε ένα προγνωστικό μοντέλο τριών γονιδίων για ασθενείς με ΚΟΚ, το οποίο αποτελείται από τα γονίδια *COL3A1*, *FOXMI* και *PLK4*. Όσον αφορά τη βιβλιογραφία, οι μεταβολές της έκφρασης του γονιδίου *COL3A1* βρέθηκαν να αποτελούν προγνωστικούς δείκτες για τον ΚΟΚ [225] και να εμπλέκονται στην ανάπτυξη του μυοδιαθητικού ΚΟΚ [226]. Επίσης, το γονίδιο αυτό προσδιορίστηκε ως πιθανός βιοδείκτης για τον ΚΟΚ στην παρούσα μελέτη και αναλύεται περαιτέρω σε επόμενη παράγραφο. Η πρωτεΐνη Forkhead box M1 (Forkhead Box M1 – *FOXMI*) έχει βρεθεί ότι συμμετέχει σε έναν άξονα που ρυθμίζει τη διαδικασία του κυτταρικού κύκλου και προάγει την εξέλιξη του ΚΟΚ [227] και ότι αποτελεί ισχυρό προγνωστικό δείκτη για την εξέλιξη της νόσου, ειδικότερα του μυοδιαθητικού ΚΟΚ [123, 228]. Έχει διαπιστωθεί επίσης ότι διαδραματίζει σημαντικό ρόλο στην υποτροπή του ΚΟΚ και στην αντίσταση των ασθενών στην αντικαρκινική φαρμακευτική θεραπεία [229]. Τέλος, το ένζυμο polo-like κινάση 4 (Polo Like Kinase 4 – *PLK4*) έχει χαρακτηριστεί από μελέτες ως σημαντικός ρυθμιστής του πολλαπλασιασμού των

κυττάρων στον ΚΟΚ και, ως εκ τούτου, ως πιθανός νέος μοριακός στόχος για τη θεραπεία του ΚΟΚ [230].

Το προγνωστικό μοντέλο των τριών γονιδίων πέτυχε τιμές AUC των χρονοεξαρτώμενων καμπυλών ROC για το 1^ο / 3^ο / 5^ο έτος μετά τη διάγνωση 0,796 / 0,779 / 0,846 για το σύνολο εκπαίδευσης (GSE13507), 0,82 / 0,806 / 0,774 για το πρώτο σύνολο ελέγχου (GSE32548), και 0,819 / 0,86 / 0,871 για το δεύτερο σύνολο ελέγχου (GSE32894), αντίστοιχα. Στο ίδιο πλαίσιο, ο L. Yang κ.α. πρότειναν ένα προγνωστικό μοντέλο εννέα γονιδίων για να βελτιώσουν την πρόβλεψη της πρόγνωσης του ΚΟΚ, επιτυγχάνοντας τιμή AUC = 0,76 για το 5^ο έτος μετά τη διάγνωση, στο σύνολο εκπαίδευσης, και τιμή AUC = 0,63 για το ίδιο χρονικό διάστημα στο σύνολο ελέγχου [231]. Σε μια άλλη μελέτη, ο Z. Xie κ.α. πρότειναν ένα προγνωστικό μοντέλο 10 γονιδίων που σχετίζονταν με τη φλεγμονώδη απόκριση, το οποίο πέτυχε τιμές AUC 0,71 και 0,67 για το πρώτο έτος μετά τη διάγνωση που αφορούσα τις ομάδες ασθενών TCGA και GSE13507, αντίστοιχα [232]. Επιπλέον, ο J. Lin κ.α. κατασκεύασαν ένα προγνωστικό μοντέλο 11 γονιδίων για την πρόβλεψη της συνολικής επιβίωσης ασθενών με ΚΟΚ, επιτυγχάνοντας τιμές AUC για το 1^ο / 3^ο / 5^ο έτος μετά τη διάγνωση 0,686 / 0,665 / 0,666 για την κοόρτη ασθενών TCGA, 0,800 / 0,742 / 0,697 για το σύνολο δεδομένων GSE13507, 0,826 / 0,792 / 0,763 για το σύνολο δεδομένων GSE32548, και 0,781 / 0,831 / 0,839 για το σύνολο δεδομένων GSE32894 [233]. Ο Tang κ.α. ανέπτυξαν μια υπογραφή επτά γονιδίων προκειμένου να προβλέψουν την πρόγνωση των ασθενών με ΚΟΚ, επιτυγχάνοντας τιμές AUC για το 1^ο / 3^ο / 5^ο έτος 0,711 / 0,714 / 0,711 στα σύνολο εκπαίδευσης και 0,608 / 0,680 / 0,638 στο σύνολο δοκιμής [234]. Επιπλέον, ο C. Zhou πρότεινε μια υπογραφή 11 γονιδίων, που σχετίζονταν με την αυτοφαγία (autophagy), για την πρόβλεψη της πρόγνωσης των ασθενών με ΚΟΚ, παρουσιάζοντας προγνωστική αποτελεσματικότητα για το 1^ο / 3^ο / 5^ο έτος μετά τη διάγνωση 0,702 / 0,744 / 0,794 στο σύνολο εκπαίδευσης και 0,695 / 0,640 / 0,658 στο σύνολο ελέγχου [235]. Επιπλέον, ο F. Xu ανέπτυξε μια προγνωστική υπογραφή έξι γονιδίων για τον ΚΟΚ, παρουσιάζοντας τιμές AUC για την ειδική της νόσου επιβίωση για το 3^ο και 5^ο έτος 0,96 / 0,967 για το σύνολο εκπαίδευσης (GSE32894), 0,744 / 0,748 για το πρώτο σύνολο ελέγχου (GSE13507) και 0,576 / 0,606 για το δεύτερο σύνολο ελέγχου (TCGA), αντίστοιχα [236]. Τέλος, ο F. Chen κ.α. κατασκεύασαν ένα προβλεπτικό μοντέλο πρόγνωσης οκτώ γονιδίων για τον ΚΟΚ, το οποίο πέτυχε μέγιστες τιμές AUC 0,795 και 0,669 για τα σύνολα εκπαίδευσης και ελέγχου (TCGA), αντίστοιχα [237]. Αξίζει να σημειωθεί ότι αυτή η σύντομη αλλά περιεκτική ανασκόπηση πρόσφατων μελετών σχετικά με τα προγνωστικά μοντέλα για τον ΚΟΚ υπογραμμίζει την ιδιαίτερα υψηλή απόδοση του απλού προτεινόμενου προγνωστικού μοντέλου τριών γονιδίων της παρούσας μελέτης και αναδεικνύει την εγκυρότητά του.

Με σκοπό να αξιολογηθεί η προβλεπτική αξία των γονιδίων – κλειδιών, όσον αφορά την ανταπόκριση των ασθενών στη θεραπεία, πραγματοποιήθηκαν η μονο-μεταβλητή ανάλυση παλινδρόμησης Cox, η ανάλυση παλινδρόμησης LASSO και η πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox σε δεδομένα ελεύθερης νόσου επιβίωσης ασθενών με μυοδιηθητικό ΚΟΚ που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη. Με βάση τα γονίδια και τους συντελεστές που προέκυψαν από την πολυ-μεταβλητή ανάλυση παλινδρόμησης Cox, δημιουργήθηκε ένα μοντέλο πρόβλεψης έξι γονιδίων για ασθενείς με μυοδιηθητικό ΚΟΚ, το οποίο αποτελείται από τα γονίδια *ANXA5*, *CD44*, *NCAM1*, *SPPI*, *CDCA8* και *KIF14*. Τα *ANXA5*, *SPPI* και *CDCA8* χαρακτηρίστηκαν ως προτεινόμενοι βιοδείκτες από την παρούσα έρευνα και η λειτουργία τους καθώς και η συσχέτισή τους με τον ΚΟΚ περιγράφονται παρακάτω. Όσον αφορά τα υπόλοιπα τρία γονίδια της προβλεπτικής υπογραφής των έξι γονιδίων, έγινε αναζήτηση στη βιβλιογραφία σχετικά με προηγούμενη μελέτη τους από άλλους ερευνητές. Έτσι, τα επίπεδα έκφρασης του αντιγόνου *CD44* βρέθηκε ότι σχετίζονται με την εξέλιξη, τη μετάσταση και την υποτροπή της νόσου του ΚΟΚ [238]. Επίσης, η έκφραση του *CD44* έχει συσχετιστεί με την επιθετικότητα των όγκων ΚΟΚ [239] και με την πρόβλεψη της απόκρισης των καρκινικών κυττάρων της ουροδόχου κύστης στην

ακτινοβολία [240]. Σε άλλες έρευνες, η έκφρασή του προτάθηκε ότι αποτελεί χρήσιμο βιοδείκτη για την πρόγνωση και τις θεραπευτικές επιλογές στον πρωτογενή KOK [241]. Το γονίδιο Neural Cell Adhesion Molecule 1 (*NCAM1*) δεν έχει διερευνηθεί εκτενώς σε σχέση με τον KOK, αλλά υπάρχουν μελέτες που συνδέουν την έκφρασή του με την αντίσταση στα φάρμακα στην οξεία μυελογενή λευχαιμία [242], στο πνευμονοβλάστωμα [243] και στην ενδοεπιθηλιακή νεοπλασία του τραχήλου της μήτρας [244]. Ακόμη, τα επίπεδα έκφρασης του γονιδίου Kinesin Family Member 14 (*KIF14*) έχουν συνδεθεί με τη χημειοευαισθησία του ηπατοκυτταρικού καρκινώματος [245] και του καρκίνου του τραχήλου της μήτρας [246], καθώς και με την πρόγνωση διαφόρων άλλων καρκίνων, όπως του καρκίνου του μαστού [247] και του παγκρέατος [248].

Το προβλεπτικό μοντέλο απέδωσε τιμές AUC για τις χρονοεξαρτώμενες καμπύλες ROC για το 1^ο / 3^ο / 5^ο έτος από την διάγνωση ίσες με 0,6 / 0,69 / 0,707 για το σύνολο εκπαίδευσης (GSE149455) και 0,61 / 0,673 / 0,611 για το σύνολο δοκιμής (GSE69795), αντίστοιχα. Η βιβλιογραφία σχετικά με τα μοντέλα γονιδιακής υπογραφής για την πρόβλεψη της ανταπόκρισης των ασθενών με μυοδιηθητικό KOK στη φάση της προεγχειρητικής θεραπείας είναι περιορισμένη. Σε μια παρόμοια μελέτη, ο W. Jiang κ.α. ανέπτυξαν μια ανοσολογική υπογραφή εννέα γονιδίων, η οποία μπορούσε να προβλέψει την ανοσοθεραπευτική ανταπόκριση των αναστολέων ανοσολογικών σημείων ελέγχου (immune checkpoint inhibitors) του ανοσοποιητικού συστήματος, επιτυγχάνοντας μέγιστη τιμή AUC 0,69 και 0,64 για τις ομάδες ασθενών TCGA και IMvigor210, αντίστοιχα [249]. Ο C. Shen κ.α. κατασκεύασαν μια υπογραφή δύο γονιδίων, που σχετίζονταν με το ανοσοποιητικό σύστημα, για την πρόβλεψη της ανταπόκρισης των ασθενών με μυοδιηθητικό KOK στην ανοσοθεραπεία, επιτυγχάνοντας τιμή AUC = 0,695 όσον αφορά την προβλεπτική της ικανότητα [250]. Οι S. J. Choi κ.α. ανέπτυξαν ένα μοντέλο βασισμένο σε ακτινομικούς δείκτες (radiomics) για την πρόβλεψη της ανταπόκρισης ασθενών με μυοδιηθητικό KOK στη νεοεπικουρική χημειοθεραπεία (NAC), επιτυγχάνοντας τιμή AUC = 0,75 για το σύνολο ελέγχου [251]. Σε παρόμοιο πλαίσιο, ο A. Parmar κ.α. χρησιμοποίησαν μια προβλεπτική υπογραφή βασισμένη σε ακτινομικούς δείκτες για την ανταπόκριση των ασθενών με μυοδιηθητικό KOK στη νεοεπικουρική χημειοθεραπεία, επιτυγχάνοντας τιμή AUC = 0,63 όσον αφορά τη διάκριση των ασθενών σε ανταποκρινόμενους και μη ανταποκρινόμενους [252]. Αυτά τα ευρήματα υποδεικνύουν ότι η προβλεπτική απόδοση του μοντέλου που αναπτύχθηκε στην παρούσα έρευνα είναι επαρκής. Μέχρι σήμερα, οι προσπάθειες ανάπτυξης μοντέλων πρόβλεψης για την ανταπόκριση των καρκινικών όγκων στη νεοεπικουρική χημειοθεραπεία συνεχίζονται και δεν έχουν ακόμη εισαχθεί βιοδείκτες του προφίλ της γονιδιακής έκφρασης βασισμένοι στο mRNA, οι οποίοι να μπορούν να προβλέψουν με ακρίβεια την ανταπόκριση των ασθενών [253]. Ωστόσο, αξίζει να τονιστεί ότι η ανταπόκριση ενός ασθενούς στη θεραπεία είναι ένα ιδιαίτερα πολύπλοκο ζήτημα το οποίο δύναται να επηρεάζεται από πολλαπλούς και ετερογενείς παράγοντες. Έτσι, η ανάπτυξη ενός ακριβούς και αποδοτικού μοντέλου πρόβλεψης που να βασίζεται στη γονιδιακή έκφραση είναι ιδιαίτερα δύσκολη.

Λαμβάνοντας υπόψη τα αποτελέσματα της ενοποιητικής μετα-ανάλυσης σχετικά με τα σημαντικά γονίδια – κόμβους που εντοπίστηκαν να εκφράζονται διαφορεικά στα ούρα ή στο πλάσμα αίματος των ασθενών με KOK και ταυτόχρονα να έχουν προγνωστική ή προβλεπτική αξία, προσδιορίστηκαν ορισμένα γονίδια – κόμβους – κλειδιά ως προτεινόμενοι βιοδείκτες σχετικά με τον KOK. Αυτά τα γονίδια περιλαμβάνουν τα *ANXA5*, *CDT1*, *COL3A1*, *SPPI*, *VEGFA*, *CDCA8*, *HJURP*, *TOP2A* και *COL6A1*.

Η αννεξίνη A5 (Annexin A5 – *ANXA5*) είναι ένας αναστολέας της πρωτεϊνικής κίνησης C και μία από τις δώδεκα αννεξίνες που έχουν ταυτοποιηθεί στον άνθρωπο (*ANXA1-11*, 13). Αποτελεί μία πρωτεΐνη με αντιπηκτικές ιδιότητες που αναστέλλει έμμεσα το σύμπλοκο της ειδικής θρομβοπλαστίνης, το οποίο συμμετέχει στον καταρράκτη της πήξης. Γενικά, οι αννεξίνες εμπλέκονται στην ομοιοστατική ρύθμιση της ενδοκυτταρικής συγκέντρωσης ιόντων ασβεστίου και διαδραματίζουν σημαντικό ρόλο στον κύκλο

ζωής των κυττάρων, την κυτταρική σηματοδότηση, τη φλεγμονή, την κυτταρική ανάπτυξη, τη διαφοροποίηση, την εξωκυττάρωση και την απόπτωση. Οι αννεξίνες βρίσκονται κανονικά στο εσωτερικό του ανθρώπινου κυττάρου. Ωστόσο, ορισμένες αννεξίνες (*ANXA1*, *ANXA2* και *ANXA5*) μπορούν να εκκρίνονται από το κυτταρόπλασμα σε εξωτερικά κυτταρικά περιβάλλοντα, όπως το αίμα. Στην παρούσα μελέτη, η έκφραση του γονιδίου *ANXA5* ήταν σημαντικά αυξημένη στο πλάσμα αίματος των ασθενών με ΚΟΚ, ενώ ήταν σημαντικά μειωμένη στον ιστό της ουροδόχου κύστης των ασθενών με ΚΟΚ. Αυτό οφείλεται στο γεγονός ότι το ενοποιημένο μετα-σύνολο δεδομένων, με τους ιστούς της ουροδόχου κύστης, περιλάμβανε περισσότερα δείγματα μη μυοδιηθητικού παρά μυοδιηθητικού ΚΟΚ και ότι το γονίδιο *ANXA5* έχει αποδειχθεί ότι υπο-εκφράζεται στα πρώιμα στάδια του ΚΟΚ, και ότι υπερ-εκφράζεται στα υψηλότερα στάδια [254]. Ως εκ τούτου, η εν λόγω πρωτεΐνη έχει προταθεί ως δείκτης της μετάβασης των όγκων από στάδιο χαμηλού σε υψηλού βαθμού για τον ΚΟΚ. Επίσης, το γονίδιο *ANXA5*, μαζί με τα υπόλοιπα μέλη της οικογένειας Annexin, έχει βρεθεί ότι εκφράζεται διαφορετικά και συνδέεται σε μεγάλο βαθμό με την πρόγνωση του ΚΟΚ [255]. Πιο συγκεκριμένα, η υψηλή έκφραση του *ANXA5* βρέθηκε να συσχετίζεται με χειρότερους χρόνους επιβίωσης ελεύθερης νόσου και χωρίς εξέλιξη νόσου, υποδεικνύοντας ότι μπορεί να εμπλέκεται στην υποτροπή και την εξέλιξη του ΚΟΚ. Σε άλλη μελέτη, επαληθεύτηκε η δυσμενής προγνωστική αξία του γονιδίου *ANXA5* και η υψηλή έκφρασή του συνδέθηκε με τον βασικό υπότυπο του μυοδιηθητικού ΚΟΚ [256]. Το γονίδιο *ANXA5* βρέθηκε επίσης να εκφράζεται διαφορετικά σε μια πληθώρα άλλων καρκίνων, όπως ο καρκίνος του μαστού [257], το ηπατοκυτταρικό καρκίνωμα [258] και το πλακώδες καρκίνωμα του πνεύμονα [259]. Έτσι, το γονίδιο *ANXA5* θεωρείται προγνωστικός βιοδείκτης για την ανάπτυξη, την εξέλιξη, την διεσόδωση και τη μετάσταση του όγκου και φαίνεται να κατέχει διαγνωστική, προγνωστική και θεραπευτική σημασία στον καρκίνο [260].

Ο παράγοντας *CDTI* αποτελεί έναν παράγοντα αδειοδότησης που λειτουργεί έτσι ώστε να περιορίζει την αντιγραφή του DNA πάνω από μία φορά ανά κυτταρικό κύκλο. Ειδικότερα, η πρωτεΐνη *CDTI* εμπλέκεται στο σχηματισμό του συμπλόκου πριν την αντιγραφή, το οποίο απαιτείται για την αντιγραφή του DNA. Η *CDTI* αναστέλλεται από τη Geminin, εμποδίζοντας τη συγκρότηση του συμπλόκου προ-αντιγραφής και την έναρξη της αντιγραφής σε ακατάλληλες θέσεις. Η πρωτεΐνη *CDTI* φωσφορυλιώνεται από κινάσες εξαρτώμενες από την κυκλίνη Α με αποτέλεσμα την αποικοδόμησή της. Ως εκ τούτου, η Chromatin licensing and DNA replication factor 1 (*CDTI*) συνδέεται σε μεγάλο βαθμό με τον κυτταρικό κύκλο, την κυτταρική διαίρεση, την αντιγραφή του DNA και τη μίτωση. Στην παρούσα μελέτη, το γονίδιο *CDTI* φαίνεται ότι ήταν υπερ-εκφρασμένο στον ιστό της ουροδόχου κύστης και στα ούρα των ασθενών με ΚΟΚ, γεγονός που επιβεβαιώθηκε επίσης από τις έρευνες των Χ. C. Μο κ.α. [119], ενώ ήταν σημαντικά υπο-εκφρασμένο στο πλάσμα αίματος των ασθενών με ΚΟΚ. Η πρωτεΐνη *CDTI* θεωρείται ότι συμβάλλει στον κυτταρικό πολλαπλασιασμό και στην αστάθεια του γονιδιώματος [261] και έχει βρεθεί ότι συχνά ρυθμίζεται εσφαλμένα στον καρκίνο [262]. Επιπλέον, όταν εκφράζεται σε υψηλά επίπεδα, έχει συσχετιστεί με κακή επιβίωση και πρόγνωση στον καρκίνο του μαστού [263], το ηπατοκυτταρικό καρκίνωμα [264], τον καρκίνο του παχέος εντέρου [265] και τον καρκίνο του προστάτη [266]. Ακόμη, η υπερ-έκφραση του γονιδίου *CDTI* συνδέεται με την ακανόνιστη κυτταρική αντιγραφή, την ενεργοποίηση των σημείων ελέγχου βλάβης του DNA και την προδιάθεση για κακοήγη πολλαπλασιασμό σε διάφορους ανθρώπινους καρκίνους. Η ανώμαλη έκφραση του *CDTI* στον ΚΟΚ και η συνακόλουθη διαγνωστική και προγνωστική σημασία του μένει να διευκρινιστεί περαιτέρω.

Το γονίδιο Collagen Type III Alpha 1 Chain (*COL3A1*) κωδικοποιεί τις προ-άλφα 1 αλυσίδες του κολλαγόνου τύπου III, μιας ινώδους πρωτεΐνης κολλαγόνου που εμφανίζεται στους περισσότερους μαλακούς συνδετικούς ιστούς, όπως οι αρτηρίες, το δέρμα και τα μαλακά όργανα, συχνά μαζί με το κολλαγόνο τύπου I. Είναι ένα βασικό γονίδιο που σχετίζεται με την εξωκυττάρια μήτρα, καθώς τα

μονομερή του διασταυρώνονται σε παχύτερα ινίδια, τα οποία συσσωρεύονται για να σχηματίσουν ίνες, παρέχοντας μια ισχυρή δομή στήριξης για τους ιστούς που απαιτούν αντοχή στον εφελκυσμό και παίζοντας ουσιαστικό ρόλο στην επεκτασιμότητά τους [267]. Τα επίπεδα του γονιδίου *COL3A1* έχει βρεθεί να είναι αξιοσημείωτα αυξημένα σε περιπτώσεις όγκων υψηλού βαθμού κακοήθειας, καθώς και στις περιπτώσεις του μυοδιηθητικού ΚΟΚ σε σύγκριση με περιπτώσεις όγκων χαμηλού βαθμού και μη μυοδιηθητικού ΚΟΚ. Η υψηλότερη αυτή έκφραση του γονιδίου έχει συνδεθεί με μικρότερη επιβίωση ελεύθερης νόσου [225], καθώς και με χειρότερη συνολική επιβίωση [268] των ασθενών με ΚΟΚ. Το γονίδιο *COL3A1* έχει βρεθεί επίσης ότι συγκαταλέγεται μεταξύ των γονιδίων – κόμβων που σχετίζονται με την εξέλιξη του μη μυοδιηθητικού σε μυοδιηθητικό ΚΟΚ [269]. Ειδικότερα, τα επίπεδα έκφρασης του *COL3A1* εντοπίστηκαν χαμηλότερα μεταξύ ασθενών με μη μυοδιηθητικό ΚΟΚ [270] και υψηλότερα μεταξύ ασθενών με διηθητική νόσο, γεγονός που καταδεικνύει ότι το γονίδιο συμβάλλει στην εξέλιξη των όγκων και τη μετάσταση [271, 272]. Αυτό συνάδει με τα αποτελέσματα της παρούσας ενοποιητικής μετα-ανάλυσης, στα οποία διαπιστώθηκε ότι το γονίδιο *COL3A1* υπο-εκφράζεται στον ιστό της ουροδόχου κύστης. Το γεγονός αυτό εξηγείται καθώς στο ενοποιημένο μετα-σύνολο δεδομένων ο αριθμός των δειγμάτων μη μυοδιηθητικού ΚΟΚ ήταν υψηλότερος σε σύγκριση με τα δείγματα μυοδιηθητικού ΚΟΚ. Σε αντίθεση με αυτό, το γονίδιο *COL3A1* βρέθηκε να υπο-εκφράζεται και στα δείγματα ούρων, παρά το γεγονός ότι η πλειονότητα των ασθενών αυτού του συνόλου δεδομένων είχε υψηλού βαθμού ουροθηλιακά καρκινώματα. Είναι προφανές πως απαιτούνται μελλοντικές μελέτες προκειμένου να διαλευκανθεί ο ρόλος του *COL3A1* στην ανάπτυξη του ΚΟΚ και στην εξέλιξη της νόσου από μη μυοδιηθητικό σε μυοδιηθητικό καρκίνωμα.

Το γονίδιο Collagen Type VI Alpha 1 Chain (*COL6A1*) κωδικοποιεί τις προ-άλφα 1 αλυσίδες του κολλαγόνου τύπου VI και ανήκει στην υπερ-οικογένεια των πρωτεϊνών κολλαγόνου, όπως και το *COL3A1*. Τα κολλαγόνα είναι πρωτεΐνες της εξωκυτταρικής μήτρας και διαδραματίζουν σημαντικό ρόλο στη διατήρηση της ακεραιότητας και της σταθερότητας διαφόρων ιστών. Συγκεκριμένα, το κολλαγόνο VI δρα ως πρωτεΐνη δέσμευσης κυττάρων και εμπλέκεται στη διαδικασία προσκόλλησης των κυττάρων. Οι αυξημένες εκφράσεις του *COL6A1* έχουν συσχετιστεί σε μεγάλο βαθμό με χειρότερη συνολική επιβίωση σε ασθενείς με ΚΟΚ [268]. Σε μια πρόσφατη βιοπληροφορική ανάλυση [273], το γονίδιο *COL6A1* βρέθηκε να αποτελεί δείκτη κινδύνου για την εξέλιξη του ΚΟΚ σε υψηλού βαθμού και να σχετίζεται δυσμενώς με την πρόγνωση των ασθενών. Στην ίδια μελέτη, αναφέρθηκε επίσης ότι μπορεί να χρησιμοποιηθεί ως μεμονωμένος αποτελεσματικός διαγνωστικός και προγνωστικός βιοδείκτης για τον ΚΟΚ, μαζί με πέντε άλλα μέλη της οικογένειας των κολλαγόνων. Στην παρούσα μελέτη, τα επίπεδα του *COL6A1* βρέθηκαν να υπο-εκφράζονται σημαντικά στο πλάσμα αίματος και στον ιστό της ουροδόχου κύστης ασθενών με ΚΟΚ, επιβεβαιώνοντας προηγούμενα ευρήματα στη βιβλιογραφία [274]. Στην τελευταία μελέτη, προτάθηκε ότι τα *COL6A1* και *COL6A2* μπορεί να δρουν ως πρότυπα κολλαγόνα κατασκευάζοντας ένα φυσικό φράγμα για την αναστολή της ανάπτυξης και της εισβολής του όγκου του ΚΟΚ. Σύμφωνα με μια άλλη μελέτη που εφάρμοσε συγκριτικό προφίλ πρωτεωμικής στα ούρα, από ασθενείς με καρκίνο του προστάτη, η πρωτεΐνη *COL6A1* φάνηκε να κατέχει ενεργό συμμετοχή στον καρκίνο του προστάτη [275]. Τα επίπεδα των κολλαγόνων στα ούρα και στο αίμα μπορεί να έχουν δυναμικά διαγνωστική και προγνωστική αξία για τον ΚΟΚ και θα πρέπει να διερευνηθούν περαιτέρω, ιδίως όσον αφορά την αλληλεπίδραση της εξωκυτταρικής μήτρας με τον όγκο. Το κολλαγόνο αποτελεί βασικό συστατικό του μικροπεριβάλλοντος του όγκου, καθώς συμμετέχει στην ίνωση σχετική με τον καρκίνο. Έτσι, η κατανόηση των δομικών ιδιοτήτων και των παθοφυσιολογικών λειτουργιών του στους ανθρώπινους καρκίνους μπορεί να οδηγήσει στην ανάπτυξη νέων αντικαρκινικών θεραπειών [276].

Η εκκρινόμενη φωσφοπρωτεΐνη 1 (Secreted Phosphoprotein 1 – *SPPI*), ή κοινώς γνωστή ως οστεοποντίνη (*OPN*), είναι μια σημαντική μη κολλαγονική δομική πρωτεΐνη της εξωκυτταρικής μήτρας

και οργανικό συστατικό των οστών. Η πρωτεΐνη *SPPI* συμμετέχει στην προσκόλληση των οστεοκλαστών στην ανοργανοποιημένη εξωκυτταρική οστική μήτρα. Εκτός από τους ευεργετικούς ρόλους της *SPPI* στην επούλωση των τραυμάτων και την ομοιόσταση των οστών, η *SPPI* θεωρείται ότι εμπλέκεται σε διάφορες παθοφυσιολογικές διεργασίες, συμπεριλαμβανομένης της εξέλιξης και της μετάστασης του καρκίνου, δρώντας ως βασικός μεσολαβητής της φλεγμονής που σχετίζεται με τον όγκο [277]. Σε προγενέστερες μελέτες, έχει διαπιστωθεί ότι υπερ-εκφράζεται σημαντικά μεταξύ των δειγμάτων όγκων της ουροδόχου κύστης συγκριτικά με τα υγιή δείγματα [278, 279] και ότι υποδεικνύει κακή πρόγνωση σε σχέση με το προχωρημένο στάδιο της νόσου [280]. Το γονίδιο *SPPI* αποδείχθηκε επίσης ότι υπερ-εκφράζεται σημαντικά στον ιστό της ουροδόχου κύστης και στον ορό αίματος ασθενών με μεταβατικό κυτταρικό καρκίνωμα της ουροδόχου κύστης [281] και σε ασθενείς με μυοδιθητικό ΚΟΚ σε σύγκριση με υγιή άτομα [186], ευρήματα τα οποία συμφωνούν με τα αποτελέσματα της παρούσας μελέτης. Έχει επίσης προταθεί ότι το γονίδιο *SPPI* μπορεί να αποτελέσει αποτελεσματικό θεραπευτικό ή διαγνωστικό στόχο σε ορισμένους καρκίνους, όπως το μελάνωμα, ο καρκίνος του μαστού, του παχέος εντέρου, του κεφαλιού και του λαιμού, αλλά και του πνεύμονα, καθώς αποδείχθηκε ότι συσχετίζεται με κακή κλινική έκβαση και ότι προάγει την εξέλιξη της νόσου, αλληλεπιδρώντας με ογκογονίδια και διευκολύνοντας τη διήθηση των ανοσοκυττάρων μέσα στους όγκους [282, 283]. Είναι σημαντικό να αναφερθεί ότι η έκφραση του *SPPI* έχει συσχετιστεί με την ανταπόκριση των ασθενών στη χημειοθεραπεία [180], γεγονός το οποίο επιβεβαιώθηκε και στην παρούσα ανάλυση. Πιο συγκεκριμένα, στην παρούσα έρευνα διαπιστώθηκε ότι οι ασθενείς που είχαν υψηλότερα επίπεδα έκφρασης του *SPPI* εμφάνισαν χαμηλότερη ανταπόκριση στη χημειοθεραπεία με βάση τη σισπλατίνη, γεγονός που υποστηρίζεται και από προηγούμενα ευρήματα της βιβλιογραφίας και έρευνες σε άλλους τύπους καρκίνου, όπως ο καρκίνος του πνεύμονα [284] και ο καρκίνος των ωοθηκών [285]. Τα αποτελέσματα μιας άλλης μελέτης διαπίστωσαν ότι το γονίδιο *SPPI* υπερ-εκφράζεται σε κύτταρα και ιστούς ουροθηλιακού καρκινώματος της ανώτερης ουροποιητικής οδού και τα υψηλά επίπεδα έκφρασης του *SPPI* στο πλάσμα αίματος συνδέονται στενά με υψηλότερο στάδιο και βαθμό κακοήθειας του όγκου [286]. Λαμβάνοντας υπόψη όλα τα παραπάνω, προτείνεται ότι τα επίπεδα του *SPPI* στην κυκλοφορία μπορεί να αποτελέσουν έναν πιθανό βιοδείκτη για τον εντοπισμό των ασθενών με ΚΟΚ και την πρόβλεψη της διηθητικής νόσου, αλλά και της ανταπόκρισης στη θεραπεία. Απαιτείται περαιτέρω έρευνα για τη διερεύνηση των ακριβών μοριακών μηχανισμών του στον ΚΟΚ και την αξιολόγηση της αξίας του ως βιοδείκτη.

Το γονίδιο του αγγειακού ενδοθηλιακού αυξητικού παράγοντα Α (Vascular endothelial growth factor A – *VEGFA*) είναι μέλος της οικογένειας των αγγειακών ενδοθηλιακών αυξητικών παραγόντων (VEGF) και του αυξητικού παράγοντα του πλακούντα (PGF), οι οποίοι διαδραματίζουν σημαντικό και συμπληρωματικό ρόλο στην αγγειογένεση. Το γονίδιο αυτό κωδικοποιεί μια πρωτεΐνη που δεσμεύει την ηπαρίνη, η οποία αποτελεί ένα δισουλφιδικό γλυκοζυλιωμένο ομοδιμερές. Επίσης, η πρωτεΐνη *VEGFA* προκαλεί τον πολλαπλασιασμό και τη μετανάστευση των αγγειακών ενδοθηλιακών κυττάρων, αποτελώντας βασικό ρυθμιστή τόσο της φυσιολογικής όσο και της παθολογικής αγγειογένεσης. Το γονίδιο *VEGFA* έχει από καιρό αναγνωριστεί ως δυνητικός αγγειακός και πολλαπλασιαστικός θεραπευτικός στόχος σε ασθενείς με καρκίνο και η μελέτη του έχει συνδράμει στην ανακάλυψη καινοτόμων θεραπευτικών προσεγγίσεων στην ογκολογία [287]. Ειδικότερα, το γονίδιο *VEGFA* υπερ-εκφράζεται σε πολλούς γνωστούς καρκίνους, συμπεριλαμβανομένου του ΚΟΚ, και η έκφρασή του έχει συσχετιστεί με το στάδιο και την εξέλιξη της νόσου, καθώς και με την πρόγνωση των ασθενών [288]. Τα επίπεδα του *VEGFA* εντοπίστηκαν να υπερ-εκφράζονται σε μεγάλο βαθμό στον ΚΟΚ [289], και το γονίδιο αυτό έχει επίσης αναγνωριστεί ως βασικό γονίδιο – κλειδί στον ΚΟΚ και ως γονίδιο που σχετίζεται με την πρόγνωση των ασθενών με ΚΟΚ [122], γεγονός που επιβεβαιώθηκε και από τα αποτελέσματα της παρούσας έρευνας. Παρόλο που ο προγνωστικός ρόλος του *VEGFA* στον ΚΟΚ

παραμένει αμφιλεγόμενος, οι περισσότερες μελέτες συγκλίνουν στο γεγονός ότι οι ασθενείς με υψηλότερα επίπεδα έκφρασης του γονιδίου *VEGFA* στους ιστούς ή στα ούρα εμφάνισαν χειρότερη έκβαση τόσο στη συνολική επιβίωση όσο και στην επιβίωση ελεύθερης νόσου [280, 290-292], γεγονός που δεν φαίνεται να επιβεβαιώνεται από τα αποτελέσματα της κοόρτης TCGA. Επίσης, σε μια μελέτη που διεξήχθη από τους Z. Zhong και M. Huang κ.α. προτάθηκε ότι το γονίδιο *MYLK*, το οποίο προσδιορίστηκε ως σημαντικό γονίδιο – κόμβος για τον KOK από την παρούσα μελέτη, μπορεί να λειτουργεί ως ανταγωνιστικό ενδογενές RNA που προάγει την εξέλιξη του KOK μέσω της διαμόρφωσης του σηματοδοτικού μονοπατιού *VEGFA/VEGFR2* [226]. Ακόμη, το γονίδιο *VEGFA* είχε προταθεί σε προηγούμενες μελέτες, μαζί με άλλους δείκτες που ανιχνεύθηκαν με ELISA, όπως τα γονίδια *IL8* και *MMP9*, ως βιοδείκτης των ούρων που μπορεί να ανιχνεύσει με ακρίβεια πρωτογενή ή υποτροπιάζοντα KOK [210, 293, 294]. Υπάρχουν λοιπόν αρκετά στοιχεία που υποδηλώνουν ότι τα επίπεδα του γονιδίου *VEGFA* στο πλάσμα αίματος έχουν ιδιαίτερη αξία και μπορούν να αποτελέσουν έναν δυνητικό διαγνωστικό και προγνωστικό βιοδείκτη για ασθενείς με KOK.

Η πρωτεΐνη Cell Division Cycle Associated 8 (*CDC48*) κωδικοποιεί το Borealin/Dasra B, το οποίο είναι ένα κρίσιμο πρωτεϊνικό συστατικό του συμπλόκου των χρωμοσωμικών επιβατών (chromosomal passenger complex – CPC), μιας σημαντικής δυναμικής δομής που λειτουργεί ως βασικός ρυθμιστής κατά τη διάρκεια της μίτωσης. Ειδικότερα, η πρωτεΐνη *CDC48* είναι απαραίτητη για τη διόρθωση των σφαλμάτων πρόσδεσης του κινητοχώρου (kinetochore) και για τη σταθερότητα της διπολικής ατράκτου στην ανθρώπινη μίτωση, ενώ, σε καταστάσεις ασθένειας, μπορεί να συμβάλει στην απομακρυσμένη μετάσταση των καρκινικών κυττάρων. Το γονίδιο *CDC48* αναφέρθηκε ως δυνητικός προγνωστικός βιοδείκτης για μια ποικιλία καρκίνων, συμπεριλαμβανομένου του καρκίνου του μαστού [295], του ήπατος [296], και του καρκίνου του προστάτη [297]. Σε μια πρόσφατη μελέτη του X. Gao κ.α. τα αποτελέσματα ανέδειξαν ότι το *CDC48* υπερ-εκφράζεται στον KOK σε αντίθεση με τους φυσιολογικούς ιστούς και η υψηλή έκφρασή του σχετίζεται σε μεγάλο βαθμό με τη δυσμενή πρόγνωση των ασθενών [298], ευρήματα που επισημάνθηκαν και από άλλους συγγραφείς [299]. Στην ίδια μελέτη, φάνηκε ότι μέσω της αναστολής της έκφρασης του *CDC48* αναστέλλονταν επίσης ο πολλαπλασιασμός, η μετανάστευση και η εισβολή των κυτταρικών σειρών KOK σε γειτονικούς ιστούς και επάχθηκε η απόπτωση των κυττάρων. Επιπλέον, ο S. Pan κ.α. διαπίστωσαν ότι το *CDC48*, μαζί με τα *KIF11*, *NCAPG* και *NEK2*, έπαιξε ουσιαστικό ρόλο στη διατήρηση των βλαστικών κυττάρων στον KOK [300]. Ο K. Chen κ.α. ανέφεραν ότι το *CDC48*, μαζί με τα *CENPF*, *AURKB*, *CCNB2*, *CDC20*, *TTK* και *ASPM*, θεωρήθηκαν ως γονίδια – κόμβοι για τον KOK και επαλήθευσαν την προγνωστική τους αξία [301]. Επιπλέον, σε προηγούμενη μελέτη έχει υπογραμμιστεί ότι το *CDC48* σε συνδυασμό με τα *MKI67*, *CENPA*, *AURKB*, *FOXMI* και *DLGAP5*, ήταν μεταξύ των κορυφαίων γονιδίων – κόμβων όσον αφορά τον KOK [302]. Επίσης, σε μια πρόσφατη βιοπληροφορική μελέτη [303], τα *CDC48* και *CDC20* αναγνωρίστηκαν ως υποψήφιοι διαγνωστικοί βιοδείκτες για τον KOK. Ο καθοριστικός τους ρόλος στον KOK υποστηρίχθηκε επίσης από τον S. Li κ.α., οι οποίοι πρότειναν ότι η υπο-έκφραση έκφραση των *CDC48*, *TOP2A*, *CENPF* και *FOXMI*, σχετιζόταν με ευνοϊκότερη συνολική επιβίωση των ασθενών με KOK [304]. Το *CDC48* προτάθηκε επίσης από τον J. Shi κ.α. ως υποψήφιο σημαντικό γονίδιο στον μη μυοδιηθητικό KOK [305]. Όλα τα προαναφερθέντα γονίδια υποδείχθηκαν ως σημαντικά γονίδια – κόμβοι για τον KOK από την παρούσα μελέτη και, έτσι, τα προηγούμενα ευρήματα στη βιβλιογραφία επιβεβαιώνουν τα αποτελέσματά αυτής της μελέτης. Στην παρούσα ανάλυση, τα υψηλά επίπεδα έκφρασης του *CDC48* βρέθηκαν να επιδεινώνουν σημαντικά τη συνολική επιβίωση των ασθενών με KOK (Πίνακας 6), αλλά να συσχετίζονται με καλύτερη συνολική επιβίωση των ασθενών με μυοδιηθητικό KOK που έλαβαν χημειοθεραπεία με βάση τη σισπλατίνη. Παρά το γεγονός ότι φαίνεται στο σημείο αυτό να υπάρχει μια ασυνέπεια, αξίζει να σημειωθεί ότι η ανταπόκριση στη χημειοθεραπεία συνδέεται με πλήθος άλλων παραμέτρων. Ο λόγος αυτής της ασυνέπειας θα μπορούσε να οφείλεται

στην παρουσία περισσότερων υποτύπων βασικοκυτταρικών / ακανθοκυτταρικών καρκινωμάτων (basal / squamous carcinomas) στο σύνολο δεδομένων GSE169455, οι οποίοι έχουν ενοχοποιηθεί από προγενέστερες μελέτες ότι καταστέλλουν την αποτελεσματικότητα της χημειοθεραπείας [306]. Ακόμη, θα μπορούσε να οφείλεται στη συσχέτιση του *CDC48* με τη διήθηση των ανοσοκυττάρων στον όγκο, γεγονός το οποίο θα μπορούσε να αποτελέσει προγνωστικό βιοδείκτη για την ανταπόκριση στη χημειοθεραπεία [307, 308]. Ως εκ τούτου, υπάρχουν πολλαπλές και ισχυρές ενδείξεις ότι το γονίδιο *CDC48* μπορεί να αποτελέσει αποτελεσματικό θεραπευτικό στόχο για την πρόγνωση και τη θεραπεία του ΚΟΚ, αλλά η ακριβής βιολογική του λειτουργία παραμένει ακόμη ασαφής και πρέπει να διαλευκανθεί.

Η πρωτεΐνη Holliday Junction Recognition (*HJURP*) είναι μια κεντρομερική συνοδός ιστόνης που απαιτείται για την πρόσληψη της παραλλαγής της πρωτεΐνης κεντρομερούς Α (*CENPA*), που μοιάζει με την ιστόνη H3, και την εναπόθεσή της στα κεντρομερή κατά την πρώιμη φάση G1. Πιο συγκεκριμένα, η πρωτεΐνη *HJURP* είναι ένας παράγοντας που ρυθμίζεται από τον κυτταρικό κύκλο και είναι υπεύθυνος για τη διατήρηση και την εναπόθεση της *CENPA* στα κεντρομερή των χρωμοσωμάτων [309]. Εκτός από τη συναρμολόγηση της χρωματίνης που περιέχει το *CENPA*, η πρωτεΐνη αυτή εμπλέκεται στη ρύθμιση της δραστηριότητας της πρόσδεσης του DNA, στον χρωμοσωμικό διαχωρισμό, στη μίτωση των κυττάρων και στη ρύθμιση της συναρμολόγησης των συμπλόκων που περιέχουν πρωτεΐνες. Η παρούσα ανάλυση επιβεβαίωσε προηγούμενες βιοπληροφορικές μελέτες στις οποίες το γονίδιο *HJURP* προτάθηκε ως γονίδιο – κλειδί για τον ΚΟΚ [118, 119]. Ο R. Cao κ.α. διαπίστωσαν ότι το *HJURP* υπερ-εκφράζεται σε μεγάλο βαθμό στους ιστούς προερχόμενους από ΚΟΚ, τόσο σε επίπεδο mRNA όσο και σε επίπεδο πρωτεΐνης, και πρότειναν ότι το *HJURP* μπορεί να ρυθμίζει τον κυτταρικό πολλαπλασιασμό και την απόπτωση στον ΚΟΚ, δρώντας στον αρνητικό βρόχο ανατροφοδότησης *PPARγ-SIRT1* [310]. Άλλες μελέτες ανέφεραν ότι τα επίπεδα του *HJURP* ήταν σημαντικά υψηλότερα στους καρκινικούς συγκριτικά με τους φυσιολογικούς ιστούς στον καρκίνο του παγκρέατος [311], του πνεύμονα [312], του μαστού [313], του προστάτη [314] και του νεφρού [315], και υπογράμμισαν ότι η υψηλή έκφρασή του ήταν άμεσα συνυφασμένη με κακή επιβίωση των ασθενών. Ο ογκογόνος ρόλος του γονιδίου διερευνήθηκε επίσης σε μια πρόσφατη παν-καρκινική ανάλυση από τον R. Su κ.α. επικυρώνοντας τα προαναφερθέντα ευρήματα [316]. Στην παρούσα μελέτη, το γονίδιο *HJURP* βρέθηκε υπερ-εκφρασμένο στους ιστούς του ΚΟΚ αλλά τα επίπεδα έκφρασής του αναφέρθηκαν χαμηλότερα στο πλάσμα αίματος των ασθενών με ΚΟΚ. Ο ρόλος του γονιδίου *HJURP* στην ανάπτυξη των όγκων, και ειδικά στον ΚΟΚ, είναι ακόμη ασαφής και μένει να διευκρινιστεί, καθώς φαίνεται ότι το *HJURP* θα μπορούσε να αποτελέσει έναν νέο διαγνωστικό και προγνωστικό βιοδείκτη για τη διαχείριση της νόσου του ΚΟΚ.

Η DNA τοποϊσομεράση ΙΙΑ (DNA Topoisomerase II Alpha – *TOP2A*) κωδικοποιεί ένα βασικό πυρηνικό ένζυμο που ρυθμίζει την κατάσταση του DNA κατά τη μεταγραφή, δημιουργεί σπασίματα μονής έλικας του DNA και επάγει τη μεταγραφή γονιδίων κατά τη διάρκεια της κυτταρικής διαίρεσης. Κατά συνέπεια, η πρωτεΐνη *TOP2A* εμπλέκεται στον σχηματισμό, τον εμπλουτισμό και τον διαχωρισμό των χρωμοσωμάτων, στην αντιγραφή και τη μεταγραφή του DNA και εικάζεται ότι εμπλέκεται στην ανάπτυξη διαφόρων τύπων καρκίνου. Το γονίδιο *TOP2A* αναφέρθηκε ότι είναι ένα από τα 10 κορυφαία γονίδια – κόμβους που εντοπίστηκαν στον ΚΟΚ, μαζί με τα γονίδια *VEGFA*, *CCNB1*, *CDC20*, *AURKB*, *UBE2C* και *CCNB2* [122]. Όλα τα προηγούμενα έχουν προσδιοριστεί κι αυτά ως γονίδια – κλειδιά για τον ΚΟΚ από την παρούσα ανάλυση. Επιπλέον, ο S. Zeng κ.α. διαπίστωσαν ότι το *TOP2A* υπερ-εκφράζεται σε μεγάλο βαθμό στον ΚΟΚ, ιδίως σε όγκους υψηλού βαθμού και προχωρημένου σταδίου, και η υπερ-έκφρασή του συνδέεται σε μεγάλο βαθμό με δυσμενέστερη ειδική της νόσου επιβίωση, επιβίωση χωρίς εξέλιξη της νόσου και επιβίωση χωρίς υποτροπή [317]. Επίσης στη μελέτη αυτή, αποδείχθηκε ότι ο πολλαπλασιασμός των κυττάρων ΚΟΚ αναστέλλεται ιδιαίτερα από την

εξουδετέρωση του γονιδίου *TOP2A* και η ικανότητα μετανάστευσης και εισβολής των καρκινικών κυττάρων καταστέλλεται σημαντικά. Στην ίδια κατεύθυνση, οι F. Zhang και H. Wu διαπίστωσαν ότι με την αναστολή του *TOP2A* καταστέλλεται η καρκινογένεση στον ΚΟΚ [318]. Στην παρούσα μελέτη, το γονίδιο *TOP2A* βρέθηκε να υπερ-εκφράζεται σημαντικά στα ούρα των ασθενών με ΚΟΚ σε σύγκριση με τους υγιείς ασθενείς. Τα αποτελέσματα αυτά είναι σε συμφωνία με τα ευρήματα του W. T. Kim κ.α. [319] και υποστηρίζουν τη θεωρία ότι τα νουκλεϊκά οξέα που κυκλοφορούν εκτός των κυττάρων στα ούρα των ασθενών μπορεί να αποτελούν αξιόλογους συμπληρωματικούς διαγνωστικούς βιοδείκτες για τον ΚΟΚ. Προγενέστερες μελέτες επιβεβαίωσαν ότι το γονίδιο *TOP2A* είναι υπερ-εκφρασμένο στα ούρα των ασθενών με ΚΟΚ συγκριτικά με τα ούρα των ασθενών ελέγχου, χρησιμοποιώντας την τεχνική Western blotting [320]. Ο G. Botti κ.α. εξέτασαν την αποτελεσματικότητα του τεστ ProEx C, το οποίο πρακτικά αποτελεί ένα ανοσοϊστοχημικό αντιδραστήριο που ενσωματώνει τα αντισώματα *TOP2A* και *MCM2*, ως βοηθητικό εργαλείο για την αξιολόγηση των ουροθηλιακών αλλοιώσεων στην κυτταρολογία ούρων και πιστοποίησαν ότι μπορεί να διαφοροποιήσει με ακρίβεια τις υψηλού βαθμού αλλοιώσεις από τις καλοήθειες και αντιδραστικές καταστάσεις [321]. Είναι αξιοσημείωτο ότι και το γονίδιο *MCM2* βρέθηκε να είναι σημαντικά υπο-εκφραζόμενο στο πλάσμα αίματος των ασθενών με ΚΟΚ στην παρούσα μελέτη (Εικόνα 48). Αξίζει να σημειωθεί ότι, σε άλλη μελέτη, ο συνδυασμός των *TOP2A* και *MCM2* αναφέρθηκε ως ο καλύτερος βιοδείκτης για τη διάκριση μεταξύ χαμηλού και υψηλού βαθμού πλακωδών ενδοεπιθηλιακών αλλοιώσεων για τον καρκίνο του τραχήλου της μήτρας [322]. Επίσης, στην παρούσα μελέτη τα επίπεδα έκφρασης του γονιδίου *TOP2A* βρέθηκαν να διαφέρουν σημαντικά μεταξύ των ασθενών που ανταποκρίθηκαν πλήρως στην προεγχειρητική χημειοθεραπεία και εκείνων που δεν ανταποκρίθηκαν. Προηγούμενες μελέτες έχουν αναφέρει ότι το γονίδιο *TOP2A* αποτελεί δείκτη πρόβλεψης της πρόγνωσης και της ανταπόκρισης σε διάφορες θεραπείες για τον καρκίνο, όπως για παράδειγμα στον καρκίνο του μαστού [323], στα σαρκώματα μαλακών μορίων [324], καθώς και στο διαυγοκυτταρικό καρκίνωμα του νεφρού [325]. Υπάρχουν λοιπόν ισχυρές ενδείξεις ότι το γονίδιο *TOP2A* διαδραματίζει λειτουργικό ρόλο στον πολλαπλασιασμό και την εισβολή του ΚΟΚ, καθώς και στην ανταπόκριση των ασθενών στη νόσο, λειτουργίες οι οποίες μένει να μελετηθούν κλινικά και να αποδειχθούν.

Τα εννέα γονίδια – κόμβοι – κλειδιά που εντοπίστηκαν και προτάθηκαν ως πιθανοί βιοδείκτες χρησιμοποιήθηκαν ως χαρακτηριστικά σε μοντέλα ταξινόμησης, με στόχο τη διάκριση των καρκινικών και των φυσιολογικών δειγμάτων. Τα μοντέλα αυτά έδειξαν πολύ υψηλή ακρίβεια πρόβλεψης για τη συντριπτική πλειονότητα των χρησιμοποιούμενων συνόλων δεδομένων, υποδεικνύοντας ότι αυτά τα γονίδια μπορούν να χρησιμοποιηθούν ως δυνητικοί διαγνωστικοί βιοδείκτες στον ΚΟΚ. Η πρωτεϊνική έκφραση αυτών των γονιδίων σε καρκινικούς και φυσιολογικούς ιστούς της ουροδόχου κύστης επιβεβαιώθηκε με την αξιοποίηση ανοσοϊστοχημικών εικόνων από τον Άτλαντα των Ανθρώπινων Πρωτεϊνών (HPA). Τα αποτελέσματα της ανάλυσης υπογραμμίζουν την ανάγκη επικύρωσης αυτών των πολλά υποσχόμενων βιοδεικτών για τον ΚΟΚ σε ανεξάρτητες προ-κλινικές δοκιμές.

Όπως είναι προφανές, υπήρχαν ορισμένοι περιορισμοί οι οποίοι είναι πιθανό να έχουν επηρεάσει τα αποτελέσματα της παρούσας μελέτης και θα πρέπει να αναφερθούν. Βασικό περιορισμό αποτελεί το γεγονός ότι τα 606 δείγματα ιστών ουροδόχου κύστης που μετα-αναλύθηκαν δεν προέρχονταν από τον αντίστοιχο αριθμό ασθενών. Σε πολλές μελέτες, τα καρκινικά δείγματα και τα δείγματα ελέγχου συλλέχθηκαν από τον ίδιο ασθενή, από καρκινικούς και παρακείμενους υγιείς ιστούς (αντιστοιχισμένα ζεύγη), γεγονός που θα μπορούσε να εισάγει ορισμένες δια-δειγματικές συσχετίσεις. Δεύτερον, τα αναπόφευκτα τεχνικά σφάλματα, όπως π.χ. τα φαινόμενα παρτίδας, επηρέασαν τα αποτελέσματα της παρούσας ανάλυσης, είτε όταν διορθώνονταν, είτε όχι. Έγινε προσπάθεια διαχείρισης των σφαλμάτων αυτών με τον πιο ενδεδειγμένο και συστηματικό τρόπο, τεκμηριώνοντας την ακολουθούμενη μεθοδολογία σε κάθε περίπτωση. Ένας άλλος σημαντικός περιορισμός ήταν το γεγονός ότι δεν

διεξήχθησαν πειράματα *in vivo* προκειμένου να επικυρωθούν οι πιθανές λειτουργίες και οι μηχανισμοί των γονιδίων που εντοπίστηκαν σχετικά με τη συμμετοχή τους στην ανάπτυξη και την εξέλιξη του ΚΟΚ. Επιπλέον, θα πρέπει να υπογραμμιστεί ότι ο ΚΟΚ αποτελεί μια εξαιρετικά ετερογενή κακοήθεια, η οποία αποτελείται από διάφορους μοριακούς υπότυπους. Η ταυτοποίηση των διαφορετικά εκφραζόμενων γονιδίων, καθώς και των γονιδίων με προγνωστική και προβλεπτική αξία μπορεί να διαφέρει σημαντικά ανάλογα με τον συγκεκριμένο υπότυπο που διερευνάται. Ωστόσο, στην παρούσα μελέτη, στόχος ήταν ο εντοπισμός των σημαντικών γονιδίων – κόμβων που διαφοροποιούν συνολικά τα κύτταρα του ΚΟΚ από τα φυσιολογικά κύτταρα. Τέλος, η παρούσα μελέτη ενσωμάτωσε δεδομένα γονιδιακής έκφρασης mRNA που προέρχονται αποκλειστικά από πειράματα μικροσυστοιχιών και επικύρωσε ορισμένα αποτελέσματα της βιβλιογραφίας αποκλειστικά με τη χρήση βιοπληροφορικών εργαλείων που ενσωματώνουν δεδομένα RNA-Seq και δεδομένα ανοσοϊστοχημείας. Μελλοντικά, σχεδιάζεται η διεξαγωγή περαιτέρω μελετών που περιλαμβάνουν ένα ευρύτερο φάσμα τύπων δεδομένων, όπως τα πρωτογενή δεδομένα RNA-Seq, τα δεδομένα μεθυλίωσης του DNA κ.α.

5. Συμπεράσματα

Ο καρκίνος της ουροδόχου κύστης (ΚΟΚ) είναι ένας από τους πιο διαδεδομένους καρκίνους παγκοσμίως, με σταθερή αύξηση της επίπτωσης και του επιπολασμού του, και συνοδεύεται από υψηλή νοσηρότητα και θνησιμότητα. Ο ΚΟΚ είναι μια σύνθετη νόσος η οποία εμφανίζει πολλαπλά μοριακά και παθολογικά προφίλ. Αυτό έχει ως αποτέλεσμα τη διαφορετική συμπεριφορά του ΚΟΚ ανάλογα με την κλινική σταδιοποίηση του όγκου και τον μοριακό του τύπο. Η διάγνωση και η παρακολούθηση του ΚΟΚ γίνεται κυρίως με επεμβατικές και δαπανηρές εξετάσεις, οι οποίες, αν και αξιόπιστες, είναι ιδιαίτερα δυσάρεστες για τον ασθενή και, προφανώς, δεν είναι απαλλαγμένες από παρενέργειες. Επί του παρόντος, δεν υπάρχει κάποιος καθιερωμένος προσυμπτωματικός έλεγχος για τη συγκεκριμένη νόσο και ούτε κάποια επίσημη ένδειξη για τη χρήση μοριακών βιοδεικτών στην κλινική πρακτική, παρόλο που υπάρχουν αρκετές διαθέσιμες εξετάσεις. Επομένως, υπάρχει επιτακτική ανάγκη για μια κλινική μη επεμβατική εξέταση (ή εξετάσεις) η οποία θα μπορεί να ανιχνεύει συγκεκριμένους βιοδείκτες και να προσφέρει την έγκαιρη ανίχνευση, την παρακολούθηση της νόσου και την ανταπόκριση του ασθενούς στη θεραπεία.

Τις τελευταίες δεκαετίες, η μελέτη του μεταγραφώματος (transcriptome) αναδείχθηκε ως μία από τις πιο υποσχόμενες τεχνολογίες στον τομέα της ογκολογίας, παρέχοντας προγνωστικά και προβλεπτικά οφέλη για τη διαχείριση του καρκίνου. Επίσης, η ανάπτυξη νέων τεχνολογιών, όπως η επαναστατική αλληλούχιση επόμενης γενιάς, επιτρέπει τον εντοπισμό καρκινικών βιοδεικτών, τον προσδιορισμό γονιδιακών υπογραφών, την ποσοτικοποίηση της γονιδιακής έκφρασης που επηρεάζει την ογκογένεση, καθώς και την ανακάλυψη μοριακών στόχων για αντικαρκινικές θεραπείες. Η μεταγραφομική (transcriptomics) έχει συμβάλει ριζικά στην καλύτερη ολιστική κατανόηση του καρκίνου, από τις ιστοπαθολογικές έως τις μοριακές ταξινομήσεις, προσφέροντας μια ελπιδοφόρα προοπτική για εξατομικευμένη διάγνωση και θεραπεία των όγκων. Η περαιτέρω πρόοδος στη διαμόρφωση του προφίλ του μεταγραφώματος μπορεί να επιτρέψει την τυποποίηση και, παράλληλα, τη μείωση του κόστους της ανάλυσής του, γεγονός που θα αποτελέσει το επόμενο βήμα ώστε η μεταγραφομική να καταστεί κανόνας της σύγχρονης ιατρικής του καρκίνου. Επίσης, η ενσωμάτωση της μεταγραφομικής με τη γονιδιωματική, την πρωτεομική, την επιγενετική και την ανοσολογική σκιαγράφηση των όγκων θα συνδράμει στην ανακάλυψη βιοδεικτών και στην κατανόηση των μηχανισμών της νόσου και, ως εκ τούτου, θα επιταχύνει την εφαρμογή της ογκολογίας ακριβείας.

Επιπρόσθετα, είναι πλέον σαφές πως οι σύγχρονες τεχνολογίες παραγωγής μεγάλου όγκου βιολογικών δεδομένων απαιτούν πολύ ισχυρά υπολογιστικά εργαλεία και αποδοτικούς αλγορίθμους για την επεξεργασία και ανάλυσή τους με σκοπό την εξαγωγή πληροφοριών. Η βιοπληροφορική είναι ένας από τους νεότερους διεπιστημονικούς τομείς της έρευνας και συμπεριλαμβάνει τη χρήση μαθηματικών, στατιστικών και υπολογιστικών μεθόδων για την επεξεργασία και ανάλυση των βιολογικών δεδομένων. Κατά την τελευταία δεκαετία, η ραγδαία ανάπτυξη στην παραγωγή δεδομένων και στις δυνατότητες της τεχνολογίας, τόσο στην εποχή της «γονιδιωματικής», όσο και στην εποχή των «omics», ήταν καταλυτική για τους επιστήμονες που ασχολούνται με την επεξεργασία των πειραματικών αποτελεσμάτων. Είναι σαφές πως οι παραδοσιακές γονιδιακές προσεγγίσεις στην έρευνα δεν επαρκούν για να ανταποκριθούν στην ανάπτυξη και τις απαιτήσεις της σύγχρονης έρευνας για την κατανόηση της πραγματικής βιολογίας. Οι τεράστιες ποσότητες δεδομένων που παράγονται από τις νέες τεχνολογίες, όπως η γονιδιωματική αλληλούχιση και τα τσιπ μικροσυστοιχιών, καθιστούν την διαχείριση των δεδομένων και την ενοποίηση πολλαπλών πλατφορμών υψίστης σημασίας. Η επιτυχημένη διαχείριση των δεδομένων θα πρέπει να καταλήγει στην ανάλυση και ερμηνεία τους με σκοπό την αναζήτηση

απαντήσεων στα ποικίλα βιολογικά ερωτήματα. Έτσι, η σύγχρονη έρευνα μετακινείται από τον περιορισμό σε ένα εργαστηριακό περιβάλλον στην εργασία με υπολογιστές, σε ένα περιβάλλον «εικονικού εργαστηρίου».

Στην παρούσα μελέτη, έγινε μια προσπάθεια αξιοποίησης του σημαντικού πλούτου των γονιδιακών, δημόσια διαθέσιμων δεδομένων, καθώς και των σύγχρονων υπολογιστικών εργαλείων με σκοπό την εξαγωγή νέας γνώσης για τον ΚΟΚ. Στόχος μας ήταν η ανεύρεση πιθανών διαγνωστικών, προγνωστικών και προβλεπτικών – ως προς την απόκριση στη θεραπεία – βιοδεικτών για τον ΚΟΚ, με την εφαρμογή βιοπληροφορικών τεχνικών και αλγορίθμων μηχανικής μάθησης. Στο πλαίσιο αυτό ενοποιήθηκαν με επιτυχία δεδομένα γονιδιακής έκφρασης από πολλαπλά σύνολα δεδομένων και, μέσω της χρήσης βιοπληροφορικών εργαλείων, εντοπίστηκε μία λίστα γονιδίων – κόμβων που φαίνεται να διαδραματίζουν ουσιαστικό ρόλο στην ανάπτυξη και την εξέλιξη του ΚΟΚ [326]. Ένα υποσύνολο αυτών των γονιδίων, και συγκεκριμένα τα *ANXA5*, *CDT1*, *COL3A1*, *SPPI*, *VEGFA*, *CDCA8*, *HJURP*, *TOP2A* και *COL6A1*, συσχετίστηκε με διαφορετική γονιδιακή έκφραση στα ούρα ή στο πλάσμα αίματος των ασθενών και αναδείχθηκε ως μια ομάδα δυνητικά διαγνωστικών βιοδεικτών για τον ΚΟΚ. Επιπλέον, στη μελέτη προσδιορίστηκε μια υπογραφή τριών γονιδίων (*COL3A1*, *FOXMI* και *PLK4*) που πέτυχε υψηλή προγνωστική απόδοση όσον αφορά τη συνολική επιβίωση των ασθενών με ΚΟΚ, αλλά και μια υπογραφή έξι γονιδίων (*ANXA5*, *CD44*, *NCAM1*, *SPPI*, *CDCA8* και *KIF14*) η οποία παρουσίασε αρκετά καλή προβλεπτική απόδοση όσον αφορά την ελεύθερη νόσου επιβίωση των ασθενών με μυοδιηθητικό ΚΟΚ οι οποίοι έλαβαν προεγχειρητική χημειοθεραπεία με βάση τη σισπλατίνη. Συνολικά, σε όλη την έκταση της παρούσας μελέτης, αναλύθηκαν 1.567 δείγματα (1.346 δείγματα ΚΟΚ και 268 υγιή δείγματα) από πειράματα μικροσυστοιχιών, αριθμός ιδιαίτερα μεγάλος για τον τύπο της έρευνας συγκριτικά με παρόμοιες μελέτες στην βιβλιογραφία. Η παρούσα μελέτη φιλοδοξεί να συμβάλει στη διαλεύκανση των γενετικών αλλαγών που συμβαίνουν στον ΚΟΚ, χρησιμοποιώντας συστηματικές βιοπληροφορικές μεθόδους. Είναι προφανές πως απαιτούνται περαιτέρω έρευνες για την επικύρωση της κλινικής αξίας αυτών των βιοδεικτών και της προοπτικής τους στη θεραπεία του ΚΟΚ.

6. Βιβλιογραφικές Αναφορές

- [1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, pp. 209-249, 2021.
- [2] M. C. Hulvat, "Cancer Incidence and Trends," *Surgical Clinics*, vol. 100, pp. 469-481, 2020.
- [3] C. de Martel, D. Georges, F. Bray, J. Ferlay, and G. M. Clifford, "Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis," *The Lancet Global Health*, vol. 8, pp. e180-e190, 2020.
- [4] K. B. Tran, J. J. Lang, K. Compton, R. Xu, A. R. Acheson, H. J. Henrikson, *et al.*, "The global burden of cancer attributable to risk factors, 2010-19: a systematic analysis for the Global Burden of Disease Study 2019," *The Lancet*, vol. 400, pp. 563-591, 2022.
- [5] J. Ferlay, M. Colombet, I. Soerjomataram, D. M. Parkin, M. Piñeros, A. Znaor, *et al.*, "Cancer statistics for the year 2020: An overview," *International Journal of Cancer*, vol. 149, pp. 778-789, 2021.
- [6] S. M. Gapstur, J. M. Drope, E. J. Jacobs, L. R. Teras, M. L. McCullough, C. E. Douglas, *et al.*, "A blueprint for the primary prevention of cancer: Targeting established, modifiable risk factors," *CA: A Cancer Journal for Clinicians*, vol. 68, pp. 446-470, 2018.
- [7] N. Nargis, A. K. M. G. Hussain, S. Asare, Z. Xue, A. Majmundar, P. Bandi, *et al.*, "Economic loss attributable to cigarette smoking in the USA: an economic modelling study," *The Lancet Public Health*, vol. 7, pp. e834-e843, 2022.
- [8] A. Hruby and F. B. Hu, "The Epidemiology of Obesity: A Big Picture," *Pharmacoeconomics*, vol. 33, pp. 673-689, 2015/07/01 2015.
- [9] A. De Lorenzo, L. Romano, L. Di Renzo, N. Di Lorenzo, G. Cennamo, and P. Gualtieri, "Obesity: A preventable, treatable, but relapsing disease," *Nutrition*, vol. 71, p. 110615, 2020/03/01/ 2020.
- [10] A. Hruby, J. E. Manson, L. Qi, V. S. Malik, E. B. Rimm, Q. Sun, *et al.*, "Determinants and Consequences of Obesity," *American Journal of Public Health*, vol. 106, pp. 1656-1662, 2016.
- [11] L. The, "The link between cancer and obesity," *The Lancet*, vol. 390, p. 1716, 2017.
- [12] S. Wehrauch-Blüher, P. Schwarz, and J.-H. Klusmann, "Childhood obesity: increased risk for cardiometabolic disease and cancer in adulthood," *Metabolism - Clinical and Experimental*, vol. 92, pp. 147-152, 2019.
- [13] T. Fiolet, B. Srour, L. Sellem, E. Kesse-Guyot, B. Allès, C. Méjean, *et al.*, "Consumption of ultra-processed foods and cancer risk: results from NutriNet-Santé prospective cohort," *BMJ*, vol. 360, p. k322, 2018.
- [14] N. Veronese, M. Solmi, M. G. Caruso, G. Giannelli, A. R. Osella, E. Evangelou, *et al.*, "Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses," *The American Journal of Clinical Nutrition*, vol. 107, pp. 436-444, 2018.
- [15] K. E. Bradbury, N. Murphy, and T. J. Key, "Diet and colorectal cancer in UK Biobank: a prospective study," *International Journal of Epidemiology*, vol. 49, pp. 246-258, 2019.
- [16] K. L. Piercy, R. P. Troiano, R. M. Ballard, S. A. Carlson, J. E. Fulton, D. A. Galuska, *et al.*, "The Physical Activity Guidelines for Americans," *JAMA*, vol. 320, pp. 2020-2028, 2018.

- [17] A. MCTIERNAN, C. M. FRIEDENREICH, P. T. KATZMARZYK, K. E. POWELL, R. MACKO, D. BUCHNER, *et al.*, "Physical Activity in Cancer Prevention and Survival: A Systematic Review," *Medicine & Science in Sports & Exercise*, vol. 51, pp. 1252-1261, 2019.
- [18] S. T. Orange, K. M. Hicks, and J. M. Saxton, "Effectiveness of diet and physical activity interventions amongst adults attending colorectal and breast cancer screening: a systematic review and meta-analysis," *Cancer Causes & Control*, vol. 32, pp. 13-26, 2021/01/01 2021.
- [19] S. Mahabir, "Association Between Diet During Preadolescence and Adolescence and Risk for Breast Cancer During Adulthood," *Journal of Adolescent Health*, vol. 52, pp. S30-S35, 2013/05/01/ 2013.
- [20] J. Pader, Y. Ruan, A. E. Poirier, K. Asakawa, C. Lu, S. Memon, *et al.*, "Estimates of future cancer mortality attributable to modifiable risk factors in Canada," *Canadian Journal of Public Health*, vol. 112, pp. 1069-1082, 2021/12/01 2021.
- [21] A. E. Poirier, Y. Ruan, K. D. Volesky, W. D. King, D. E. O'Sullivan, P. Gogna, *et al.*, "The current and future burden of cancer attributable to modifiable risk factors in Canada: Summary of results," *Preventive Medicine*, vol. 122, pp. 140-147, 2019/05/01/ 2019.
- [22] J. C. Pardo, V. Ruiz de Porras, A. Plaja, C. Carrato, O. Etxaniz, O. Buisan, *et al.*, "Moving towards Personalized Medicine in Muscle-Invasive Bladder Cancer: Where Are We Now and Where Are We Going?," *International Journal of Molecular Sciences*, vol. 21, 2020 2020.
- [23] J. Dobruch and M. Oszczudłowski, "Bladder Cancer: Current Challenges and Future Directions," *Medicina*, vol. 57, p. 749, 2021/07/24 2021.
- [24] International Agency for Research on Cancer (IARC) | Global Cancer Observatory (GCO), "Cancer Today," 2020.
- [25] American Cancer Society, "Cancer Facts & Figures 2022," American Cancer Society, Atlanta2022.
- [26] A. T. Lenis, P. M. Lec, and K. Chamie, "Bladder Cancer: A Review," *JAMA*, vol. 324, pp. 1980-1991, 2020/11/17 2020.
- [27] C. T. Lutz, L. Livas, S. R. Presnell, M. Sexton, and P. Wang, "Gender Differences in Urothelial Bladder Cancer: Effects of Natural Killer Lymphocyte Immunity," *Journal of Clinical Medicine*, vol. 10, p. 5163, 2021/11/04 2021.
- [28] Cancer Research UK. *Bladder cancer incidence statistics*. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/incidence>
- [29] A. Richters, K. K. H. Aben, and L. A. L. M. Kiemeny, "The global burden of urinary bladder cancer: an update," *World Journal of Urology*, vol. 38, pp. 1895-1904, 2020/08/01 2020.
- [30] S. L. Washington III, S. E. Gregorich, M. V. Meng, A. M. Suskind, and S. P. Porten, "Race modifies survival benefit of guideline-based treatment: Implications for reducing disparities in muscle invasive bladder cancer," *Cancer Medicine*, vol. 9, pp. 8310-8317, 2020.
- [31] K. Saginala, A. Barsouk, J. S. Aluru, P. Rawla, S. A. Padala, and A. Barsouk, "Epidemiology of Bladder Cancer," *Medical Sciences*, vol. 8, p. 15, 2020.
- [32] International Agency for Research on Cancer (IARC) | Global Cancer Observatory (GCO), "Cancer Tomorrow," 2020.
- [33] C. Samanic, M. Kogevinas, M. Dosemeci, N. r. Malats, F. X. Real, M. Garcia-Closas, *et al.*, "Smoking and Bladder Cancer in Spain: Effects of Tobacco Type, Timing, Environmental Tobacco Smoke, and Gender," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 15, pp. 1348-1354, 2006.

- [34] F. Bravi, M.-E. Spei, J. Polesel, M. Di Maso, M. Montella, M. Ferraroni, *et al.*, "Mediterranean Diet and Bladder Cancer Risk in Italy," *Nutrients*, vol. 10, p. 1061, 2018.
- [35] M. G. K. Cumberbatch, I. Jubber, P. C. Black, F. Esperto, J. D. Figueroa, A. M. Kamat, *et al.*, "Epidemiology of Bladder Cancer: A Systematic Review and Contemporary Update of Risk Factors in 2018," *European Urology*, vol. 74, pp. 784-795, 2018/12/01/ 2018.
- [36] E. P. Christoforidou, E. Riza, S. N. Kales, K. Hadjistavrou, M. Stoltidi, A. N. Kastania, *et al.*, "Bladder cancer and arsenic through drinking water: A systematic review of epidemiologic evidence," *Journal of Environmental Science and Health, Part A*, vol. 48, pp. 1764-1775, 2013/12/06 2013.
- [37] C. Ciocan, A. Godono, N. Franco, C. La Vecchia, E. Negri, P. Boffetta, *et al.*, "Mortality from bladder cancer in dyestuff workers exposed to aromatic amines: A 73-year follow-up," *La Medicina del Lavoro | Work, Environment and Health*, vol. 113, p. e2022017, 04/26 2022.
- [38] N. Albarakati, D. Khayyat, A. Dallol, J. Al-Maghrabi, and T. Nedjadi, "The prognostic impact of GSTM1/GSTP1 genetic variants in bladder Cancer," *BMC Cancer*, vol. 19, p. 991, 2019/10/23 2019.
- [39] Z. Guo, X. Niu, G. Fu, B. Yang, G. Chen, and S. a. Sun, "SLC14A1 (UT-B) gene rearrangement in urothelial carcinoma of the bladder: a case report," *Diagnostic Pathology*, vol. 15, p. 94, 2020/07/23 2020.
- [40] S. Cheng, A. S. Andrew, P. C. Andrews, and J. H. Moore, "Complex systems analysis of bladder cancer susceptibility reveals a role for decarboxylase activity in two genome-wide association studies," *BioData Mining*, vol. 9, p. 40, 2016/12/12 2016.
- [41] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology E-Book*: Elsevier Health Sciences, 2015.
- [42] O. Sanli, J. Dobruch, M. A. Knowles, M. Burger, M. Alemozaffar, M. E. Nielsen, *et al.*, "Bladder cancer," *Nature Reviews Disease Primers*, vol. 3, p. 17022, 2017/04/13 2017.
- [43] M. Minoli, M. Kiener, G. N. Thalmann, D. M. K. Julio, and R. Seiler, "Evolution of urothelial bladder cancer in the context of molecular classifications," *International Journal of Molecular Sciences*, vol. 21, pp. 1-25, 2020 2020.
- [44] Surveillance Epidemiology and End Results (SEER) Program. (2022). *Cancer Stat Facts: Bladder Cancer*. Available: <https://seer.cancer.gov/statfacts/html/urinb.html>
- [45] A. N. Gonzalez, M. J. Lipsky, G. Li, M. P. Rutman, K. L. Cooper, D. M. Weiner, *et al.*, "The Prevalence of Bladder Cancer During Cystoscopy for Asymptomatic Microscopic Hematuria," *Urology*, vol. 126, pp. 34-38, 2019.
- [46] G. Wang and J. K. McKenney, "Urinary Bladder Pathology: World Health Organization Classification and American Joint Committee on Cancer Staging Update," *Archives of Pathology & Laboratory Medicine*, vol. 143, pp. 571-577, 2018.
- [47] A. V. D. MEIJDEN, R. SYLVESTER, L. COLLETTE, A. BONO, and F. T. KATE, "The role and impact of pathology review on stage and grade assessment of stages Ta and T1 bladder tumors: a combined analysis of 5 European Organization for Research and Treatment of Cancer Trials," *Journal of Urology*, vol. 164, pp. 1533-1537, 2000.
- [48] S. F. Shariat, C. Savage, T. F. Chromecki, M. Sun, D. S. Scherr, R. K. Lee, *et al.*, "Assessing the clinical benefit of nuclear matrix protein 22 in the surveillance of patients with nonmuscle-invasive bladder cancer and negative cytology," *Cancer*, vol. 117, pp. 2892-2897, 2011.
- [49] National Cancer Institute at the National Institutes of Health, "Bladder and Other Urothelial Cancers Screening (PDQ®): Patient Version," in *PDQ Cancer Information Summaries*, ed, 2021.

- [50] C. D. Fankhauser, M. S. Wettstein, L. Afferi, N. C. Grossmann, and H. Mostafid, "En Bloc Resection of Bladder Tumor—Is It the Way Forward?," *Frontiers in Surgery*, vol. 8, 2021-May-31 2021.
- [51] B. Calò, M. Chirico, F. Fortunato, F. Sanguedolce, E. Carvalho-Dias, R. Autorino, *et al.*, "Is Repeat Transurethral Resection Always Needed in High-Grade T1 Bladder Cancer?," *Frontiers in Oncology*, vol. 9, 2019-June-04 2019.
- [52] J. Han, X. Gu, Y. Li, and Q. Wu, "Mechanisms of BCG in the treatment of bladder cancer-current understanding and the prospect," *Biomedicine & Pharmacotherapy*, vol. 129, p. 110393, 2020/09/01/ 2020.
- [53] C. C. Peyton, J. Chipollini, M. Azizi, A. M. Kamat, S. M. Gilbert, and P. E. Spiess, "Updates on the use of intravesical therapies for non-muscle invasive bladder cancer: how, when and what," *World Journal of Urology*, vol. 37, pp. 2017-2029, 2019/10/01 2019.
- [54] L. Stroman, B. Tschobotko, H. Abboudi, D. Ellis, E. Mensah, H. Kaneshayogan, *et al.*, "Improving Compliance With a Single Post-Operative Dose of Intravesical Chemotherapy After Transurethral Resection of Bladder Tumour," vol. 8, p. e29967, 2016.
- [55] "Robotic Partial Cystectomy for Bladder Cancer: A Single-Institutional Pilot Study," *Journal of Endourology*, vol. 24, pp. 223-227, 2010.
- [56] B. H. Bochner, G. Dalbagni, D. D. Sjoberg, J. Silberstein, G. E. Keren Paz, S. M. Donat, *et al.*, "Comparing Open Radical Cystectomy and Robot-assisted Laparoscopic Radical Cystectomy: A Randomized Clinical Trial," *European Urology*, vol. 67, pp. 1042-1050, 2015/06/01/ 2015.
- [57] D. J. Parekh, I. M. Reis, E. P. Castle, M. L. Gonzalgo, M. E. Woods, R. S. Svatek, *et al.*, "Robot-assisted radical cystectomy versus open radical cystectomy in patients with bladder cancer (RAZOR): an open-label, randomised, phase 3, non-inferiority trial," *The Lancet*, vol. 391, pp. 2525-2536, 2018.
- [58] T. N. Clinton, C. Huang, and A. C. Goh, "Is there an oncological benefit to extended lymphadenectomy for muscle-invasive bladder cancer?," *Translational Andrology and Urology*, vol. 9, pp. 2956-2964, 2020.
- [59] R. K. Lee, H. Abol-Enein, W. Artibani, B. Bochner, G. Dalbagni, S. Daneshmand, *et al.*, "Urinary diversion after radical cystectomy for bladder cancer: options, patient selection, and outcomes," *BJU International*, vol. 113, pp. 11-23, 2014.
- [60] Y. Cerantola, M. Valerio, B. Persson, P. Jichlinski, O. Ljungqvist, M. Hubner, *et al.*, "Guidelines for perioperative care after radical cystectomy for bladder cancer: Enhanced Recovery After Surgery (ERAS^{–}) society recommendations," *Clinical Nutrition*, vol. 32, pp. 879-887, 2013.
- [61] F. Abdollah, G. Gandaglia, R. Thuret, J. Schmitges, Z. Tian, C. Jeldres, *et al.*, "Incidence, survival and mortality rates of stage-specific bladder cancer in United States: A trend analysis," *Cancer Epidemiology*, vol. 37, pp. 219-225, 2013/06/01/ 2013.
- [62] P.-U. Malmström, T. Gårdmark, A. Sherif, V. Ströck, A. Hosseini-Aliabad, S. Jahnson, *et al.*, "Incidence, survival and mortality trends of bladder cancer in Sweden 1997–2016," *Scandinavian Journal of Urology*, vol. 53, pp. 193-199, 2019/07/04 2019.
- [63] C. L. Vale, "Neoadjuvant Chemotherapy in Invasive Bladder Cancer: Update of a Systematic Review and Meta-Analysis of Individual Patient Data: Advanced Bladder Cancer (ABC) Meta-analysis Collaboration," *European Urology*, vol. 48, pp. 202-206, 2005/08/01/ 2005.
- [64] C. Pfister, G. Gravis, A. Fléchon, M. Soulié, L. Guy, B. Laguerre, *et al.*, "Randomized Phase III Trial of Dose-dense Methotrexate, Vinblastine, Doxorubicin, and Cisplatin, or Gemcitabine and Cisplatin as Perioperative Chemotherapy for Patients with Muscle-invasive Bladder Cancer.

- Analysis of the GETUG/AFU V05 VESPER Trial Secondary Endpoints: Chemotherapy Toxicity and Pathological Responses," *European Urology*, vol. 79, pp. 214-221, 2021/02/01/ 2021.
- [65] C. N. Sternberg, I. Skoneczna, J. M. Kerst, P. Albers, S. D. Fossa, M. Agerbaek, *et al.*, "Immediate versus deferred chemotherapy after radical cystectomy in patients with pT3–pT4 or N+ M0 urothelial carcinoma of the bladder (EORTC 30994): an intergroup, open-label, randomised phase 3 trial," *The Lancet Oncology*, vol. 16, pp. 76-86, 2015.
 - [66] J. J. Leow, W. Martin-Doyle, P. S. Rajagopal, C. G. Patel, E. M. Anderson, A. T. Rothman, *et al.*, "Adjuvant Chemotherapy for Invasive Bladder Cancer: A 2013 Updated Systematic Review and Meta-Analysis of Randomized Trials," *European Urology*, vol. 66, pp. 42-54, 2014/07/01/ 2014.
 - [67] S. Burdett, D. J. Fisher, C. L. Vale, C. N. Sternberg, N. W. Clarke, M. K. B. Parmar, *et al.*, "Adjuvant Chemotherapy for Muscle-invasive Bladder Cancer: A Systematic Review and Meta-analysis of Individual Participant Data from Randomised Controlled Trials," *European Urology*, vol. 81, pp. 50-61, 2022/01/01/ 2022.
 - [68] E. Solsona, I. Iborra, A. Collado, J. Rubio-Briones, J. Casanova, and A. Calatrava, "Feasibility of Radical Transurethral Resection as Monotherapy for Selected Patients With Muscle Invasive Bladder Cancer," *Journal of Urology*, vol. 184, pp. 475-481, 2010.
 - [69] J. J. Knodler, S. A. Boorjian, S. P. Kim, C. J. Weight, P. Thapa, R. F. Tarrell, *et al.*, "Does Partial Cystectomy Compromise Oncologic Outcomes for Patients with Bladder Cancer Compared to Radical Cystectomy? A Matched Case-Control Analysis," *Journal of Urology*, vol. 188, pp. 1115-1119, 2012.
 - [70] J. Mathes, S. Rausch, T. Todenhöfer, and A. Stenzl, "Trimodal therapy for muscle-invasive bladder cancer," *Expert Review of Anticancer Therapy*, vol. 18, pp. 1219-1229, 2018/12/02 2018.
 - [71] J. A. Witjes, H. M. Bruins, R. Cathomas, E. M. Compérat, N. C. Cowan, G. Gakis, *et al.*, "European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines," *European Urology*, vol. 79, pp. 82-104, 2021/01/01/ 2021.
 - [72] L. P. Rhea, S. Mendez-Marti, D. Kim, and J. B. Aragon-Ching, "Role of immunotherapy in bladder cancer," *Cancer Treatment and Research Communications*, vol. 26, p. 100296, 2021/01/01/ 2021.
 - [73] B. A. Inman, T. A. Longo, S. Ramalingam, and M. R. Harrison, "Atezolizumab: A PD-L1–Blocking Antibody for Bladder Cancer," *Clinical Cancer Research*, vol. 23, pp. 1886-1890, 2017.
 - [74] J. W. F. Catto, A. Downing, S. Mason, P. Wright, K. Absolom, S. Bottomley, *et al.*, "Quality of Life After Bladder Cancer: A Cross-sectional Survey of Patient-reported Outcomes," *European Urology*, vol. 79, pp. 621-632, 2021/05/01/ 2021.
 - [75] R. E. Hautmann, H. Abol-Enein, T. Davidsson, S. Gudjonsson, S. H. Hautmann, H. V. Holm, *et al.*, "ICUD-EAU International Consultation on Bladder Cancer 2012: Urinary Diversion," *European Urology*, vol. 63, pp. 67-80, 2013/01/01/ 2013.
 - [76] E. Rammant, K. Decaestecker, R. Bultijnck, N. Sundahl, P. Ost, N. S. Pauwels, *et al.*, "A systematic review of exercise and psychosocial rehabilitation interventions to improve health-related outcomes in patients with bladder cancer undergoing radical cystectomy," *Clinical Rehabilitation*, vol. 32, pp. 594-606, 2018.

- [77] B. J. Danna, M. J. Metcalfe, E. L. Wood, and J. B. Shah, "Assessing Symptom Burden in Bladder Cancer: An Overview of Bladder Cancer Specific Health-Related Quality of Life Instruments," *Bladder Cancer*, vol. 2, pp. 329-340, 2016.
- [78] B. T. Jensen, "Organization Factors in the ERAS Bladder Cancer Pathway: The Multifarious Role of the ERAS Nurse, Why and What Is Important?," *Seminars in Oncology Nursing*, vol. 37, p. 151106, 2021/02/01/ 2021.
- [79] S. Siracusano, A. Zaka, P. Bassi, P. Gontero, E. Mearini, C. Imbimbo, *et al.*, "Quality-of-Life Outcomes in Female Patients With Ileal Conduit or Orthotopic Neobladder Urinary Diversion: 6-Month Results of a Multicenter Prospective Study," *Frontiers in Oncology*, vol. 12, 2022-April-12 2022.
- [80] S. Iivanainen, T. Alanko, K. Peltola, T. Konkola, J. Ekström, H. Virtanen, *et al.*, "ePROs in the follow-up of cancer patients treated with immune checkpoint inhibitors: a retrospective study," *Journal of Cancer Research and Clinical Oncology*, vol. 145, pp. 765-774, 2019/03/01 2019.
- [81] T. Kondo and K. Tanabe, "Role of lymphadenectomy in the management of urothelial carcinoma of the bladder and the upper urinary tract," *International Journal of Urology*, vol. 19, pp. 710-721, 2012.
- [82] P. Mandel, D. Tilki, and G. D. Eslick, "Extent of lymph node dissection and recurrence-free survival after radical cystectomy: A meta-analysis," *Urologic Oncology: Seminars and Original Investigations*, vol. 32, pp. 1184-1190, 2014/11/01/ 2014.
- [83] H. M. Bruins, E. Veskimäe, V. Hernandez, M. Imamura, M. M. Neuberger, P. Dahm, *et al.*, "The Impact of the Extent of Lymphadenectomy on Oncologic Outcomes in Patients Undergoing Radical Cystectomy for Bladder Cancer: A Systematic Review," *European Urology*, vol. 66, pp. 1065-1077, 2014/12/01/ 2014.
- [84] J. Nix, A. Smith, R. Kurpad, M. E. Nielsen, E. M. Wallen, and R. S. Pruthi, "Prospective Randomized Controlled Trial of Robotic versus Open Radical Cystectomy for Bladder Cancer: Perioperative and Pathologic Results," *European Urology*, vol. 57, pp. 196-201, 2010/02/01/ 2010.
- [85] J. Messer and D. J. Parekh, "Perioperative outcomes and Oncologic Efficacy from a Pilot Prospective Randomized Clinical Trial of Open Versus Robotic Assisted Radical Cystectomy," *Journal of Urology*, vol. 0, 2013.
- [86] N. E. Annels, M. Arif, G. R. Simpson, M. Denyer, C. Moller-Levet, D. Mansfield, *et al.*, "Oncolytic Immunotherapy for Bladder Cancer Using Coxsackie A21 Virus," *Molecular Therapy - Oncolytics*, vol. 9, pp. 1-12, 2018.
- [87] P. Sidaway, "IFN α -2b gene transfer: a new approach for BCG-resistant disease," *Nature Reviews Urology*, vol. 14, pp. 583-583, 2017/10/01 2017.
- [88] C. P. N. Dinney, M. B. Fisher, N. Navai, M. A. O'Donnell, D. Cutler, A. Abraham, *et al.*, "Phase I Trial of Intravesical Recombinant Adenovirus Mediated Interferon- β Formulated in Syn3 for Bacillus Calmette-Guérin Failures in Nonmuscle Invasive Bladder Cancer," *Journal of Urology*, vol. 190, pp. 850-856, 2013.
- [89] Ö. Şanlı and Y. Lotan, "Alternative therapies in patients with non-muscle invasive bladder cancer," *Turkish journal of urology*, vol. 43, p. 414, 2017.
- [90] W. M. Stadler, S. P. Lerner, S. Groshen, J. P. Stein, S.-R. Shi, D. Raghavan, *et al.*, "Phase III Study of Molecularly Targeted Adjuvant Therapy in Locally Advanced Urothelial Cancer of the Bladder Based on p53 Status," *Journal of Clinical Oncology*, vol. 29, pp. 3443-3449, 2011.
- [91] Q. Li, A. W. Damish, Z. Frazier, D. Liu, E. Reznichenko, A. Kamburov, *et al.*, "ERCC2 Helicase Domain Mutations Confer Nucleotide Excision Repair Deficiency and Drive Cisplatin

- Sensitivity in Muscle-Invasive Bladder Cancer," *Clinical Cancer Research*, vol. 25, pp. 977-988, 2019.
- [92] T. W. Flaig, C. M. Tangen, S. Daneshmand, A. Alva, S. P. Lerner, M. S. Lucia, *et al.*, "A Randomized Phase II Study of Coexpression Extrapolation (COXEN) with Neoadjuvant Chemotherapy for Bladder Cancer (SWOG S1314; NCT02177695)," *Clinical Cancer Research*, vol. 27, pp. 2435-2441, 2021.
 - [93] J. E. Rosenberg, J. Kim, A. Cherniack, C. Sougnez, M. Meyerson, G. Getz, *et al.*, "Somatic genomic alterations in urothelial cancer: Results of the Cancer Genome Atlas (TCGA) bladder cancer (BC) analysis," *Journal of Clinical Oncology*, vol. 32, pp. 285-285, 2014.
 - [94] K. T. Flaherty, R. Gray, A. Chen, S. Li, D. Patton, S. R. Hamilton, *et al.*, "The Molecular Analysis for Therapy Choice (NCI-MATCH) Trial: Lessons for Genomic Trial Design," *JNCI: Journal of the National Cancer Institute*, vol. 112, pp. 1021-1029, 2020.
 - [95] M. Z. Wojtukiewicz, M. M. Rek, K. Karpowicz, M. Górski, B. Polityńska, A. M. Wojtukiewicz, *et al.*, "Inhibitors of immune checkpoints—PD-1, PD-L1, CTLA-4—new opportunities for cancer patients and a new challenge for internists and general practitioners," *Cancer and Metastasis Reviews*, vol. 40, pp. 949-982, 2021/09/01 2021.
 - [96] M. D. Galsky, C. J. Hoimes, A. Necchi, N. Shore, J. A. Witjes, G. Steinberg, *et al.*, "Perioperative pembrolizumab therapy in muscle-invasive bladder cancer: Phase III KEYNOTE-866 and KEYNOTE-905/EV-303," *Future Oncology*, vol. 17, pp. 3137-3150, 2021.
 - [97] C. Cattrini and F. Boccardo, "Atezolizumab and bladder cancer: facing a complex disease," *The Lancet*, vol. 391, pp. 305-306, 2018.
 - [98] L. Hadjiiski, A. Z. Weizer, A. Alva, E. M. Caoili, R. H. Cohan, K. Cha, *et al.*, "Treatment Response Assessment for Bladder Cancer on CT Based on Computerized Volume Analysis, World Health Organization Criteria, and RECIST," *American Journal of Roentgenology*, vol. 205, pp. 348-352, 2015/08/01 2015.
 - [99] G. Petrella, G. Ciufolini, R. Vago, and D. O. Cicero, "Urinary Metabolic Markers of Bladder Cancer: A Reflection of the Tumor or the Response of the Body?," *Metabolites*, vol. 11, p. 756, 2021/10/31 2021.
 - [100] M. C. de Oliveira, H. R. Caires, M. J. Oliveira, A. Fraga, M. H. Vasconcelos, and R. Ribeiro, "Urinary Biomarkers in Bladder Cancer: Where Do We Stand and Potential Role of Extracellular Vesicles," *Cancers*, vol. 12, p. 1400, 2020/05/29 2020.
 - [101] J. Rubio-briones, F. Algaba, E. Gallardo, J. A. Marcos-rodíguez, M. Á. Climent, A. G. Caamaño, *et al.*, "Recent Advances in the Management of Patients with Non-Muscle-Invasive Bladder Cancer Using a Multidisciplinary Approach: Practical Recommendations from the Spanish Oncology Genitourinary (SOGUG) Working Group," *Cancers*, vol. 13, p. 4762, 2021/09/23 2021.
 - [102] E. Oeyen, L. Hoekx, S. De Wachter, M. Baldewijns, F. Ameye, and I. Mertens, "Bladder Cancer Diagnosis and Follow-Up: The Current Status and Possible Role of Extracellular Vesicles," *International Journal of Molecular Sciences*, vol. 20, p. 821, 2019/02/14 2019.
 - [103] K. Ng, A. Stenzl, A. Sharma, and N. Vasdev, "Urinary biomarkers in bladder cancer: A review of the current landscape and future directions," *Urologic Oncology: Seminars and Original Investigations*, vol. 39, pp. 41-51, 2021/01/01 2021.
 - [104] L. Tran, J. F. Xiao, N. Agarwal, J. E. Duex, and D. Theodorescu, "Advances in bladder cancer biology and therapy," *Nature Reviews Cancer*, vol. 21, pp. 104-121, 2020/12/02 2020.

- [105] K. Inamura, "Bladder Cancer: New Insights into Its Molecular Pathology," *Cancers*, vol. 10, p. 100, 2018/04/01 2018.
- [106] C. Lourenço, V. Constâncio, R. Henrique, Â. Carvalho, and C. Jerónimo, "Urinary Extracellular Vesicles as Potential Biomarkers for Urologic Cancers: An Overview of Current Methods and Advances," *Cancers*, vol. 13, p. 1529, 2021/03/26 2021.
- [107] F. Crick, "Central Dogma of Molecular Biology," *Nature*, vol. 227, pp. 561-563, 1970/08/01 1970.
- [108] M. K. Deyholos and D. W. Galbraith, "High-density microarrays for gene expression analysis," *Cytometry*, vol. 43, pp. 229-238, 2001.
- [109] M. Lohse, A. Nunes-Nesi, P. Krüger, A. Nagel, J. Hannemann, F. M. Giorgi, *et al.*, "Robin: An Intuitive Wizard Application for R-Based Expression Microarray Quality Assessment and Analysis " *Plant Physiology*, vol. 153, pp. 642-651, 2010.
- [110] A. Cambrosio, J. Campbell, E. Vignola-Gagné, P. Keating, B. R. Jordan, and P. Bourret, "'Overcoming the Bottleneck': Knowledge Architectures for Genomic Data Interpretation in Oncology," in *Data Journeys in the Sciences*, S. Leonelli and N. Tempini, Eds., ed Cham: Springer International Publishing, 2020, pp. 305-327.
- [111] D. Toro-Domínguez, J. A. Villatoro-García, J. Martorell-Marugán, Y. Román-Montoya, M. E. Alarcón-Riquelme, and P. Carmona-Sáez, "A survey of gene expression meta-analysis: methods and applications," *Briefings in Bioinformatics*, vol. 00, pp. 1-12, 2020 2020.
- [112] P. I. Kontou, A. Pavlopoulou, and P. G. Bagos, "Methods of Analysis and Meta-Analysis for Identifying Differentially Expressed Genes," in *Genetic Epidemiology: Methods and Protocols, Methods in Molecular Biology*. vol. 1793, E. Evangelou, Ed., ed, 2018, pp. 183-210.
- [113] T. E. Sweeney, W. A. Haynes, F. Vallania, J. P. Ioannidis, and P. Khatri, "Methods to increase reproducibility in differential gene expression via meta-analysis," *Nucleic Acids Research*, vol. 45, pp. 1-14, 2017 2017.
- [114] K. Sielemann, A. Hafner, and B. Pucker, "The reuse of public datasets in the life sciences: potential risks and rewards," *PeerJ*, vol. 8, p. e9954, 2020/09/22 2020.
- [115] N. A. Vasilevsky, J. Minnier, M. A. Haendel, and R. E. Champieux, "Reproducible and reusable research: are journal data sharing policies meeting the mark?," *PeerJ*, vol. 5, p. e3208, 2017/04/25 2017.
- [116] D. G. Roche, L. E. B. Kruuk, R. Lanfear, and S. A. Binning, "Public Data Archiving in Ecology and Evolution: How Well Are We Doing?," *PLOS Biology*, vol. 13, p. e1002295, 2015.
- [117] J. P. Wang, J. Y. Leng, R. K. Zhang, L. Zhang, B. Zhang, W. Y. Jiang, *et al.*, "Functional analysis of gene expression profiling-based prediction in bladder cancer," *Oncol Lett*, vol. 15, pp. 8417-8423, 2018/06/01 2018.
- [118] C. Zhang, M. Berndt-Paetz, and J. Neuhaus, "Identification of Key Biomarkers in Bladder Cancer: Evidence from a Bioinformatics Analysis," *Diagnostics*, vol. 10, p. 66, 2020/01/24 2020.
- [119] X. C. Mo, Z. T. Zhang, M. J. Song, Z. Q. Zhou, J. X. Zeng, Y. F. Du, *et al.*, "Screening and identification of hub genes in bladder cancer by bioinformatics analysis and KIF11 is a potential prognostic biomarker," *Oncology Letters*, vol. 21, pp. 1-10, 2021 2021.
- [120] M. Yan, X. Jing, Y. Liu, and X. Cui, "Screening and identification of key biomarkers in bladder carcinoma: Evidence from bioinformatics analysis," *Oncol Lett*, vol. 16, pp. 3092-3100, 2018/09/01 2018.

- [121] Y. Xu, G. Wu, J. Li, J. Li, N. Ruan, L. Ma, *et al.*, "Screening and Identification of Key Biomarkers for Bladder Cancer: A Study Based on TCGA and GEO Data," *BioMed Research International*, vol. 2020, p. 8283401, 2020/01/25 2020.
- [122] X. Gao, Y. Chen, M. Chen, S. Wang, X. Wen, and S. Zhang, "Identification of key candidate genes and biological pathways in bladder cancer," *PeerJ*, vol. 6, 2018 2018.
- [123] S. Verma, E. Shankar, S. Lin, V. Singh, E. R. Chan, S. Cao, *et al.*, "Identification of key genes associated with progression and prognosis of bladder cancer through integrated bioinformatics analysis," *Cancers*, vol. 13, p. 5931, 2021/12/01 2021.
- [124] B. M. Zeeshan Hameed, V. L. S. Aiswarya Dhavileswarapu, S. Z. Raza, H. Karimi, H. S. Khanuja, D. K. Shetty, *et al.*, "Artificial Intelligence and Its Impact on Urological Diseases and Management: A Comprehensive Review of the Literature," *Journal of Clinical Medicine*, vol. 10, p. 1864, 2021/04/26 2021.
- [125] T. Barrett, T. Suzek, D. Troup, S. Wilhite, W.-C. Ngau, P. Ledoux, *et al.*, "NCBI GEO: mining millions of expression profiles--database and tools," *Nucleic acids research*, vol. 33, 2005/01/01 2005.
- [126] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021/03/29 2021.
- [127] T. Bammiller, R. P. Beyer, S. Bhattacharya, G. A. Boorman, A. Boyles, B. U. Bradford, *et al.*, "Standardizing global gene expression analysis between laboratories and across platforms," *Nature Methods*, vol. 2, pp. 351-356, 2005/04/21 2005.
- [128] C. Walsh, P. Hu, J. Batt, and C. Santos, "Microarray Meta-Analysis and Cross-Platform Normalization: Integrative Genomics for Robust Biomarker Discovery," *Microarrays*, vol. 4, pp. 389-406, 2015 2015.
- [129] L. Gautier, L. Cope, B. M. Bolstad, and R. A. Irizarry, "affy—analysis of Affymetrix GeneChip data at the probe level," *Bioinformatics*, vol. 20, pp. 307-315, 2004/02/12 2004.
- [130] B. S. Carvalho and R. A. Irizarry, "A framework for oligonucleotide microarray preprocessing," *Bioinformatics*, vol. 26, pp. 2363-2367, 2010/10/01 2010.
- [131] R. A. Irizarry, B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs, and T. P. Speed, "Summaries of Affymetrix GeneChip probe level data," *Nucleic Acids Research*, vol. 31, pp. e15-e15, 2003/02/15 2003.
- [132] R. A. Irizarry, D. Warren, F. Spencer, I. F. Kim, S. Biswal, B. C. Frank, *et al.*, "Multiple-laboratory comparison of microarray platforms," *Nature Methods*, vol. 2, pp. 345-350, 2005/04/21 2005.
- [133] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, *et al.*, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, pp. e47-e47, 2015/04/20 2015.
- [134] S. Calza and Y. Pawitan, "Normalization of Gene-Expression Microarray Data," *Methods in molecular biology*, vol. 673, pp. 37-52, 2010 2010.
- [135] A. Kauffmann, R. Gentleman, and W. Huber, "arrayQualityMetrics--a bioconductor package for quality assessment of microarray data," *Bioinformatics*, vol. 25, pp. 415-416, 2009 2009.
- [136] A. Kauffmann and W. Huber, "Microarray data quality control improves the detection of differentially expressed genes," *Genomics*, vol. 95, pp. 138-142, 2010/03/01 2010.
- [137] S. Tweedie, B. Braschi, K. Gray, T. E. M. Jones, R. L. Seal, B. Yates, *et al.*, "Genenames.org: the HGNC and VGNC resources in 2021," *Nucleic Acids Research*, vol. 49, pp. D939-D946, 2021/01/08 2021.

- [138] B. Braschi, R. L. Seal, S. Tweedie, T. E. M. Jones, and E. A. Bruford, "The risks of using unapproved gene symbols," *The American Journal of Human Genetics*, vol. 108, pp. 1813-1816, 2021/10/07 2021.
- [139] B. L. Aken, S. Ayling, D. Barrell, L. Clarke, V. Curwen, S. Fairley, *et al.*, "The Ensembl gene annotation system," *Database: The Journal of Biological Databases and Curation*, vol. 2016, 2016 2016.
- [140] M. R. J. Carlson, H. Pagès, S. Arora, V. Obenchain, and M. Morgan, "Genomic Annotation Resources in R/Bioconductor," *Methods in molecular biology*, vol. 1418, pp. 67-90, 2016 2016.
- [141] B. Ballester, N. Johnson, G. Proctor, and P. Flicek, "Consistent annotation of gene expression arrays," *BMC Genomics*, vol. 11, pp. 1-14, 2010/05/11 2010.
- [142] B. W. W. Goh, W. Wang, and L. Wong, "Why Batch Effects Matter in Omics Data, and How to Avoid Them," *Trends in Biotechnology*, vol. 35, pp. 498-507, 2017/06/01 2017.
- [143] C. Lazar, S. Meganck, J. Taminiau, D. Steenhoff, A. Coletta, C. Molter, *et al.*, "Batch effect removal methods for microarray gene expression data integration: A survey," *Briefings in Bioinformatics*, vol. 14, pp. 469-490, 2013 2013.
- [144] C. Cheadle, M. P. Vawter, W. J. Freed, and K. G. Becker, "Analysis of microarray data using Z score transformation," *Journal of Molecular Diagnostics*, vol. 5, pp. 73-81, 2003 2003.
- [145] H. Yasrebi, "Comparative study of joint analysis of microarray gene expression data in survival prediction and risk assessment of breast cancer patients," *Briefings in Bioinformatics*, vol. 17, pp. 771-785, 2016/09/01 2016.
- [146] B. Zhou and R. Guo, "Integrative Analysis of Genomic and Clinical Data Reveals Intrinsic Characteristics of Bladder Urothelial Carcinoma Progression," *Genes*, vol. 10, p. 464, 2019 2019.
- [147] S. Balivada, C. K. Ganta, Y. Zhang, H. N. Pawar, R. J. Ortiz, K. G. Becker, *et al.*, "Microarray analysis of aging-associated immune system alterations in the rostral ventrolateral medulla of F344 rats," *Physiological genomics*, vol. 49, pp. 400-415, 2017 2017.
- [148] V. Nygaard, E. A. Rødland, and E. Hovig, "Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses," *Biostatistics*, vol. 17, pp. 29-39, 2016 2016.
- [149] M. R. Dalman, A. Deeter, G. Nimishakavi, and Z. H. Duan, "Fold change and p-value cutoffs significantly alter microarray interpretations," *BMC bioinformatics*, vol. 13 Suppl 2, 2012 2012.
- [150] A. Statnikov, L. Wang, and C. F. Aliferis, "A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification," *BMC Bioinformatics*, vol. 9, pp. 1-10, 2008/07/22 2008.
- [151] T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, *et al.*, "clusterProfiler 4.0: A universal enrichment tool for interpreting omics data," *The Innovation*, vol. 2, 2021/08/28 2021.
- [152] S. Carbon, E. Douglass, N. Dunn, B. Good, N. L. Harris, S. E. Lewis, *et al.*, "The Gene Ontology Resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, pp. D330-D338, 2019/01/08 2019.
- [153] J. Z. Wang, Z. Du, R. Payattakool, P. S. Yu, and C. F. Chen, "A new method to measure the semantic similarity of GO terms," *Bioinformatics*, vol. 23, pp. 1274-1281, 2007/05/15 2007.
- [154] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Research*, vol. 45, pp. D353-D361, 2017/01/04 2017.
- [155] B. Jassal, L. Matthews, G. Viteri, C. Gong, P. Lorente, A. Fabregat, *et al.*, "The reactome pathway knowledgebase," *Nucleic Acids Research*, vol. 48, pp. D498-D503, 2020/01/08 2020.

- [156] G. Yu and Q. Y. He, "ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization," *Molecular bioSystems*, vol. 12, pp. 477-479, 2016 2016.
- [157] D. Merico, R. Isserlin, O. Stueker, A. Emili, and G. D. Bader, "Enrichment map: a network-based method for gene-set enrichment visualization and interpretation," *PLOS ONE*, vol. 5, 2010 2010.
- [158] L. M. Schriml, E. Mitraha, J. Munro, B. Tauber, M. Schor, L. Nickle, *et al.*, "Human Disease Ontology 2018 update: classification, content and workflow expansion," *Nucleic Acids Research*, vol. 47, pp. D955-D962, 2019/01/08 2019.
- [159] G. Yu, L. G. Wang, G. R. Yan, and Q. Y. He, "DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis," *Bioinformatics*, vol. 31, pp. 608-609, 2015/02/15 2015.
- [160] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, *et al.*, "STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, pp. D607-D613, 2019/01/08 2019.
- [161] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, *et al.*, "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks," *Genome Research*, vol. 13, pp. 2498-2504, 2003/11/01 2003.
- [162] C. H. Chin, S. H. Chen, H. H. Wu, C. W. Ho, M. T. Ko, and C. Y. Lin, "cytoHubba: Identifying hub objects and sub-networks from complex interactome," *BMC Systems Biology*, vol. 8, pp. 1-7, 2014/12/08 2014.
- [163] G. D. Bader and C. W. V. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC bioinformatics*, vol. 4, 2003/01/13 2003.
- [164] B. Zhang and S. Horvath, "A general framework for weighted gene co-expression network analysis," *Statistical applications in genetics and molecular biology*, vol. 4, 2005/08/12 2005.
- [165] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC bioinformatics*, vol. 9, 2008/12/29 2008.
- [166] J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey, "The sva package for removing batch effects and other unwanted variation in high-throughput experiments," *Bioinformatics*, vol. 28, p. 882, 2012/03 2012.
- [167] A. Perez, A. Loizaga, R. Arceo, I. Lacasa, A. Rabade, K. Zorroza, *et al.*, "A Pilot Study on the Potential of RNA-Associated to Urinary Vesicles as a Suitable Non-Invasive Source for Diagnostic Purposes in Bladder Cancer," *Cancers*, vol. 6, pp. 179-192, 2014/03 2014.
- [168] O. G. Troyanskaya, M. E. Garber, P. O. Brown, D. Botstein, and R. B. Altman, "Nonparametric methods for identifying differentially expressed genes in microarray data," *Bioinformatics*, vol. 18, pp. 1454-1461, 2002/11/01 2002.
- [169] J.-S. Lee, S.-H. Leem, S.-Y. Lee, S.-C. Kim, E.-S. Park, S.-B. Kim, *et al.*, "Expression signature of E2F1 and its associated genes predict superficial to invasive progression of bladder tumors," *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 28, pp. 2660-2667, 2010/06/01 2010.
- [170] T. M. Therneau and P. M. Grambsch, *Modeling Survival Data: Extending the Cox Model*, 1 ed. New York, NY, 2000.
- [171] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," *Journal of statistical software*, vol. 33, p. 1, 2010 2010.
- [172] P. J. Heagerty, T. Lumley, and M. S. Pepe, "Time-dependent ROC curves for censored survival data and a diagnostic marker," *Biometrics*, vol. 56, pp. 337-344, 2000 2000.

- [173] M. C. Sachs, "plotROC: A Tool for Plotting ROC Curves," *Journal of statistical software*, vol. 79, 2017 2017.
- [174] G. Sjö Dahl, M. Lauss, K. Lövgren, G. Chebil, S. Gudjonsson, S. Veerla, *et al.*, "A molecular taxonomy for urothelial carcinoma," *Clinical Cancer Research*, vol. 18, pp. 3377-3386, 2012/06/15 2012.
- [175] D. Lindgren, G. Sjö Dahl, M. Lauss, J. Staaf, G. Chebil, K. Lövgren, *et al.*, "Integrated Genomic and Gene Expression Profiling Identifies Two Major Genomic Circuits in Urothelial Carcinoma," *PLOS ONE*, vol. 7, p. e38863, 2012/06/07 2012.
- [176] Z. Tang, B. Kang, C. Li, T. Chen, and Z. Zhang, "GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis," *Nucleic Acids Research*, vol. 47, pp. W556-W560, 2019/07/02 2019.
- [177] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge," *Contemporary oncology*, vol. 19, pp. A68-A77, 2015 2015.
- [178] J. Lonsdale, J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, *et al.*, "The Genotype-Tissue Expression (GTEx) project," *Nature genetics*, vol. 45, pp. 580-585, 2013/06 2013.
- [179] M. L. Iacovino, C. C. Miceli, M. De Felice, B. Barone, L. Pompella, F. Chiancone, *et al.*, "Novel Therapeutic Opportunities in Neoadjuvant Setting in Urothelial Cancers: A New Horizon Opened by Molecular Classification and Immune Checkpoint Inhibitors," *International Journal of Molecular Sciences*, vol. 23, p. 1133, 2022/01/20 2022.
- [180] G. Sjö Dahl, J. Abrahamsson, K. Holmsten, C. Bernardo, G. Chebil, P. Eriksson, *et al.*, "Different Responses to Neoadjuvant Chemotherapy in Urothelial Carcinoma Molecular Subtypes," *European Urology*, 2021/11/13 2021.
- [181] W. E. Johnson, C. Li, and A. Rabinovic, "Adjusting batch effects in microarray expression data using empirical Bayes methods," *Biostatistics*, vol. 8, pp. 118-127, 2007/01/01 2007.
- [182] R. Seiler, H. A. D. Ashab, N. Erho, B. W. G. van Rhijn, B. Winters, J. Douglas, *et al.*, "Impact of Molecular Subtypes in Muscle-invasive Bladder Cancer on Predicting Response and Survival after Neoadjuvant Chemotherapy," *European urology*, vol. 72, pp. 544-554, 2017/10/01 2017.
- [183] D. J. McConkey, W. Choi, Y. Shen, I. L. Lee, S. Porten, S. F. Matin, *et al.*, "A Prognostic Gene Expression Signature in the Molecular Classification of Chemotherapy-naïve Urothelial Cancer is Predictive of Clinical Outcomes from Neoadjuvant Chemotherapy: A Phase 2 Trial of Dose-dense Methotrexate, Vinblastine, Doxorubicin, and Cisplatin with Bevacizumab in Urothelial Cancer," *European urology*, vol. 69, pp. 855-862, 2016/05/01 2016.
- [184] M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, *et al.*, "Tissue-based map of the human proteome," *Science*, vol. 347, 2015/01/23 2015.
- [185] A. Brazma, H. Parkinson, U. Sarkans, M. Shojatalab, J. Vilo, N. Abeygunawardena, *et al.*, "ArrayExpress—a public repository for microarray gene expression data at the EBI," *Nucleic Acids Research*, vol. 31, pp. 68-71, 2003/01/01 2003.
- [186] S. A. Hussain, D. H. Palmer, W. K. Syn, J. J. Sacco, R. M. D. Greensmith, T. Elmetwali, *et al.*, "Gene expression profiling in bladder cancer identifies potential therapeutic targets," *International journal of oncology*, vol. 50, pp. 1147-1159, 2017/04/01 2017.
- [187] L. Dyrskjöt, M. Kruhøffer, T. Thykjaer, N. Marcussen, J. Jensen, K. Møller, *et al.*, "Gene expression in the urinary bladder: a common carcinoma in situ gene expression signature exists disregarding histopathological classification," *Cancer Research*, vol. 64, pp. 4040-4048, 2004/06/01 2004.

- [188] L. Mengual, M. Burset, E. Ars, J. J. Lozano, H. Villavicencio, M. J. Ribal, *et al.*, "DNA Microarray Expression Profiling of Bladder Cancer Allows Identification of Noninvasive Diagnostic Markers," *Journal of Urology*, vol. 182, pp. 741-748, 2009/08 2009.
- [189] U. Gabriel, L. Li, C. Bolenz, A. Steidler, B. Kränzlin, M. Saile, *et al.*, "New insights into the influence of cigarette smoking on urothelial carcinogenesis: smoking-induced gene expression in tumor-free urothelium might discriminate muscle-invasive from nonmuscle-invasive urothelial bladder cancer," *Molecular carcinogenesis*, vol. 51, pp. 907-915, 2012/11 2012.
- [190] Z. Zhang, K. A. Furge, X. J. Yang, B. T. Teh, and D. E. Hansel, "Comparative gene expression profiling analysis of urothelial carcinoma of the renal pelvis and bladder," *BMC Medical Genomics*, vol. 3, pp. 1-10, 2010/12/15 2010.
- [191] V. Urquidi, S. Goodison, Y. Cai, Y. Sun, and C. Rosser, "A candidate molecular biomarker panel for the detection of bladder cancer," *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, vol. 21, pp. 2149-2158, 2012/12 2012.
- [192] Y.-J. Kim, H.-Y. Yoon, J. S. Kim, H. W. Kang, B.-D. Min, S.-K. Kim, *et al.*, "HOXA9, ISL1 and ALDH1A3 methylation patterns as prognostic markers for nonmuscle invasive bladder cancer: array-based DNA methylation and expression profiling," *International journal of cancer*, vol. 133, pp. 1135-1142, 2013/09/01 2013.
- [193] M. Santos, M. Martínez-Fernández, M. Dueñas, R. García-Escudero, B. Alfaya, F. Villacampa, *et al.*, "In Vivo Disruption of an Rb–E2F–Ezh2 Signaling Loop Causes Bladder Cancer," *Cancer Research*, vol. 74, pp. 6565-6577, 2014/11/15 2014.
- [194] N. Hecker, C. Stephan, H.-J. Mollenkopf, K. Jung, R. Preissner, and H.-A. Meyer, "A new algorithm for integrated analysis of miRNA-mRNA interactions based on individual classification reveals insights into bladder cancer," *PLOS ONE*, vol. 8, 2013/05/24 2013.
- [195] F. Roudnický, C. Poyet, P. Wild, S. Krampitz, F. Negrini, R. Huggenberger, *et al.*, "Endocan is upregulated on tumor vessels in invasive bladder cancer where it mediates VEGF-A-induced angiogenesis," *Cancer Research*, vol. 73, pp. 1097-1106, 2013/02/01 2013.
- [196] N. Zhou, K. Singh, M. Mir, Y. Parker, D. Lindner, R. Dreicer, *et al.*, "The investigational Aurora kinase A inhibitor MLN8237 induces defects in cell viability and cell-cycle progression in malignant bladder cancer cells in vitro and in vivo," *Clinical cancer research: an official journal of the American Association for Cancer Research*, vol. 19, pp. 1717-1728, 2013/04/01 2013.
- [197] W. He, Q. Cai, F. Sun, G. Zhong, P. Wang, H. Liu, *et al.*, "linc-UBC1 physically associates with polycomb repressive complex 2 (PRC2) and acts as a negative prognostic factor for lymph node metastasis and survival in bladder cancer," *Biochimica et biophysica acta*, vol. 1832, pp. 1528-1537, 2013/10 2013.
- [198] N. Borisov, V. Tkachev, M. Suntsova, O. Kovalchuk, A. Zhavoronkov, I. Muchnik, *et al.*, "A method of gene expression data transfer from cell lines to cancer patients for machine-learning prediction of drug efficiency," *Cell cycle*, vol. 17, pp. 486-491, 2018/02/16 2018.
- [199] L. Chen, L. Yuan, G. Wang, R. Cao, J. Peng, B. Shu, *et al.*, "Identification and bioinformatics analysis of miRNAs associated with human muscle invasive bladder cancer," *Molecular medicine reports*, vol. 16, pp. 8709-8720, 2017/12/01 2017.
- [200] W. He, G. Zhong, N. Jiang, B. Wang, X. Fan, C. Chen, *et al.*, "Long noncoding RNA BLACAT2 promotes bladder cancer-associated lymphangiogenesis and lymphatic metastasis," *The Journal of clinical investigation*, vol. 128, pp. 861-875, 2018/02/01 2018.

- [201] A. Loras, C. Suárez-Cabrera, C. Martínez-Bisbal, G. Quintás, J. M. Paramio, R. Martínez-Máñez, *et al.*, "Integrative Metabolomic and Transcriptomic Analysis for the Study of Bladder Cancer," *Cancers*, vol. 11, 2019/05/01 2019.
- [202] K. Ng, K. Vinnakota, A. Sharma, J. Kelly, P. Dasgupta, and N. Vasdev, "Urinary biomarkers to mitigate diagnostic delay in bladder cancer during the COVID-19 era," *Nature Reviews Urology*, vol. 18, pp. 185-187, 2021 2021.
- [203] D. Hanahan and R. A. Weinberg, "The Hallmarks of Cancer," *Cell*, vol. 100, pp. 57-70, 2000/01/07 2000.
- [204] D. Hanahan, "Hallmarks of Cancer: New DimensionsHallmarks of Cancer: New Dimensions," *Cancer Discovery*, vol. 12, pp. 31-46, 2022/01/01 2022.
- [205] M. Charpentier, C. Gutierrez, T. Guillaudeux, G. Verhoest, and R. Pedeux, "Noninvasive urine-based tests to diagnose or detect recurrence of bladder cancer," *Cancers*, vol. 13, p. 1650, 2021/04/01 2021.
- [206] M. A. Icer, M. Gezmen-Karadag, and S. Sozen, "Can urine osteopontin levels, which may be correlated with nutrition intake and body composition, be used as a new biomarker in the diagnosis of nephrolithiasis?," *Clinical biochemistry*, vol. 60, pp. 38-43, 2018/09/01 2018.
- [207] F. Yao, X. Hong, S. Li, Y. Zhang, Q. Zhao, W. Du, *et al.*, "Urine-Based Biomarkers for Alzheimer's Disease Identified Through Coupling Computational and Experimental Methods," *Journal of Alzheimer's Disease*, vol. 65, pp. 421-431, 2018/01/01 2018.
- [208] W. Jiang, T. Ma, C. Zhang, X. Tang, Q. Xu, X. Meng, *et al.*, "Identification of urinary candidate biomarkers of cisplatin-induced nephrotoxicity in patients with carcinoma," *Journal of Proteomics*, vol. 210, p. 103533, 2020/01/06 2020.
- [209] Y. Shimizu, H. Furuya, P. Bryant Greenwood, O. Chan, Y. Dai, M. D. Thornquist, *et al.*, "A multiplex immunoassay for the non-invasive detection of bladder cancer," *Journal of Translational Medicine*, vol. 14, pp. 1-9, 2016/01/30 2016.
- [210] C. J. Rosser, M. Chang, Y. Dai, S. Ross, L. Mengual, A. Alcaraz, *et al.*, "Urinary protein biomarker panel for the detection of recurrent bladder cancer," *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, vol. 23, pp. 1340-1345, 2014 2014.
- [211] W. T. Kim, P. Jeong, C. Yan, Y. H. Kim, I. S. Lee, H. W. Kang, *et al.*, "UBE2C cell-free RNA in urine can discriminate between bladder cancer and hematuria," *Oncotarget*, vol. 7, p. 58193, 2016 2016.
- [212] A. S. Kononikhin, V. N. Zakharova, V. A. Sergeeva, M. I. Indeykina, N. L. Starodubtseva, A. E. Bugrova, *et al.*, "Differential Diagnosis of Preeclampsia Based on Urine Peptidome Features Revealed by High Resolution Mass Spectrometry," *Diagnostics*, vol. 10, 2020/12/01 2020.
- [213] K. Stoeber, R. Swinn, A. T. Prevost, P. De Clive-Lowe, I. Halsall, S. M. Dilworth, *et al.*, "Diagnosis of Genito-Urinary Tract Cancer by Detection of Minichromosome Maintenance 5 Protein in Urine Sediments," *JNCI: Journal of the National Cancer Institute*, vol. 94, pp. 1071-1079, 2002/07/17 2002.
- [214] S. H. Lee, P. H. Lee, B. G. Kim, J. Hong, and A. S. Jang, "Annexin A5 Protein as a Potential Biomarker for the Diagnosis of Asthma," *Lung*, vol. 196, pp. 681-689, 2018/12/01 2018.
- [215] K. C. Ang, S. Kathirgamanathan, E. S. Ch'ng, W. Z. Abdullah, N. M. Yusoff, C. M. Jahnke, *et al.*, "Elevated annexin A5 plasma levels in term pregnancies of M2/ANXA5 carriers," *Thrombosis Research*, vol. 156, pp. 87-90, 2017/08/01 2017.

- [216] W. M. Serag, B. S. e. Mohammed, M. M. Mohamed, and B. E. Elsayed, "Predicting the risk of portal vein thrombosis in patients with liver cirrhosis and hepatocellular carcinoma," *Heliyon*, vol. 6, 2020/08/01 2020.
- [217] N. Laprovitera, I. Salamon, F. Gelsomino, E. Porcellini, M. Riefolo, M. Garonzi, *et al.*, "Genetic Characterization of Cancer of Unknown Primary Using Liquid Biopsy Approaches," *Frontiers in cell and developmental biology*, vol. 9, 2021/06/10 2021.
- [218] W. Marlicz, K. Sielatycka, K. Serwin, E. Kubis, M. Tkacz, R. Guszko, *et al.*, "Effect of colorectal cancer on the number of normal stem cells circulating in peripheral blood," *Oncology reports*, vol. 36, pp. 3635-3642, 2016/12/01 2016.
- [219] D. Zhou, W. Tang, X. Liu, H. X. An, and Y. Zhang, "Clinical verification of plasma messenger RNA as novel noninvasive biomarker identified through bioinformatics analysis for lung cancer," *Oncotarget*, vol. 8, pp. 43978-43989, 2017 2017.
- [220] R. Garcia, C. Bouleti, M. Sirol, D. Logeart, C. Monnot, C. Ardidie-Robouant, *et al.*, "VEGF-A plasma levels are associated with microvascular obstruction in patients with ST-segment elevation myocardial infarction," *International journal of cardiology*, vol. 291, pp. 19-24, 2019/09/15 2019.
- [221] D. M. Smadja, S. J. Mentzer, M. Fontenay, M. A. Laffan, M. Ackermann, J. Helms, *et al.*, "COVID-19 is a systemic vascular hemopathy: insight for mechanistic and clinical aspects," *Angiogenesis*, vol. 24, pp. 755-788, 2021/11/01 2021.
- [222] F. Innocenti, C. Jiang, A. B. Sibley, A. S. Etheridge, A. J. Hatch, S. Denning, *et al.*, "Genetic variation determines VEGF-A plasma levels in cancer patients," *Scientific reports*, vol. 8, 2018/12/01 2018.
- [223] A. Periyasamy, G. Gopisetty, M. J. Subramaniam, S. Velusamy, and T. Rajkumar, "Identification and validation of differential plasma proteins levels in epithelial ovarian cancer," *Journal of proteomics*, vol. 226, 2020/08/30 2020.
- [224] M. M. Karsten, M. H. Beck, A. Rademacher, J. Knabl, J. U. Blohmer, J. Jückstock, *et al.*, "VEGF-A165b levels are reduced in breast cancer patients at primary diagnosis but increase after completion of cancer treatment," *Scientific reports*, vol. 10, 2020/12/01 2020.
- [225] M. Ingenwerth, P. Nyirády, B. Hadaschik, T. Szarvas, and H. Reis, "The prognostic value of cytokeratin and extracellular collagen expression in urinary bladder cancer," *Current molecular medicine*, vol. 21, 2021/02/26 2021.
- [226] Z. Zhong, M. Huang, M. Lv, Y. He, C. Duan, L. Zhang, *et al.*, "Circular RNA MYLK as a competing endogenous RNA promotes bladder cancer progression through modulating VEGFA/VEGFR2 signaling pathway," *Cancer letters*, vol. 403, pp. 305-317, 2017/09/10 2017.
- [227] L. Yi, H. Wang, W. Li, K. Ye, W. Xiong, H. Yu, *et al.*, "The FOXM1/RNF26/p57 axis regulates the cell cycle to promote the aggressiveness of bladder cancer," *Cell Death & Disease*, vol. 12, 2021/10/01 2021.
- [228] S. Rinaldetti, R. Wirtz, T. S. Worst, A. Hartmann, J. Breyer, L. Dyrskjot, *et al.*, "FOXM1 predicts disease progression in non-muscle invasive bladder cancer," *Journal of cancer research and clinical oncology*, vol. 144, pp. 1701-1709, 2018/09/01 2018.
- [229] Y. G. Roh, J. Y. Mun, S. K. Kim, W. Park, M. S. Jeong, T. N. Kim, *et al.*, "Fanconi Anemia Pathway Activation by FOXM1 Is Critical to Bladder Cancer Recurrence and Anticancer Drug Resistance," *Cancers*, vol. 12, pp. 1-20, 2020/06/01 2020.
- [230] Z. Yang, H. Sun, W. Ma, K. Wu, G. Peng, T. Ou, *et al.*, "Down-regulation of Polo-like kinase 4 (PLK4) induces G1 arrest via activation of the p38/p53/p21 signaling pathway in bladder cancer," *FEBS open bio*, vol. 11, pp. 2631-2646, 2021/09/01 2021.

- [231] L. Yang, C. Li, Y. Qin, G. Zhang, B. Zhao, Z. Wang, *et al.*, "A Novel Prognostic Model Based on Ferroptosis-Related Gene Signature for Bladder Cancer," *Frontiers in Oncology*, vol. 11, p. 3070, 2021/08/06 2021.
- [232] Z. Xie, J. Cai, W. Sun, S. Hua, X. Wang, A. Li, *et al.*, "Development and Validation of Prognostic Model in Transitional Bladder Cancer Based on Inflammatory Response-Associated Genes," *Frontiers in Oncology*, vol. 11, p. 4033, 2021/10/07 2021.
- [233] J. Lin, J. Yang, X. Xu, Y. Wang, M. Yu, and Y. Zhu, "A robust 11-genes prognostic model can predict overall survival in bladder cancer patients based on five cohorts," *Cancer Cell International*, vol. 20, pp. 1-14, 2020/08/20 2020.
- [234] F. Tang, Z. Li, Y. Lai, Z. Lu, H. Lei, C. He, *et al.*, "A 7-gene signature predicts the prognosis of patients with bladder cancer," *BMC Urology*, vol. 22, pp. 1-12, 2022/12/01 2022.
- [235] C. Zhou, A. H. Li, S. Liu, and H. Sun, "Identification of an 11-Autophagy-Related-Gene Signature as Promising Prognostic Biomarker for Bladder Cancer Patients," *Biology*, vol. 10, p. 375, 2021/04/27 2021.
- [236] F. Xu, Q. Tang, Y. Wang, G. Wang, K. Qian, L. Ju, *et al.*, "Development and Validation of a Six-Gene Prognostic Signature for Bladder Cancer," *Frontiers in Genetics*, vol. 12, p. 2395, 2021/12/06 2021.
- [237] F. Chen, Q. Wang, and Y. Zhou, "The construction and validation of an RNA binding protein-related prognostic model for bladder cancer," *BMC Cancer*, vol. 21, pp. 1-13, 2021/12/01 2021.
- [238] Y. Hu, Y. Zhang, J. Gao, X. Lian, and Y. Wang, "The clinicopathological and prognostic value of CD44 expression in bladder cancer: a study based on meta-analysis and TCGA data," *Bioengineered*, vol. 11, pp. 572-581, 2020/01/01 2020.
- [239] T. C. Wu, W. Y. Lin, W. C. Chen, and M. F. Chen, "Predictive Value of CD44 in Muscle-Invasive Bladder Cancer and Its Relationship with IL-6 Signaling," *Annals of surgical oncology*, vol. 25, pp. 3518-3526, 2018/11/01 2018.
- [240] T. C. Wu, W. Y. Lin, Y. H. Chang, W. C. Chen, and M. F. Chen, "Impact of CD44 expression on radiation response for bladder cancer," *Journal of Cancer*, vol. 8, pp. 1137-1144, 2017/04/09 2017.
- [241] J. L. Sottnik, L. Vanderlinden, M. Joshi, A. Chauca-Diaz, C. Owens, D. E. Hansel, *et al.*, "Androgen Receptor Regulates CD44 Expression in Bladder Cancer," *Cancer Research*, vol. 81, pp. 2833-2846, 2021/06/01 2021.
- [242] D. Sasca, J. Szybinski, A. Schöler, V. Shah, J. Heidelberger, P. S. Haehnel, *et al.*, "NCAM1 (CD56) promotes leukemogenesis and confers drug resistance in AML," *Blood*, vol. 133, pp. 2305-2319, 2019/05/23 2019.
- [243] R. Shukrun, H. Golan, R. Caspi, N. Pode-Shakked, O. Pleniceanu, E. Vax, *et al.*, "NCAM1/FGF module serves as a putative pleuropulmonary blastoma therapeutic target," *Oncogenesis*, vol. 8, 2019/09/01 2019.
- [244] I. T. Øvestad, B. Engesæter, M. K. Halle, S. Akbari, B. Bicskei, M. Lapin, *et al.*, "High-Grade Cervical Intraepithelial Neoplasia (CIN) Associates with Increased Proliferation and Attenuated Immune Signaling," *International Journal of Molecular Sciences*, vol. 23, 2021/01/01 2021.
- [245] C. Cheng, X. Wu, Y. Shen, and Q. Li, "KIF14 and KIF23 Promote Cell Proliferation and Chemoresistance in HCC Cells, and Predict Worse Prognosis of Patients with HCC," *Cancer management and research*, vol. 12, pp. 13241-13257, 2020 2020.
- [246] W. Wang, Y. Shi, J. Li, W. Cui, and B. Yang, "Up-regulation of KIF14 is a predictor of poor survival and a novel prognostic biomarker of chemoresistance to paclitaxel treatment in cervical cancer," *Bioscience reports*, vol. 36, 2016/04/01 2016.

- [247] T. F. Li, H. J. Zeng, Z. Shan, R. Y. Ye, T. Y. Cheang, Y. J. Zhang, *et al.*, "Overexpression of kinesin superfamily members as prognostic biomarkers of breast cancer," *Cancer cell international*, vol. 20, 2020/04/15 2020.
- [248] A. Klimaszewska-Wiśniewska, I. Neska-Długosz, K. Buchholz, J. Durślewicz, D. Grzanka, A. Kasperska, *et al.*, "Prognostic Significance of KIF11 and KIF14 Expression in Pancreatic Adenocarcinoma," *Cancers*, vol. 13, 2021/06/02 2021.
- [249] W. Jiang, D. Zhu, C. Wang, and Y. Zhu, "An immune relevant signature for predicting prognoses and immunotherapeutic responses in patients with muscle-invasive bladder cancer (MIBC)," *Cancer Medicine*, vol. 9, p. 2774, 2020/04/01 2020.
- [250] C. Shen, T. Xu, Y. Sun, L. Wang, Z. Liang, H. Niu, *et al.*, "Construction of an Immune-Associated Gene-Based Signature in Muscle-Invasive Bladder Cancer," *Disease markers*, vol. 2020, 2020 2020.
- [251] S. J. Choi, K. J. Park, C. Heo, B. W. Park, M. Kim, and J. K. Kim, "Radiomics-based model for predicting pathological complete response to neoadjuvant chemotherapy in muscle-invasive bladder cancer," *Clinical Radiology*, vol. 76, pp. 627.e13-627.e21, 2021/08/01 2021.
- [252] A. Parmar, A. A. Qazi, A. Stundzia, H. W. Sim, J. Lewin, U. Metser, *et al.*, "Development of a radiomic signature for predicting response to neoadjuvant chemotherapy in muscle-invasive bladder cancer," *Canadian Urological Association Journal*, vol. 16, p. E113, 2022/03/01 2022.
- [253] M. P. Scholtes, A. R. Alberts, I. G. Iflé, P. C. M. S. Verhagen, A. A. M. van der Veldt, and T. C. M. Zuiverloon, "Biomarker-Oriented Therapy in Bladder and Renal Cancer," *International Journal of Molecular Sciences*, vol. 22, p. 2832, 2021/03/11 2021.
- [254] A. Valdés, A. Bitzios, E. Kassa, G. Shevchenko, A. Falk, P. U. Malmström, *et al.*, "Proteomic comparison between different tissue preservation methods for identification of promising biomarkers of urothelial bladder cancer," *Scientific Reports*, vol. 11, p. 7595, 2021/12/01 2021.
- [255] Y. Fan, C. Jiang, S. Li, X. Yao, X. Qi, Y. Wang, *et al.*, "Identification and Validation of an Annexin-Related Prognostic Signature and Therapeutic Targets for Bladder Cancer: Integrative Analysis," *Biology*, vol. 11, p. 259, 2022/02/07 2022.
- [256] W. B. Wu, G. Z. Jia, L. Chen, H. T. Liu, and S. J. Xia, "Analysis of the Expression and Prognostic Value of Annexin Family Proteins in Bladder Cancer," *Frontiers in Genetics*, vol. 12, p. 1501, 2021/08/13 2021.
- [257] S. Deng, J. Wang, L. Hou, J. Li, G. Chen, B. Jing, *et al.*, "Annexin A1, A2, A4 and A5 play important roles in breast cancer, pancreatic cancer and laryngeal carcinoma, alone and/or synergistically," *Oncology Letters*, vol. 5, pp. 107-112, 2012/01/01 2012.
- [258] W. M. Serag and B. E. Elsayed, "Annexin A5 as a marker for hepatocellular carcinoma in cirrhotic hepatitis C virus patients," *Egyptian Liver Journal*, vol. 11, pp. 1-5, 2021/12/01 2021.
- [259] B. Sun, Y. Bai, L. Zhang, L. Gong, X. Qi, H. Li, *et al.*, "Quantitative Proteomic Profiling the Molecular Signatures of Annexin A5 in Lung Squamous Carcinoma Cells," *PLOS ONE*, vol. 11, p. e0163622, 2016/09/01 2016.
- [260] B. Peng, C. Guo, H. Guan, S. Liu, and M. Z. Sun, "Annexin A5 as a potential marker in tumors," *Clinica Chimica Acta*, vol. 427, pp. 42-48, 2014/01/01 2014.
- [261] P. N. Pozo and J. G. Cook, "Regulation and Function of Cdt1; A Key Factor in Cell Proliferation and Genome Stability," *Genes*, vol. 8, p. 2, 2017/12/22 2017.
- [262] A. Kanellou, N. N. Giakoumakis, A. Panagopoulos, S. C. Tsaniras, and Z. Lygerou, "The Licensing Factor Cdt1 Links Cell Cycle Progression to the DNA Damage Response," *Anticancer Research*, vol. 40, pp. 2449-2456, 2020/05/01 2020.

- [263] R. Mahadevappa, H. Neves, S. M. Yuen, Y. Bai, C. M. McCrudden, H. F. Yuen, *et al.*, "The prognostic significance of Cdc6 and Cdt1 in breast cancer," *Scientific Reports*, vol. 7, 2017/12/01 2017.
- [264] C. Cai, Y. Zhang, X. Hu, W. Hu, S. Yang, H. Qiu, *et al.*, "CDT1 Is a Novel Prognostic and Predictive Biomarkers for Hepatocellular Carcinoma," *Frontiers in Oncology*, vol. 11, p. 3803, 2021/09/24 2021.
- [265] V. Bravou, H. Nishitani, S. Y. Song, S. Taraviras, and J. Varakis, "Expression of the licensing factors, Cdt1 and Geminin, in human colon cancer," *International Journal of Oncology*, vol. 27, pp. 1511-1518, 2005/12/01 2005.
- [266] C. Wang, J. Che, Y. Jiang, P. Chen, G. Bao, and C. Li, "CDT1 facilitates metastasis in prostate cancer and correlates with cell cycle regulation," *Cancer biomarkers : section A of Disease markers*, pp. 1-11, 2022/02/24 2022.
- [267] H. Kuivaniemi and G. Tromp, "Type III collagen (COL3A1): Gene and protein structure, tissue distribution, and associated diseases," *Gene*, vol. 707, p. 151, 2019/07/30 2019.
- [268] S. Shi and B. Tian, "Identification of biomarkers associated with progression and prognosis in bladder cancer via co-expression analysis," *Cancer biomarkers : section A of Disease markers*, vol. 24, pp. 183-193, 2019 2019.
- [269] H. Zhang, G. Shan, J. Song, Y. Tian, L. Y. An, Y. Ban, *et al.*, "Extracellular matrix-related genes play an important role in the progression of NMIBC to MIBC: a bioinformatics analysis study," *Bioscience reports*, vol. 40, 2020/05/01 2020.
- [270] J. Y. Lee, S. J. Yun, P. Jeong, X. M. Piao, Y. H. Kim, J. Kim, *et al.*, "Identification of differentially expressed miRNAs and miRNA-targeted genes in bladder cancer," *Oncotarget*, vol. 9, p. 27656, 2018/06/12 2018.
- [271] L. Yuan, B. Shu, L. Chen, K. Qian, Y. Wang, G. Qian, *et al.*, "Overexpression of COL3A1 confers a poor prognosis in human bladder cancer identified by co-expression analysis," *Oncotarget*, vol. 8, pp. 70508-70520, 2017 2017.
- [272] J. A. Ewald, T. M. Downs, J. P. Cetnar, and W. A. Ricke, "Expression microarray meta-analysis identifies genes associated with Ras/MAPK and related pathways in progression of muscle-invasive bladder transition cell carcinoma," *PLOS ONE*, vol. 8, 2013/02/01 2013.
- [273] H. Zhu, H. Chen, J. Wang, L. Zhou, and S. Liu, "Collagen stiffness promoted non-muscle-invasive bladder cancer progression to muscle-invasive bladder cancer," *OncoTargets and therapy*, vol. 12, pp. 3441-3457, 2019 2019.
- [274] X. M. Piao, B. Hwang, P. Jeong, Y. J. Byun, H. W. Kang, S. P. Seo, *et al.*, "Collagen type VI- α 1 and 2 repress the proliferation, migration and invasion of bladder cancer cells," *International journal of oncology*, vol. 59, 2021/07/01 2021.
- [275] K. Davalieva, S. Kiprijanovska, I. M. Kostovska, S. Stavridis, O. Stankov, S. Komina, *et al.*, "Comparative Proteomics Analysis of Urine Reveals Down-Regulation of Acute Phase Response Signaling and LXR/RXR Activation Pathways in Prostate Cancer," *Proteomes*, vol. 6, p. 1, 2017/12/29 2017.
- [276] S. Xu, H. Xu, W. Wang, S. Li, H. Li, T. Li, *et al.*, "The role of collagen in cancer: from bench to bedside," *Journal of Translational Medicine*, vol. 17, pp. 1-22, 2019/09/14 2019.
- [277] A. S. Lamort, I. Giopanou, I. Psallidas, and G. T. Stathopoulos, "Osteopontin as a Link between Inflammation and Cancer: The Thorax in the Spotlight," *Cells*, vol. 8, p. 815, 2019/08/02 2019.
- [278] A. Zaravinos, G. I. Lambrou, D. Volanis, D. Delakas, and D. A. Spandidos, "Spotlight on Differentially Expressed Genes in Urinary Bladder Cancer," *PLoS ONE*, vol. 6, p. e18255, 2011 2011.

- [279] G. Pignot, A. Vieillefond, S. Vacher, M. Zerbib, B. Debre, R. Lidereau, *et al.*, "Hedgehog pathway activation in human transitional cell carcinoma of the bladder," *British journal of cancer*, vol. 106, pp. 1177-1186, 2012/03/13 2012.
- [280] A. Zaravinos, D. Volanis, G. I. Lambrou, D. Delakas, and D. A. Spandidos, "Role of the angiogenic components, VEGFA, FGF2, OPN and RHOC, in urothelial cell carcinoma of the urinary bladder," *Oncology reports*, vol. 28, pp. 1159-1166, 2012/10 2012.
- [281] H. Ghasemi, S. H. Mousavibahar, M. Hashemnia, J. Karimi, I. Khodadadi, and H. Tavalani, "Transitional cell carcinoma matrix stiffness regulates the osteopontin and YAP expression in recurrent patients," *Molecular biology reports*, vol. 48, pp. 4253-4262, 2021/05/01 2021.
- [282] Y. Tu, C. Chen, and G. Fan, "Association between the expression of secreted phosphoprotein - related genes and prognosis of human cancer," *BMC cancer*, vol. 19, 2019/12/18 2019.
- [283] T. Wei, G. Bi, Y. Bian, S. Ruan, G. Yuan, H. Xie, *et al.*, "The Significance of Secreted Phosphoprotein 1 in Multiple Human Cancers," *Frontiers in Molecular Biosciences*, vol. 7, p. 251, 2020/11/24 2020.
- [284] H. Tang, J. Chen, X. Han, Y. Feng, and F. Wang, "Upregulation of SPP1 Is a Marker for Poor Lung Cancer Prognosis and Contributes to Cancer Progression and Cisplatin Resistance," *Frontiers in cell and developmental biology*, vol. 9, 2021/04/29 2021.
- [285] J. Qian, B. L. LeSavage, K. M. Hubka, C. Ma, S. Natarajan, J. T. Eggold, *et al.*, "Cancer-associated mesothelial cells promote ovarian cancer chemoresistance through paracrine osteopontin signaling," *The Journal of clinical investigation*, vol. 131, 2021/08/16 2021.
- [286] Y. Li, S. He, A. He, B. Guan, G. Ge, Y. Zhan, *et al.*, "Identification of plasma secreted phosphoprotein 1 as a novel biomarker for upper tract urothelial carcinomas," *Biomedicine & Pharmacotherapy*, vol. 113, p. 108744, 2019/05/01 2019.
- [287] N. Ferrara and A. P. Adamis, "Ten years of anti-vascular endothelial growth factor therapy," *Nature Reviews Drug Discovery*, vol. 15, pp. 385-403, 2016/01/18 2016.
- [288] Z. Huang, M. Zhang, G. Chen, W. Wang, P. Zhang, Y. Yue, *et al.*, "Bladder cancer cells interact with vascular endothelial cells triggering EGFR signals to promote tumor progression," *International journal of oncology*, vol. 54, pp. 1555-1566, 2019/04/01 2019.
- [289] W. Cao, Y. Zhao, L. Wang, and X. Huang, "Circ0001429 regulates progression of bladder cancer through binding miR-205-5p and promoting VEGFA expression," *Cancer biomarkers : section A of Disease markers*, vol. 25, pp. 101-113, 2019 2019.
- [290] G. Pignot, I. Bieche, S. Vacher, C. Güet, A. Vieillefond, B. Debré, *et al.*, "Large-scale real-time reverse transcription-PCR approach of angiogenic pathways in human transitional cell carcinoma of the bladder: identification of VEGFA as a major independent prognostic marker," *European urology*, vol. 56, pp. 678-689, 2009/10 2009.
- [291] Y. J. Huang, W. X. Qi, A. N. He, Y. J. Sun, Z. Shen, and Y. Yao, "Prognostic value of tissue vascular endothelial growth factor expression in bladder cancer: a meta-analysis," *Asian Pacific journal of cancer prevention : APJCP*, vol. 14, pp. 645-649, 2013 2013.
- [292] M. Sankhwar, S. N. Sankhwar, A. Abhishek, and S. Rajender, "Clinical significance of the VEGF level in urinary bladder carcinoma," *Cancer biomarkers : section A of Disease markers*, vol. 15, pp. 349-355, 2015/06/02 2015.
- [293] M. De Paoli, P. Perco, I. Mühlberger, A. Lukas, H. Pandha, R. Morgan, *et al.*, "Disease map-based biomarker selection and pre-validation for bladder cancer diagnostic," *Biomarkers: biochemical indicators of exposure, response, and susceptibility to chemicals*, vol. 20, pp. 328-337, 2015/07/04 2015.

- [294] Y. Hirasawa, I. Pagano, R. Chen, Y. Sun, Y. Dai, A. Gupta, *et al.*, "Diagnostic performance of Oncuria™, a urinalysis test for bladder cancer," *Journal of translational medicine*, vol. 19, 2021/12/01 2021.
- [295] Y. Bu, L. Shi, D. Yu, Z. Liang, and W. Li, "CDCA8 is a key mediator of estrogen-stimulated cell proliferation in breast cancer cells," *Gene*, vol. 703, pp. 1-6, 2019/06/30 2019.
- [296] Y. Shuai, E. Fan, Q. Zhong, Q. Chen, G. Feng, X. Gou, *et al.*, "CDCA8 as an independent predictor for a poor prognosis in liver cancer," *Cancer cell international*, vol. 21, 2021/12/01 2021.
- [297] P. Gu, D. Yang, J. Zhu, M. Zhang, X. He, and K. Mojumdar, "Bioinformatics analysis of the clinical relevance of CDCA gene family in prostate cancer," *Medicine*, vol. 101, p. E28788, 2022/02/04 2022.
- [298] X. Gao, X. Wen, H. He, L. Zheng, Y. Yang, J. Yang, *et al.*, "Knockdown of CDCA8 inhibits the proliferation and enhances the apoptosis of bladder cancer cells," *PeerJ*, vol. 8, 2020 2020.
- [299] Y. Bi, S. Chen, J. Jiang, J. Yao, G. Wang, Q. Zhou, *et al.*, "CDCA8 expression and its clinical relevance in patients with bladder cancer," *Medicine*, vol. 97, 2018/08/01 2018.
- [300] S. Pan, Y. Zhan, X. Chen, B. Wu, and B. Liu, "Identification of Biomarkers for Controlling Cancer Stem Cell Characteristics in Bladder Cancer by Network Analysis of Transcriptome Data Stemness Indices," *Frontiers in Oncology*, vol. 9, pp. 1-11, 2019 2019.
- [301] K. Chen, J. Xing, W. Yu, Y. Xia, Y. Zhang, F. Cheng, *et al.*, "Identification and Validation of Hub Genes Associated with Bladder Cancer by Integrated Bioinformatics and Experimental Assays," *Frontiers in oncology*, vol. 11, 2021/12/20 2021.
- [302] H. C. Lu, J. Q. Yao, X. Yang, J. Han, J. Z. Wang, K. Xu, *et al.*, "Identification of a potentially functional circRNA-miRNA-mRNA regulatory network for investigating pathogenesis and providing possible biomarkers of bladder cancer," *Cancer Cell International*, vol. 20, pp. 1-15, 2020/01/29 2020.
- [303] P. Shen, X. He, L. Lan, Y. Hong, and M. Lin, "Identification of cell division cycle 20 as a candidate biomarker and potential therapeutic target in bladder cancer using bioinformatics analysis," *Bioscience reports*, vol. 40, 2020/07/01 2020.
- [304] S. Li, X. Liu, T. Liu, X. Meng, X. Yin, C. Fang, *et al.*, "Identification of Biomarkers Correlated with the TNM Staging and Overall Survival of Patients with Bladder Cancer," *Frontiers in physiology*, vol. 8, 2017/11/28 2017.
- [305] J. Shi, P. Zhang, L. Liu, X. Min, and Y. Xiao, "Weighted gene coexpression network analysis identifies a new biomarker of CENPF for prediction disease prognosis and progression in nonmuscle invasive bladder cancer," *Molecular genetics & genomic medicine*, vol. 7, 2019/11/01 2019.
- [306] A. Taber, E. Christensen, P. Lamy, I. Nordentoft, F. Prip, S. V. Lindskrog, *et al.*, "Molecular correlates of cisplatin-based chemotherapy response in muscle invasive bladder cancer by integrated multi-omics analysis," *Nature Communications*, vol. 11, 2020/12/01 2020.
- [307] Z. Liu, Q. Zhou, Z. Wang, H. Zhang, H. Zeng, Q. Huang, *et al.*, "Intratumoral TIGIT+ CD8+ T-cell infiltration determines poor prognosis and immune evasion in patients with muscle-invasive bladder cancer," *Journal for ImmunoTherapy of Cancer*, vol. 8, p. e000978, 2020/08/01 2020.
- [308] D. Jiang, Y. Li, J. Cao, L. Sheng, X. Zhu, and M. Xu, "Cell Division Cycle-Associated Genes Are Potential Immune Regulators in Nasopharyngeal Carcinoma," *Frontiers in Oncology*, vol. 12, p. 84, 2022/02/14 2022.

- [309] E. M. Dunleavy, D. Roche, H. Tagami, N. Lacoste, D. Ray-Gallet, Y. Nakamura, *et al.*, "HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres," *Cell*, vol. 137, pp. 485-497, 2009/05/01 2009.
- [310] R. Cao, G. Wang, K. Qian, L. Chen, G. Qian, C. Xie, *et al.*, "Silencing of HJURP induces dysregulation of cell cycle and ROS metabolism in bladder cancer cells via PPAR γ -SIRT1 feedback loop," *Journal of Cancer*, vol. 8, pp. 2282-2295, 2017 2017.
- [311] C.-j. Wang, X. Li, P. Shi, H. y. Ding, Y. p. Liu, T. Li, *et al.*, "Holliday junction recognition protein promotes pancreatic cancer growth and metastasis via modulation of the MDM2/p53 signaling," *Cell Death & Disease*, vol. 11, pp. 1-14, 2020/05/21 2020.
- [312] Y. Wei, G. L. Ouyang, W. X. Yao, Y. J. Zhu, X. Li, L. X. Huang, *et al.*, "Knockdown of HJURP inhibits non-small cell lung cancer cell proliferation, migration, and invasion by repressing Wnt/ β -catenin signaling," *European review for medical and pharmacological sciences*, vol. 23, pp. 3847-3856, 2019 2019.
- [313] Z. Hu, G. Huang, A. Sadanandam, S. Gu, M. E. Lenburg, M. Pai, *et al.*, "The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer," *Breast cancer research : BCR*, vol. 12, 2010/03/08 2010.
- [314] W. Lai, W. Zhu, C. Xiao, X. Li, Y. Wang, Y. Han, *et al.*, "HJURP promotes proliferation in prostate cancer cells through increasing CDKN1A degradation via the GSK3 β /JNK signaling pathway," *Cell Death & Disease*, vol. 12, pp. 1-12, 2021/06/07 2021.
- [315] F. Zhang, D. B. Yuan, J. K. Song, W. M. Chen, W. Wang, G. H. Zhu, *et al.*, "HJURP is a prognostic biomarker for clear cell renal cell carcinoma and is linked to immune infiltration," *International Immunopharmacology*, vol. 99, p. 107899, 2021/10/01 2021.
- [316] R. Su, H. Huang, X. Gao, Y. Zhou, S. Yin, H. Xie, *et al.*, "A pan-cancer analysis of the oncogenic role of Holliday junction recognition protein in human tumors," *Open Medicine*, vol. 17, pp. 317-328, 2022/02/16 2022.
- [317] S. Zeng, A. Liu, L. Dai, X. Yu, Z. Zhang, Q. Xiong, *et al.*, "Prognostic value of TOP2A in bladder urothelial carcinoma and potential molecular mechanisms," *BMC cancer*, vol. 19, 2019/06/19 2019.
- [318] F. Zhang and H. Wu, "MiR-599 targeting TOP2A inhibits the malignancy of bladder cancer cells," *Biochemical and biophysical research communications*, vol. 570, pp. 154-161, 2021/09/17 2021.
- [319] W. T. Kim, Y. H. Kim, P. Jeong, S. P. Seo, H. W. Kang, Y. J. Kim, *et al.*, "Urinary cell-free nucleic acid IQGAP3: a new non-invasive diagnostic marker for bladder cancer," *Oncotarget*, vol. 9, pp. 14354-14365, 2018 2018.
- [320] M. Lindén, U. Segersten, M. Runeson, K. Wester, C. Busch, U. Pettersson, *et al.*, "Tumour expression of bladder cancer-associated urinary proteins," *BJU international*, vol. 112, pp. 407-415, 2013/08 2013.
- [321] G. Botti, M. G. Malzone, E. La Mantia, M. Montanari, D. Vanacore, S. Rossetti, *et al.*, "ProEx C as Diagnostic Marker for Detection of Urothelial Carcinoma in Urinary Samples: A Review," *International Journal of Medical Sciences*, vol. 14, p. 554, 2017/04/21 2017.
- [322] O. Del Moral-Hernández, D. Hernández-Sotelo, L. d. C. Alarcón-Romero, M. A. Mendoza-Catalán, E. Flores-Alfaro, Y. Castro-Coronel, *et al.*, "TOP2A/MCM2, p16INK4a, and cyclin E1 expression in liquid-based cytology: a biomarkers panel for progression risk of cervical premalignant lesions," *BMC Cancer*, vol. 21, pp. 1-13, 2021/12/01 2021.
- [323] J. Li, P. Sun, T. Huang, S. He, L. Li, and G. Xue, "Individualized chemotherapy guided by the expression of ERCC1, RRM1, TUBB3, TYMS and TOP2A genes versus classic chemotherapy

- in the treatment of breast cancer: A comparative effectiveness study," *Oncology letters*, vol. 21, 2021/01/01 2021.
- [324] L. M. Berclaz, A. Altendorf-Hofmann, H. R. Dürr, A. Klein, M. K. Angele, M. Albertsmeier, *et al.*, "Expression Patterns of TOP2A and SIRT1 Are Predictive of Survival in Patients with High-Risk Soft Tissue Sarcomas Treated with a Neoadjuvant Anthracycline-Based Chemotherapy," *Cancers*, vol. 13, 2021/10/01 2021.
 - [325] X. Yin, Z. Wang, J. Wang, Y. Xu, W. Kong, and J. Zhang, "Development of a novel gene signature to predict prognosis and response to PD-1 blockade in clear cell renal cell carcinoma," *Oncoimmunology*, vol. 10, 2021 2021.
 - [326] M. Sarafidis, G. I. Lambrou, V. Zoumpourlis, and D. Koutsouris, "An Integrated Bioinformatics Analysis towards the Identification of Diagnostic, Prognostic, and Predictive Key Biomarkers for Urinary Bladder Cancer," *Cancers*, vol. 14, p. 3358, 2022.

Λίστα δημοσιεύσεων

Περιοδικά

1. **M. Sarafidis**, G. I. Lambrou, V. Zoumpourlis, and D. Koutsouris, "An Integrated Bioinformatics Analysis towards the Identification of Diagnostic, Prognostic, and Predictive Key Biomarkers for Urinary Bladder Cancer," *Cancers*, vol. 14, p. 3358, 2022.
2. M. Tarousi, **M. Sarafidis**, T. Androutsou, O. Manta, O. Giannakopoulou, G. Koutsouri, et al., "Serious games for Parkinson's Disease management as implemented in PROPHETIC platform," *Health Informatics Journal*, vol. 27, p. 14604582211011231, 2021.
3. K. Koklonis, **M. Sarafidis**, M. Vastardi, S. Philippakis, and D. Koutsouris, "Managing the Intervention Costs of Musculoskeletal Disorders in the Hospital Workplace," *IISE Transactions on Occupational Ergonomics and Human Factors*, vol. 9, pp. 33-48, 2021.
4. K. Koklonis, **M. Sarafidis**, M. Vastardi, and D. Koutsouris, "Utilization of Machine Learning in Supporting Occupational Safety and Health Decisions in Hospital Workplace," *Engineering, Technology & Applied Science Research*, vol. 11, pp. 7262-7272, 2021.
5. **M. Sarafidis**, P. Toumpaniaris, I. Kouris, N. Karanasiou, K. Moirogiorgou, M. Kalla, et al., "eMass: A Smart Digital Tool for Efficient Management and Medical Assessment of The New Recruits In Greek Armed Forces," *Research in Medical & Engineering Sciences*, vol. 9, pp. 985-987, 2021.
6. P. Vartholomaïos, N. Ramdani, C. Christophorou, ..., **M. Sarafidis**, et al., "ENDORSE Concept: An Integrated Indoor Mobile Robotic System for Medical Diagnostic Support," *International Journal of Reliable and Quality E-Healthcare (IJRQEH)*, vol. 8, pp. 47-59, 2019.
7. M. Tarousi, **M. Sarafidis**, P. Katrakazas, I. Kouris, S. Pitoglou, O. Petropoulou, et al., "A Cloud Based Decision Support System Aimed to Contribute in Policy Making for Natural Disaster Related Incidents," *International Journal of Reliable and Quality E-Healthcare (IJRQEH)*, vol. 8, pp. 36-46, 2019.
8. W. Schlee, S. Schoisswohl, ..., **M. Sarafidis**, et al., "Towards a unification of treatments and interventions for tinnitus patients: The EU research and innovation action UNITI," *Progress in brain research*, vol. 260, pp. 441-451, 2021.
9. W. Schlee, B. Langguth, R. Pryss, ..., **M. Sarafidis**, et al., "Using big data to develop a clinical decision support system for tinnitus treatment," in *The Behavioral Neuroscience of Tinnitus*, ed: Springer, Cham, 2021, pp. 175-189.

10. Y. Dionyssiotis, E. Vellidou, S. T. Konstantinidis, P. Sarafis, S. Artemi, ..., **M. Sarafidis**, et al., "Education Program for Carers in Facilities with Neuro Disabled Subjects EPoCFiNDS," *Journal of Frailty, Sarcopenia and Falls*, vol. 4, p. 45, 2019.

Συνέδρια

1. **M. Sarafidis**, G. I. Lambrou, G. Matsopoulos, D. Koutsouris, "Integrative Bioinformatics Analysis of Transcriptomic Data Reveals Hub Genes as Diagnostic Biomarkers for Non-Muscle vs. Muscle Invasive Bladder Cancer," 2022 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), 2022 (under publication).
2. **M. Sarafidis**, A. Zaravinos, D. Iliopoulou, D. Koutsouris, and G. I. Lambrou, "Regressions of Clustered Gene Expression Data Manifest Tumor-Specific Genes in Urinary Bladder Cancer," in 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019, pp. 127-131.
3. **M. Sarafidis**, M. Tarousi, A. Anastasiou, S. Pitoglou, E. Lampoukas, A. Spetsariasis, et al., "Data Quality Challenges in a Learning Health System," *Digital Personalized Health and Medicine: Proceedings of MIE 2020*, vol. 270, p. 143, 2020.
4. **M. Sarafidis**, O. Manta, I. Kouris, W. Schlee, D. Kikidis, E. Vellidou, et al., "Why a Clinical Decision Support System is needed for Tinnitus?," in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 2075-2078.
5. O. Manta, **M. Sarafidis**, W. Schlee, C. Consoulas, D. Kikidis, and D. Koutsouris, "Electrophysiological differences in distinct hearing threshold level individuals with and without tinnitus distress," in 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022 (under publication).
6. K. Bromis, **M. Sarafidis**, O. Manta, I. Kouris, E. Vellidou, W. Schlee, and D. Koutsouris, "Predicting the optimal therapeutic intervention for tinnitus patients using random forest regression: A preliminary study of UNITI's decision support system model," in 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, (under publication).
7. M. Braoudaki, **M. Sarafidis**, D. D. Koutsouris, G. Koutsouri, and G. I. Lambrou, "Bioinformatics Analysis Reveals Ki-67 Specific microRNA Functions in Pediatric Embryonal Tumors," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 1346-1349.
8. G. I. Lambrou, **M. Sarafidis**, P. Bizopoulos, D. Iliopoulou, A. Paidi, and D. D. Koutsouris, "The Prisoner's Dilemma" in the Tumor Biology Context: The Warburg Effect, Proliferation Dynamics and Equilibria," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 1342-1345.

9. V. Costarides, T. Guilcher, M. Rebiaï, C. Christophorou, ..., **M. Sarafidis**, et al., "The ENDORSE concept, an integrated and mobile robotic solution, targeted at Medical Diagnostic support," in Journées d'Etude sur la Télésanté, 2019.
10. P. Toumpaniaris, P. Katrakazas, G. I. Lambrou, A. Papanikitas, **M. Sarafidis**, S. Pavlopoulos, et al., "Design and Implementation of a Precision Oxygen Delivery Control System in Facial Masks," in 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019, pp. 894-898.
11. K. Kardaras, N. Apostolou, G. I. Lambrou, **M. Sarafidis**, and D. Koutsouris, "Development and Evaluation of Advanced Image Analysis Techniques for Pediatric Deep Vein Thrombosis Imaging Scans," in 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019, pp. 296-300.
12. I. Kouris, **M. Sarafidis**, T. Androutsou, and D. Koutsouris, "HOLOBALANCE: an augmented reality virtual trainer solution for balance training and fall prevention," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 4233-4236.
13. M. Tarousi, **M. Sarafidis**, A. Anastasiou, I. Kouris, O. Petropoulou, and D. Koutsouris, "A cloud computing platform for the holistic approach to the needs of children and their families in natural disasters incidents," in 2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2018, pp. 9-15.
14. P. Ioannidou, P. Katrakazas, S. Kollias, **M. Sarafidis**, and D. Koutsouris, "HEAR? INFO: A Modern Mobile-Web Platform Addressed to Hard-of-Hearing Elderly Individuals," in 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), 2017, pp. 570-575.

Ελληνικά συνέδρια

1. M. Georgiopoulos, G. Kotsidis, G. Ladakis, G. Stromatias, **M. Sarafidis**, D. Koutsouris and O. Petropoulou, "Review of user acceptance factors of health monitoring wearable devices," in 8th Panhellenic Conference on Biomedical Technology, 2019.
2. **M. Sarafidis**, O. Petropoulou, A. Anastasiou, S. Kepetzi, D. Koutsouris, "Application of the Think – Pair – Share (TPS) Teaching Strategy in Big Data course: An Innovative Collaborative Learning Approach for Students to Activate Data-Driven Analytical Thinking," in 2nd Panhellenic Conference on Technology, Economics & Management, 2018, pp. 21.
3. **M. Sarafidis**, D. Koutsouris, "PROPHETIC: An innovating Personal Healthcare Service for a holistic remote management and treatment of Parkinson patients," in 24th Scientific Congress of Hellenic Medical Students (SCHMS), 2018.

Κεφάλαια σε βιβλία

1. **M. Sarafidis**, O. Petropoulou, D. Iliopoulou, and D. Koutsouris, "Biomedical engineering contribution to modern surgery" in I. Bramis, "History of Greek

Surgery throughout the centuries," 3rd Volume, pp. 225-253, Athens, 2019, Militos Publications.