



ΕΘΝΙΚΟ ΙΔΡΥΜΑ ΕΡΕΥΝΩΝ
National Hellenic Research Foundation

Προς: Ίδρυμα Κρατικών Υποτροφιών

Θεοδώρα Κατσίλα
Εντεταλμένη Ερευνήτρια
Εθνικό Ίδρυμα Ερευνών

☎ : +30.210.7273752

☎ : +30 210 727 3831

✉ thkatsila@eie.gr

http://www.eie.gr/nhrf/institutes/ibrb/programmes/BiomarkerDiscovery_TranslationalResearch_en.html

ΒΕΒΑΙΩΣΗ ΠΕΡΑΤΩΣΗΣ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ ΕΡΕΥΝΑΣ

Αγαπητή(έ) Κυρία(ε),

Βεβαιώνω ότι η Δρ. Αγγελική Μπαλασοπούλου περάτωσε τη μεταδιδακτορική της έρευνα, με τίτλο «Ο ρόλος της Th1 κυτταρικής ανοσίας και των ιικών ερεθισμάτων στην κοιλιοκάκη και τη συννοσηρότητα με αιματολογικές κακοήθειες», στις 31/01/2022 σύμφωνα με την εγκεκριμένη πρόταση με αριθμό πρωτοκόλλου πρότασης 17600.

Η έρευνα αυτή διενεργήθηκε στο **Εργαστήριο Ανακάλυψης Βιοδεικτών και Μεταφραστικής Έρευνας, Ινστιτούτο Χημικής Βιολογίας, Εθνικό Ίδρυμα Ερευνών** και ενέπιπτε στον Επιστημονικό τομέα «ΕΠΙΣΤΗΜΕΣ ΖΩΗΣ» και τον Επιστημονικό Κλάδο «ΙΑΤΡΙΚΗ ΚΑΙ ΕΠΙΣΤΗΜΕΣ ΥΓΕΙΑΣ» με εξειδίκευση στην «Γαστρεντερολογία, Ηπατολογία» σύμφωνα με την πρόσκληση Α.Π. 11047 του Ιδρύματος Κρατικών Υποτροφιών.

Με εκτίμηση

Η επιβλέπουσα καθηγήτρια,

Δρ. Θεοδώρα Κατσίλα

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
«Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση Και Διά Βίου
Μάθηση»
(ΕΣΠΑ 2014-2020)

ΠΡΑΞΗ
«Ενίσχυση Μεταδιδασκτόρων Ερευνητών/Ερευνητριών - Β΄
Κύκλος»(MIS 5033021)

ΤΕΛΙΚΗ ΕΚΘΕΣΗ ΠΕΠΡΑΓΜΕΝΩΝ
ΑΠΟ 01/02/2020 ΕΩΣ 31/01/2022

Α ΕΝΟΤΗΤΑ: ΣΤΟΙΧΕΙΑ ΥΠΟΤΡΟΦΟΥ & ΤΑΥΤΟΤΗΤΑ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ
ΕΡΕΥΝΑΣ

ΟΝΟΜΑΤΕΠΩΝΥΜΟ	Αγγελική Μπαλασοπούλου
ΠΑΤΡΩΝΥΜΟ	Μπαλάσης
ΑΡΙΘΜΟΣ ΣΥΜΒΑΣΗΣ	2019-050-0503-17600
ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΤΟΜΕΑΣ	Επιστήμες Ζωής
ΘΕΜΑ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ ΕΡΕΥΝΑΣ (σύμφωνα με την εγκεκριμένη πρόταση για χρηματοδότηση)	Ο ρόλος της Th1 κυτταρικής ανοσίας και των ικών ερεθισμάτων στη κοιλιοκάκη και τη συννοσηρότητα με αιματολογικές κακοήθειες
ΙΔΡΥΜΑ ΥΛΟΠΟΙΗΣΗΣ ΤΗΣ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ ΕΡΕΥΝΑΣ (αναφέρονται τα πλήρη στοιχεία ΑΕΙ/ Ερ. Κέντρο/Τμήμα/Εργαστήριο)	Εθνικό Ίδρυμα Ερευνών, Ινστιτούτο Χημικής Βιολογίας, Εργαστήριο Ανακάλυψης Βιοδεικτών & Μεταφραστικής Έρευνας

<p>ΗΜ/ΝΙΑ ΕΝΑΡΞΗΣ ΥΠΟΤΡΟΦΙΑΣ(σύμφωνα με την υπογεγραμμένη σύμβαση μεταξύ υποτρόφου και ΙΚΥ)</p>	<p>01/02/2020</p>	<p>ΗΜ/ΝΙΑ ΛΗΞΗΣ ΥΠΟΤΡΟΦΙΑΣ (σύμφωνα με την υπογεγραμμένη σύμβαση μεταξύ υποτρόφου και ΙΚΥ)</p>	<p>31/01/2022</p>
--	-------------------	--	-------------------

<p>ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ (συμπληρώνονται τα πλήρη στοιχεία καθώς και το Τμήμα/ΑΕΙ/Ερ. Κέντρο στο οποίο ανήκει, τηλέφωνο, email)</p>	<p>Δρ. Κασιίλα Θεοδώρα, Εργαστήριο Ανακάλυψης Βιοδεικτών & Μεταφραστικής Έρευνας, Ινστιτούτο Χημικής Βιολογίας, Εθνικό Ίδρυμα Ερευνών, +30 210 72 73 752 (thkatsila@ie.gr)</p>
---	--

B ΕΝΟΤΗΤΑ: ΤΕΛΙΚΗ ΕΚΘΕΣΗ ΠΕΠΡΑΓΜΕΝΩΝ

Εισαγωγή

Οι σύνθετες ασθένειες δε συμμορφώνονται με τα πρότυπα Μενδελικής κληρονομικότητας. Τα σύνθετα χαρακτηριστικά και οι κλινικοί φαινότυποι αυτών προέρχονται από γενετικές (π.χ. γονίδια, γενετικοί τόποι, γονιδιωματικές παραλλαγές) και περιβαλλοντικές επιρροές (π.χ. διαίτα, ιικά βιομόρια/ερεθίσματα), με τις περισσότερες από αυτές τις επιρροές να παραμένουν αδιευκρίνιστες ως προς την ταυτοποίησή τους και το ρόλο τους στην (παθο)βιολογία των σύνθετων ασθενειών. Μια πληθώρα ασθενειών εμπίπτουν στην κατηγορία των σύνθετων ασθενειών, όπως ο καρκίνος, ο σακχαρώδης διαβήτης, η νόσος του Αλτσχάιμερ (Liu ZP, 2012) και πρωτίστως, συνολικά, τα λοιμώδη και αυτοάνοσα νοσήματα (Matzaraki V, 2017).

Οι σύνθετες ασθένειες μπορούν να αποσαφηνιστούν με την εφαρμογή -ομικών προσεγγίσεων και τη βιοπληροφορική (Sun YV, 2016), όπως η τελευταία «καθοδηγείται» από τη βιοστατιστική. Στις μέρες μας, ειδικότερα με την τεχνολογική ανάπτυξη και τις μεθοδολογίες υψηλής απόδοσης, παρέχονται οι βάσεις προς τη βέλτιστη διαφορική διάγνωση, τη διαστρωμάτωση των ασθενών και τη λήψη αποφάσεων βάσει δεδομένων στην κλινική πράξη. Η χρήση, μάλιστα, προσεγγίσεων τεχνητής νοημοσύνης (AI), μοριακής μοντελοποίησης και μηχανικής μάθησης (ML), σε συνδυασμό με την απεριόριστη κλιμακούμενη αποθήκευση και τα μεγέθη των συνόλων δεδομένων (datasets) μπορεί να οδηγήσει στην ανακάλυψη και ανάπτυξη βιοδεικτών (Vamathevan J, 2016; Nayarisseri A, 2021).

Το φλεγμονόσωμα (inflammasome) είναι ένα πρωτεϊνικό σύμπλεγμα, που αποτελεί ορόσημο στον τομέα της φυσικής ανοσίας, καθώς δρα ως ρυθμιστής της ενδοκυτταρικής ανοσοαπόκρισης. Πρόσφατα, έχει αναφερθεί ο ρόλος του φλεγμονοσώματος σε φλεγμονώδεις ασθένειες, αλλά και συννοσηρότητες αυτών (Rathinam VA, 2016; Guo H, 2015), εξαιτίας της προβληματικής αυτορρύθμισής του, που μπορεί να μετατοπίσει τις φλεγμονώδεις αποκρίσεις από το πλαίσιο του αμυντικού μηχανισμού προς μια διαιωνιζόμενη φλεγμονώδη απόκριση (Awad F, 2018). Στη γαστρεντερική οδό, τα φλεγμονοσώματα ενορχηστρώνουν την ανοσολογική ανοχή σε μικροβιακά και σχετιζόμενα με τα τρόφιμα δυνητικά σήματα ή οδηγούν στην έναρξη φλεγμονωδών αποκρίσεων έναντι παθογόνων, που έχουν κατορθώσει να διεισδύσουν στο ξενιστή. Όταν το φλεγμονόσωμα ρυθμίζεται ανεπαρκώς, η ενεργοποίηση του φλεγμονώδους εντέρου οδηγεί σε μια διαιωνιζόμενη φλεγμονώδη απόκριση, με παθολογία του ανοσοποιητικού συστήματος και βλάβη ιστού (Zmora N, 2017). Με την παρουσία της επιβλαβούς αιτίας (μικρόβια, γλουτένη) (Bauernfeind F, 2013; Paloná-Jelínková L, 2013), το φλεγμονόσωμα ενεργοποιεί τις προφλεγμονώδεις κυτοκίνες, που εξέρχονται από το κύτταρο και προκαλούν μια φλεγμονώδη αντίδραση (Chan AH, 2020). Αναφορικά με την κοιλιοκάκη, τον κλινικό φαινότυπο ενδιαφέροντος της παρούσας ερευνητικής στρατηγικής, το NLRP3 είναι μακράν το πιο μελετημένο φλεγμονώδες σώμα (Gómez Castro MF, 2019).

Στην παρούσα μεταδιδακτορική έρευνα, αναπτύχθηκε, βελτιστοποιήθηκε και εφαρμόστηκε μια καινοτόμος προγνωστική προσέγγιση για την αξιολόγηση κινδύνου συννοσηροτήτων, που υπογραμμίζει τις μοριακές υπογραφές της σοβαρότητας της ασθένειας. Ως πρότυπο-μοντέλο ασθένειας χρησιμοποιήθηκε η κοιλιοκάκη. Ειδικότερα, απαντήθηκαν τα ερευνητικά

ερωτήματα της έρευνας (ο ρόλος της Th1 κυτταρικής ανοσίας και των ικών βιομορίων-ερεθισμάτων στη διαφοροδιάγνωση της κοιλιοκάκης στον παιδιατρικό πληθυσμό και στη συννοσηρότητα της κοιλιοκάκης και της αιματολογικής κακοήθειας) και επιτεύχθηκαν οι τρεις κύριοι στόχοι της έρευνας (α. ανακάλυψη και επικύρωση μοριακής υπογραφής δριμείας νόσου, β. βιοδεικτών συννοσηρότητας κοιλιοκάκης και αιματολογικής κακοήθειας και γ. μηχανιστική διερεύνηση της Th1 κυτταρικής ανοσίας και των ικών βιομορίων ερεθισμάτων, υπό το πρίσμα των επιλεγμένων βιοδεικτών).

Μεθοδολογία

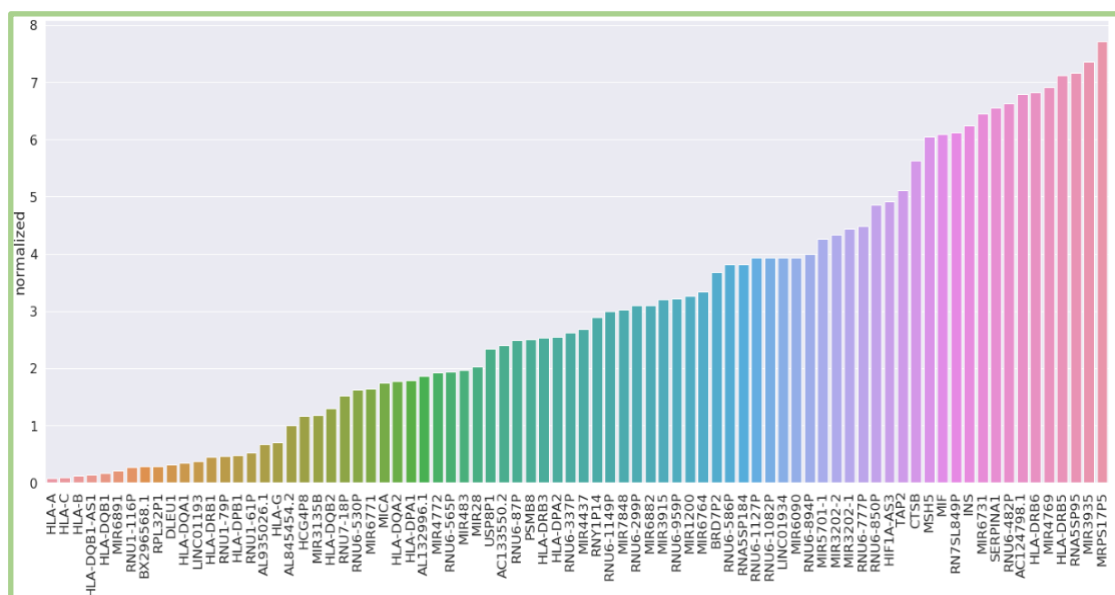
Πραγματοποιήθηκε μια εν τω βάθει *in silico* ανάλυση σε ελεύθερα διαθέσιμα δεδομένα αποθετηρίων και βάσεις δεδομένων (data repositories and databases), σε παγκόσμια κλίμακα με χρήση της γλώσσας προγραμματισμού R και Python. Εφαρμόστηκαν δύο στρατηγικές: η αγνωστική (μη-στοχευμένη) ανάλυση και η στοχευμένη ανάλυση, στη βάση των πλεονεκτημάτων τους, ώστε να έχουμε τη μέγιστη γνωσιακή απόδοση κατά την ενσωμάτωση των δεδομένων. Εν συνεχεία, ακολουθήθηκε εκτεταμένη εξόρυξη δεδομένων και κειμένου προς α. αυστηρή ανάλυση των απορρεόντων ευρημάτων και β. ανάλυση συσχέτισης πρωτεϊνών και μοριακών δικτύων, με χρήση του υπολογιστικού εργαλείου STRING v11.5 (<https://string-db.org/>) για την ανακάλυψη και επικύρωση της μοριακής υπογραφής της δριμείας νόσου της κοιλιοκάκης. Η ανάλυση επικεντρώθηκε σε παρανοηματικές παραλλαγές, καθώς είναι εκείνες που οδηγούν σε ποικίλες πρωτεομορφές ή/ και πρωτεϊνικά μεταλλάγματα, με διαφορετική δομή κι εν δυνάμει, λειτουργία (ή απώλεια της τελευταίας). Τέλος, διεξήχθη εκτεταμένη εξόρυξη δεδομένων και ανασκόπηση της τρέχουσας βιβλιογραφίας και των ελεύθερα διαθέσιμων δεδομένων, με μια ποικιλία υπολογιστικών εργαλείων, όχι μόνο για την ανακάλυψη και επικύρωση υποψήφιων βιοδεικτών συννοσηρότητας κοιλιοκάκης/αιματολογικών κακοηθειών, αλλά και για να εντοπιστούν ποια ικά βιομόρια δύνανται να πυροδοτούν την κοιλιοκάκη μέσω της Th1 απόκρισης στην παιδική ηλικία.

Αποτελέσματα

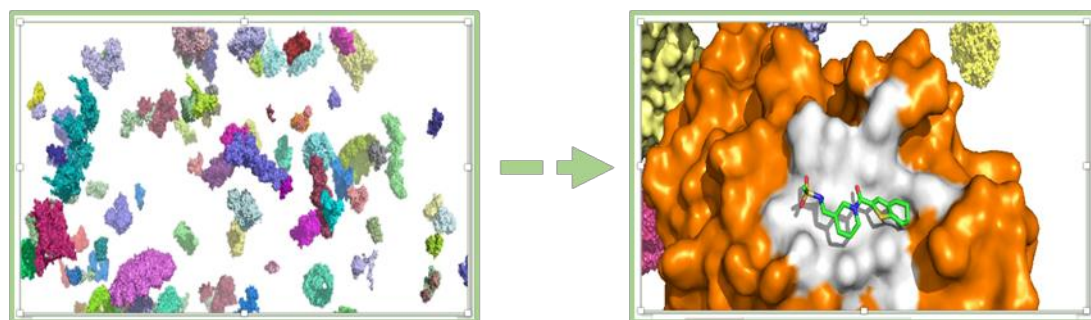
Η ερευνητική προσέγγιση που ακολουθήθηκε είναι καινοτόμος, καθώς αποτελεί τη συνέργεια μεθοδολογιών τεχνητής νοημοσύνης, μοριακής μοντελοποίησης και μηχανικής μάθησης, αλλά και μιας πληθώρας βιοπληροφορικών και βιοστατιστικών εργαλείων προς έλεγχο των συγχυτικών, συστηματικών ή/και σφαλμάτων στη βάση όλων των ελεύθερα διαθέσιμων δεδομένων και κειμένων. Η στρατηγική αυτή, αν και φιλόδοξη, απέδωσε τα μέγιστα, καθώς ανίχνευσε πρότυπα συννοσηρότητας ως τη συνισταμένη γενετικών και περιβαλλοντικών επιρροών (π.χ. δίαιτα, ικά βιομόρια/ερεθίσματα) (Γράφημα 1). Οι εν τω βάθει *in silico* αναλύσεις κατέληξαν σε 1.156 παρανοηματικές παραλλαγές, αναδεικνύοντας πρωτεΐνες μεγίστου ενδιαφέροντος και σημαντικά, τα πρωτεϊνικά τους μεταλλάγματα ή/και τις πρωτεομορφές τους. Ειδικότερα, στην Εικόνα 1 απεικονίζεται η προσομοίωση των εν λόγω πρωτεϊνών ενδιαφέροντος σε τρισδιάστατη μορφή, συνολικά, καθώς και ενδεικτική αλληλεπίδραση προσδέτη-πρωτεΐνης, προς φαρμακευτική στόχευση.

Αξιοσημείωτα, ανακαλύφθηκε και επικυρώθηκε η μοριακή υπογραφή της δριμείας κλινικής κοιλιοκάκης. Ειδικότερα, στον Πίνακα 1 παρουσιάζονται κρίσιμες πληροφορίες των παρανοηματικών παραλλαγών εκείνων των γονιδίων, που υπολογίστηκαν κι εν τέλει,

ταυτοποιήθηκαν να συνιστούν την υποψήφια μοριακή υπογραφή της δριμείας κλινικής εικόνας της νόσου. Αν και η πληροφορία είναι πλούσια, στο παρόν κείμενο καταχωρούνται τα ακόλουθα στοιχεία: ονομασία παρανοηματικής παραλλαγής, θέση, αλληλόμορφα, επίδραση, συχνότητα ελάσσονος αλληλομόρφου. Προκειμένου να διερευνηθούν οι μοριακοί μηχανισμοί της δριμείας κλινικής κοιλιοκάκης, δόθηκε έμφαση σε αλληλεπιδράσεις πρωτεΐνης-πρωτεΐνης και μοριακά δίκτυα, με τη χρήση του υπολογιστικού εργαλείου STRING v11.5 (<https://string-db.org/>). Παρακάτω, παρατίθενται τα σχήματα του απορρέοντος πρωτεϊνικού δικτύου για τις πρωτεΐνες της δριμείας κλινικής εικόνας βάσει δεδομένων (Σχήμα 1) και βάσει εμπιστοσύνης (Σχήμα 2). Επιπλέον, παρατίθενται οι πληροφορίες στατιστικής ανάλυσης του δικτύου αυτού (Πίνακας 2). Συγκεκριμένα, το παραπάνω πρωτεϊνικό δίκτυο αποτελείται από 23 κόμβους (πρωτεΐνες) και 22 αλληλεπιδράσεις μεταξύ αυτών. Ο συντελεστής ομαδοποίησης του δικτύου είναι 1.91, υποδηλώνοντας ένα έντονα συνδεδεμένο δίκτυο και η τιμή «PPI enrichment p-value» είναι ίση προς $5.85e-08$, υποδεικνύοντας ότι στο συγκεκριμένο δίκτυο οι πρωτεΐνες είναι μερικώς, βιολογικά συνδεδεμένες, ως ομάδα.



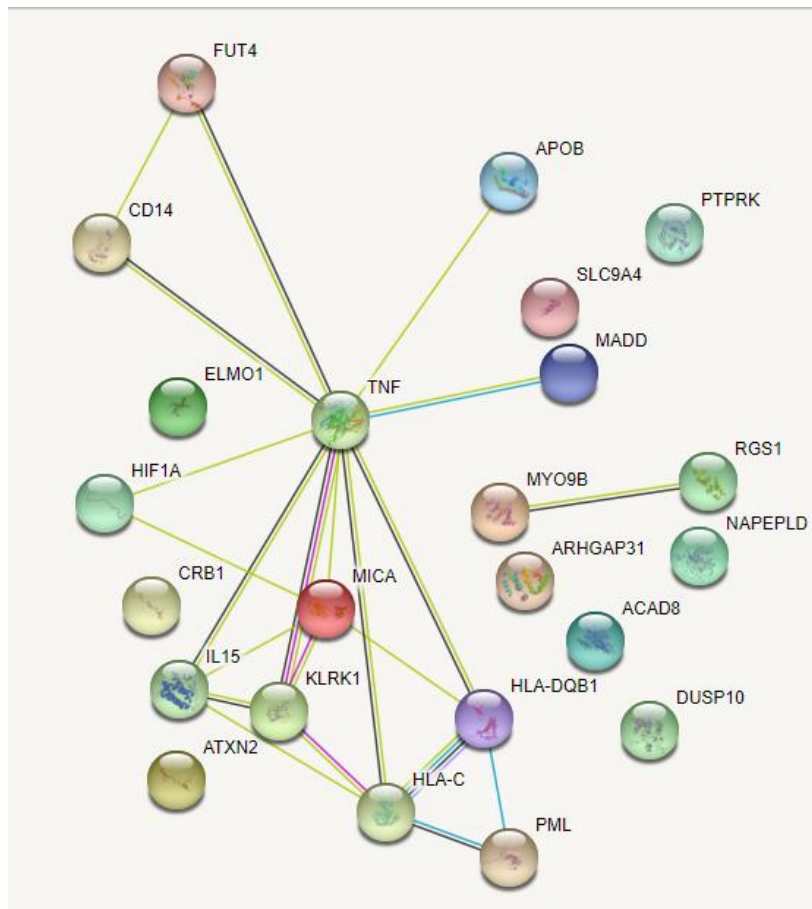
Γράφημα 1. Πρότυπα συννοσηρότητας ως η απορρέουσα συνισταμένη γενετικών και περιβαλλοντικών επιρροών.



Εικόνα 1. Προσομοίωση πρωτεϊνών σε τρισδιάστατη μορφή (αριστερά) και ενδεικτική αλληλεπιδραση προσδέτη-πρωτεΐνης, προς φαρμακευτική στόχευση (δεξιά).

Πίνακας 1. Μοριακή υπογραφή δριμείας κλινικής κοιλιοκάκης, απόρροια μεθοδολογιών τεχνητής νοημοσύνης, μοριακής μοντελοποίησης και μηχανικής μάθησης, αλλά και μιας πληθώρας βιοπληροφορικών και βιοστατιστικών εργαλείων

#Uploaded variation	Location	Allele	IMPACT	SYMBOL	mino acid	genet gene	variant zigosity	mean	MAF	MAF(EUR)	MAF(TSI)
rs1146593	1:75769201-75769201	C	MODERATE	ACADM	T/P	ACAD8	1,684210526	0.12 (A)	0.06 (A)	0.05 (A)	
rs1144566	1:182600491-182600491	A	MODERATE	RGS16	H/L	RGS1	1,973684211	0.01 (T)	0.02 (T)	0.00 (T)	
rs1443549	11:20601610-20601610	A	MODERATE	SLC6A5	A/D	SLC6A4	1,842105263	0.01 (C)	0.00 (C)	0.00 (C)	
rs4268467	11:46681370-46681370	A	MODERATE	ARHGAP1	K/N	ARHGAP31	1,894736842	< 0.01 (T)	0.00 (T)	0.00 (T)	
rs1017594	11:47309565-47309565	A	MODERATE	MADD	L/M	MADD	2	0.01 (T)	0.00 (T)	0.00 (T)	
rs2253849	12:10451071-10451071	A	MODERATE	KLRK1	N/I	KLRK1	1,815789474	0.01 (T)	0.00 (T)	0.00 (T)	
rs582954	12:70534944-70534944	T	MODERATE	PTPRB	D/E	PTPRK	2	0.02 (A)	0.06 (A)	0.09 (A)	
rs4482094	12:80460849-80460849	A	MODERATE	PTPRQ	A/D	PTPRK	2	0.17 (C)	0.12 (C)	0.08 (C)	
rs1169305	12:120999579-120999579	C	MODERATE	HNF1A	S/R	HIF1A	1,815789474	0.01 (A)	0.00 (A)	0.00 (A)	
rs4238526	15:81308654-81308654	T	MODERATE	IL16	L/F	IL15	1,894736842	0.01 (A)	0.00 (A)	0.00 (A)	
rs2471844	16:4890808-4890808	C	MODERATE	PPL	C/W	PML	2	0.07 (A)	0.00 (A)	0.00 (A)	
rs4344749	16:28830990-28830990	T	MODERATE	ATXN2L	Q/H	ATXN2	1,973684211	< 0.01 (A)	0.00 (A)	0.00 (A)	
rs584542	2:21009931-21009931	A	MODERATE	APOB	I/F	APOB	2	0.01 (T)	0.00 (T)	0.00 (T)	
rs1878529	2:102932971-102932971	G	MODERATE	TMEM182	R/G	TMEM187	2	0.03 (A)	0.08 (A)	0.07 (A)	
rs2271767	2:191414601-191414601	A	MODERATE	MYO1B	L/M	MYO9B	2	0.25 (T)	0.04 (T)	0.02 (T)	
rs2257495	20:46367543-46367543	T	MODERATE	ELMO2	D/E	ELMO1	2	0.01 (A)	0.00 (A)	0.00 (A)	
rs2427240	20:61910571-61910571	G	MODERATE	CDH4	D/E	CD14	1,842105263	0.30 (C)	0.21 (C)	0.07 (C)	
rs1352882	3:62243879-62243879	G	MODERATE	PTPRG	I/M	PTPRK	1,973684211	0.02 (C)	0.00 (C)	0.00 (C)	
rs2692696	3:133766289-133766289	C	MODERATE	TF	I/L	TNF	1,973684211	0.01 (A)	0.00 (A)	0.00 (A)	
rs958415	4:125408634-125408634	G	MODERATE	FAT4	D/E	FUT4	1,842105263	0.01 (T)	0.00 (T)	0.01 (T)	
rs2431663	5:172769708-172769708	C	MODERATE	DUSP1	I/M	DUSP10	2	0.15 (G)	0.06 (G)	0.07 (G)	
rs1065076	6:31509904-31509904	G	MODERATE	MICB	T/A	MICA	2	0.20 (A)	0.27 (A)	0.21 (A)	
rs1049086	6:32662127-32662127	T	MODERATE	HLA-DQB	D/E	HLA-DQA1	1,842105263	0.37 (A)	0.40 (A)	0.37 (A)	
rs3181009	7:103103446-103103446	G	MODERATE	NAPEPLD	D/H	NAPEPLD	2	0.14 (C)	0.00 (C)	0.00 (C)	
rs1757095	9:115086115-115086115	A	MODERATE	TNC	Q/L	TNF	2	< 0.01 (T)	0.06 (T)	0.07 (T)	
rs2488602	9:123371268-123371268	A	MODERATE	CRB2	V/E	CRB1	1,973684211	< 0.01 (T)	0.00 (T)	0.00 (T)	
rs1065711	CHR_HSCHR6_MHC_COXG	G	MODERATE	HLA-C	V/A	HLA-C	1,921052632	0.11 (G)	0.12 (G)	0.09 (G)	

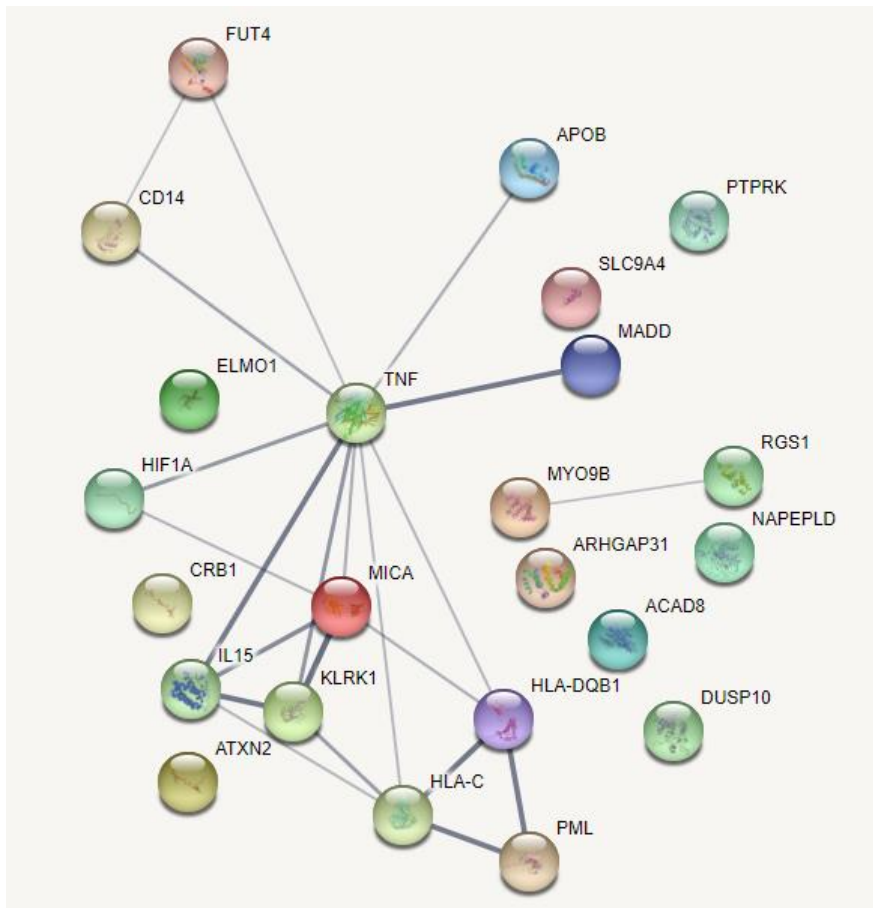


Σχήμα 1. Απεικόνιση του απορρέοντος πρωτεϊνικού δικτύου για τις πρωτεΐνες της δριμείας κλινικής κοιλιοκάκης. Γαλάζιες γραμμές: γνωστές αλληλεπιδράσεις, βάσει δεδομένων. Μωβ γραμμές: γνωστές αλληλεπιδράσεις από πειραματικά δεδομένα. Μαύρες γραμμές: αλληλεπιδράσεις, βάσει δεδομένων κοινής έκφρασης. Κίτρινες γραμμές: αλληλεπιδράσεις, βάσει εξόρυξης κειμένου. Γκρι γραμμές: δεδομένα πρωτεϊνικής ομολογίας.

Πίνακας 2. Στατιστική ανάλυση πρωτεϊνικού δικτύου, STRING version 11.5 (<https://string-db.org/>).

Network Stats	
number of nodes: 23	expected number of edges: 5
number of edges: 22	PPI enrichment p-value: 5.85e-08
average node degree: 1.91	<i>your network has significantly more interactions than expected (what does that mean?)</i>
avg. local clustering coefficient: 0.494	

number of nodes: αριθμός κόμβων στο δίκτυο, number of edges: αριθμός αλληλεπιδράσεων μεταξύ κόμβων, average node degree: πόσες αλληλεπιδράσεις έχει κατά μέσο όρο μία πρωτεΐνη στο απορρέον δίκτυο, clustering coefficient: συντελεστής ομαδοποίησης δικτύου-υψηλές τιμές υποδηλώνουν υψηλή σύνδεση, PPI enrichment p-value: τιμή p-value της αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης



Σχήμα 2. Απεικόνιση του απορρέοντος πρωτεϊνικού δικτύου για τις πρωτεΐνες της δριμείας κλινικής κολιοκάκης, βάσει εμπιστοσύνης. Το πάχος των γραμμών μεταξύ των κόμβων είναι ανάλογο της ισχύος των δεδομένων. Αχνό γκρι χρώμα: χαμηλή τιμή εμπιστοσύνης, 0.15. Γκρι χρώμα: μέση τιμή εμπιστοσύνης, 0.4. Σκούρο γκρι χρώμα: υψηλή τιμή εμπιστοσύνης, 0.7. Έντονο σκούρο γκρι χρώμα: ύψιστη τιμή εμπιστοσύνης, 0.9.

Επιπλέον, ανακαλύφθηκαν και επικυρώθηκαν υποψήφιοι βιοδείκτες συννοσηρότητας κοιλιοκάκης/αιματολογικών κακοηθειών (Πίνακας 3). Τέλος, εντοπίστηκαν ιοί και ιικά βιομόρια αυτών (ενδεικτικά: πρωτεΐνες, γονίδια, γονιδιακές παραλλαγές, κυτταροκίνες), που δύναται να πυροδοτούν την κοιλιοκάκη, μέσω της Th1 απόκρισης στην παιδική ηλικία (Πίνακας 4).

Πίνακας 3. Υποψήφιοι βιοδείκτες συννοσηρότητας κοιλιοκάκης με αιματολογικές κακοήθειες

Genes/Proteins	miRNAs	Hematologic malignancies
IL-15	n.a	T cell large granular lymphocytic leukemia (T-LGL); gD/D T cell lymphoma; enteropathy-associated T cell lymphoma; enteropathy-associated T-cell-lymphoma (EATL)
IL-15R	n.a	T cell large granular lymphocytic leukemia (T-LGL); gD/D T cell lymphoma; enteropathy-associated T cell lymphoma; enteropathy-associated T-cell-lymphoma (EATL)
SH2B3	n.a	chronic myeloid leukemia (CML)
ATXN2	n.a	chronic myeloid leukemia (CML)
STAT5	n.a	T-cell leukemia; Sézary lymphoma; type I EATL
IL2	n.a	T-cell leukemia; Sézary lymphoma; type I EATL
TNF	n.a	T-cell leukemia; Sézary lymphoma; type I EATL
IL-21	n.a	T-cell leukemia; Sézary lymphoma; type I EATL
SMAD3; MDM2; STAT3;	miR-200 and miR-192/215; human chromosome 19 miRNA cluster (C19MC) and miR-17/92	enteropathy-associated T-cell-lymphoma (EATL)
KLRK1	n.a	acute myeloid leukemia (AML)
CD14	n.a	T cell leukemia myeloid leukemia
HLA-DQB1	n.a	enteropathy-associated T-cell-lymphoma (EATL)
TNF	n.a	Non-Hodgkin Lymphoma
HLA-C	n.a	acute leukaemia

n.a (not available) = μη διαθέσιμη πληροφορία

Πίνακας 4. Συσχέτιση μεταξύ κοιλιοκάκης και ιών (ικών ερεθισμάτων)

Proteins	Genes	Variants	Cytokines	Viruses
CD86	M1 (viral gene); M2 (viral gene)	n.a	IL-27; IL-12p40; Type 1 IFNs; IRF1	Reovirus type 1 Lang (T1L)
MMP12; TLR27	n.a	n.a	IL-10; IFN-α	Rotavirus
EBNA2 (viral protein); CD14	RMI2; TAGAP; SH2D1A (SAP); ITK; MAGT1; CD27; CD70; CTPS1; RASGRP1; CORO1A; FOXP3	rs34437200, rs1738074	n.a	Epstein-Barr virus (EBV)
Zonulin	IDO	n.a	n.a	Enterovirus
KLRK1	n.a	n.a	n.a	Herpes Simplex Virus (HSV)
M2.2 (viral protein); G (viral protein); ULBP2; MICA; MICB; KLRK1	n.a	n.a	n.a	Respiratory syncytial virus (RSV)

n.a (not available) = μη διαθέσιμη πληροφορία

Συζήτηση

Η καινοτόμος προγνωστική στρατηγική που περιγράφηκε προηγουμένα αποτέλεσε τον θεμέλιο λίθο της παρούσας μεταδιδακτορικής έρευνας. Η εν λόγω προσέγγιση, κατά την ανάπτυξη, βελτιστοποίηση και εφαρμογή της, έκανε χρήση μεθοδολογιών τεχνητής νοημοσύνης, μοριακής μοντελοποίησης και μηχανικής μάθησης, αλλά και μιας πληθώρας βιοπληροφορικών και βιοστατιστικών εργαλείων.

Ως πρότυπο-μοντέλο ασθένειας χρησιμοποιήθηκε μια πολυγονιδιακή και πολυπαραγοντική, σύνθετη, χρόνια ανοσομεσολαβούμενη εντεροπάθεια, η κοιλιοκάκη. Η εν λόγω επιλογή βασίστηκε στο σύνολο της τρέχουσας, ελεύθερα διαθέσιμης γνώσης, καθώς η κοιλιοκάκη συνοδεύεται από διαγνωστική σύγχυση. Πράγματι, η κοιλιοκάκη «μοιράζεται» παρόμοια συμπτώματα με έτερες αυτοάνοσες ασθένειες, αλλά και γενετικό υπόβαθρο (ενδεικτικά: νόσος ευερέθιστου εντέρου) (Nadhem ON, 2015). Σημαντικά, η κοιλιοκάκη στερείται σήμερα ενός «Disease Activity Index» για την ταξινόμησή της σε ήπια/μέτρια/δριμεία νόσο προς βέλτιστη διαστροφμάτωση των ασθενών, αλλά και τη διαχείριση της νόσου. Εξίσου σημαντικά, η κοιλιοκάκη, ούσα μια χρόνια νόσος, αλλά και στην αδιάγνωστη ή μακροχρόνια μη θεραπευθείσα μορφή της, αποτελεί πρόσφορο έδαφος συννοσηρότητας με αυτοάνοσες ασθένειες και κακοήθεις νεοπλασίες (Fessatou S, 2003; Fracchia M, 2004; Sharp J, 2012; Bhat AS, 2013; Malamut G, 2013). Έως σήμερα, καθίσταται αδύνατη η εκτίμηση κινδύνου της εν λόγω συννοσηρότητας για τους ασθενείς με κοιλιοκάκη, τυπική ή άτυπη.

Αξιοσημείωτα, ανακαλύφθηκε και επικυρώθηκε η μοριακή υπογραφή της δριμείας κλινικής κοιλιοκάκης. Συγκεκριμένα, εντοπίστηκαν παρανοηματικές παραλλαγές (n=27) σε επιλεγμένα γονίδια (n=24) (Πίνακας 1). Από τη διερεύνηση του απορρέοντος πρωτεϊνικού

δικτύου (Σχήματα 1 και 2) και τη στατιστική και βιοπληροφορική μας ανάλυση (Πίνακας 2), υποδηλώνεται ότι στο συγκεκριμένο δίκτυο οι πρωτεΐνες είναι μερικώς, βιολογικά συνδεδεμένες ως ομάδα. Συνεπώς, η μοριακή υπογραφή της δριμείας κλινικής κοιλιοκάκης αποτελεί υποψήφιο μέσο προς τη βέλτιστη διαχείριση της νόσου.

Επιπρόσθετα, η κοιλιοκάκη φάνηκε να σχετίζεται με αυξημένο κίνδυνο εμφάνισης αιματολογικών κακοηθειών, όπως η λευχαιμία των μεγάλων κοκκιδίων Τ-λεμφοκυττάρων (T-cell large granular lymphocytic leukemia, T-LGL), το Τ-λέμφωμα εντεροπαθούς εκδήλωσης (enteropathy-associated T-cell lymphoma, EATL), η χρόνια μυελοειδής λευχαιμία (chronic myeloid leukemia, CML) και η Τ-κυτταρική λευχαιμία (T-cell leukemia) (Πίνακας 3). Η αυξημένη έκκριση IL-15 στο φλεγμονώδες περιβάλλον της κοιλιοκάκης μπορεί να επάγει την ανάπτυξη λεμφωμάτων, αφού ευνοεί την ανάπτυξη των Β- και ΝΚ- κυττάρων, αλλά και τον πολλαπλασιασμό των κακοηθών Β-κυττάρων (Fehniger TA, 2001). Η Th1 κυτταρική ανοσία, λόγω της έκκρισης κυτταροκινών, όπως η IFN- γ και η IL-2, που προάγουν τη διαφοροποίηση των CD8+ κυττάρων προς τα Τ-κυτταροτοξικά κύτταρα, μπορεί να θεωρηθεί κατάλληλη για την απόκριση έναντι παθογόνων και ιών (Kooy-Winkelaar YM, 2017).

Πέραν της γλουτένης, σημαντικός - αν και ακόμη αρκετά ανεξερεύνητος - εμφανίζεται ο ρόλος των ιών και των βιομορίων τους (ικά ερεθίσματα) στην κοιλιοκάκη. Από την ερευνητική μας προσέγγιση προέκυψε η συμμετοχή ιών (και βιομορίων τους) στην ανάπτυξη και εκδήλωση της κοιλιοκάκης. Στον Πίνακα 4, παρουσιάζονται οι ιοί με κρίσιμο ρόλο στην ανάπτυξη της κοιλιοκάκης. Ο ρεοϊός (Reovirus), αν και φαινομενικά είναι ένας αβλαβής ιός, έχει την ικανότητα να διαταράσσει την εντερική ομοιόσταση, να ενεργοποιεί την Th1 ανοσία και να πυροδοτεί την ανάπτυξη της κοιλιοκάκης (Bouziat R, 2017). Στο ίδιο πλαίσιο, ο έτερος ιός της οικογένειας *Reoviridae*, ο ροταϊός (Rotavirus) αποτελεί την πιο κοινή αιτία γαστρεντερικών λοιμώξεων στα παιδιά, που μπορεί να τους προκαλέσει σοβαρή διάρροια και εντερική δυσφορία, με μελέτες να υποδηλώνουν πως συχνές λοιμώξεις από ροταϊό αυξάνουν τον κίνδυνο εμφάνισης κοιλιοκάκης (Stene LC, 2006; Nanayakkara M, 2018). Τόσο ο αναπνευστικός συγκυτιακός ιός, ένας πολύ κοινός ιός της βρεφικής και παιδικής ηλικίας, που προσβάλλει όλα σχεδόν τα παιδιά μέχρι την ηλικία των 2-3 ετών (Tjernberg AR, 2014), όσο και ο εντεροϊός (Kahrs CR, 2019), που αποτελεί τη συνηθέστερη αιτία λοίμωξης του ανώτερου αναπνευστικού συστήματος στα παιδιά, φαίνεται να συνδέεται με την εκδήλωση της κοιλιοκάκης. Η αρχική λοίμωξη με τον ιό Epstein-Barr αφορά, κυρίως, παιδιά και εμφανίζεται με το κλινικό σύνδρομο της λοίμωξης μονοκυττάρων. Τόσο η κοιλιοκάκη, όσο και τα λεμφώματα, έχουν συσχετιστεί με τον ιό Epstein-Barr (Harley JB, 2018). Τέλος, ο ιός του απλού έρπητα μπορεί να οδηγήσει σε δριμεία κλινική κοιλιοκάκη (Chen A, 2016).

Τα αποτελέσματα της παρούσας ερευνητικής μελέτης παρουσιάστηκαν στο 71^ο Πανελλήνιο Συνέδριο της Ελληνικής Εταιρείας Βιοχημείας και Μοριακής Βιολογίας και δημοσιεύτηκαν στα πρακτικά του συνεδρίου.

Εν κατακλείδι, μια προγνωστική αξιολόγηση του κινδύνου ανάπτυξης συννοσηρότητας μπορεί να εντοπίσει αυτούς τους ασθενείς υψηλού κινδύνου και επομένως, να ενισχύσει τη βέλτιστη διαστρωμάτωση των ασθενών μέσω μοριακών υπογραφών με επίκεντρο τη φλεγμονή, χαρτογραφώντας, παράλληλα, τους μοριακούς μηχανισμούς της νόσου υπό μελέτη. Σημειώνεται πως η προσέγγισή μας δύναται να εφαρμοστεί σε εύρος κλινικών φαινοτύπων, βάσει του συνόλου των τρέχοντων, ελεύθερα διαθέσιμων δεδομένων και κειμένων, με ικανότητα επικαιροποίησης.

Βιβλιογραφία

- Awad F, A. E.-L. (2018). Inflammasome biology, molecular pathology and therapeutic implications. *Pharmacol Ther.*, 187, pp. 133-149.
- Bauernfeind F, H. V. (2013). Of inflammasomes and pathogens--sensing of microbes by the inflammasome. *EMBO Mol Med.*, 5(6), pp. 814-26.
- Bhat AS, C. M. (2013). Prevalence of celiac disease in Indian children with Down syndrome and its clinical and laboratory predictors. *Indian J Pediatr.*, 80(2), pp. 114-7.
- Bouziat R, H. R.-B. (2017). Reovirus infection triggers inflammatory responses to dietary antigens and development of celiac disease. *Science.*, 356(6333), pp. 44-50.
- Chan AH, S. K. (2020). Inflammasome signaling and regulation of interleukin-1 family cytokines. *J Exp Med.*, 217(1), p. e20190314.
- Chen A, L. C.-D. (2016). Celiac Crisis Associated with Herpes Simplex Virus Esophagitis. *ACG Case Rep J.*, 3(4), p. e159.
- Fehniger TA, C. M. (2001). Interleukin 15: biology and relevance to human disease. *Blood.*, 97(1), pp. 14-32.
- Fessatou S, K. M. (2003). Coeliac disease and alopecia areata in childhood. *J Paediatr Child Health.*, 39(2), pp. 152-4.
- Fracchia M, G. G. (2004). Coeliac disease associated with Sjögren's syndrome, renal tubular acidosis, primary biliary cirrhosis and autoimmune hyperthyroidism. *Dig Liver Dis.*, 36(7), pp. 489-91.
- Gómez Castro MF, M. E. (2019). p31-43 Gliadin Peptide Forms Oligomers and Induces NLRP3 Inflammasome/Caspase 1- Dependent Mucosal Damage in Small Intestine. *Front Immunol.*, 10, p. 31.
- Guo H, C. J. (2015). Inflammasomes: mechanism of action, role in disease, and therapeutics. *Nat Med.*, 21(7), pp. 677-87.
- Harley JB, C. X. (2018). Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat Genet.*, 50(5), pp. 699-707.
- Kahrs CR, C. K. (2019). Enterovirus as trigger of coeliac disease: nested case-control study within prospective birth cohort. *BMJ.*, 364, p. l231.
- Kooy-Winkelaar YM, B. D. (2017). CD4 T-cell cytokines synergize to induce proliferation of malignant and nonmalignant innate intraepithelial lymphocytes. *Proc Natl Acad Sci U S A.*, 114(6), pp. E980-E989.
- Liu ZP, W. Y. (2012). Network-based analysis of complex diseases. *IET Syst Biol.*, 6(1), pp. 22-33.
- Malamut G, C. O.-B. (2013). Enteropathy associated T cell lymphoma in celiac disease: a large retrospective study. *Dig Liver Dis.*, 45(5), pp. 377-84.

- Matzaraki V, K. V. (2017). The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biol.*, 18(1), p. 76.
- Nadhem ON, A. G. (2015). Review and practice guidelines for celiac disease in 2014. *Review and practice guidelines for celiac disease in 2014.*, 127(3), pp. 259-65.
- Nanayakkara M, L. G. (2018). P31-43, an undigested gliadin peptide, mimics and enhances the innate immune response to viruses and interferes with endocytic trafficking: a role in celiac disease. *Sci Rep.*, 8(1), p. 10821.
- Nayarisseri A, K. R.-P. (2021). Artificial Intelligence, Big Data and Machine Learning Approaches in Precision Medicine & Drug Discovery. *Curr Drug Targets.*, 22(6), pp. 631-655.
- Palová-Jelínková L, D. K.-K. (2013). Pepsin digest of wheat gliadin fraction increases production of IL-1 β via TLR4/MyD88/TRIF/MAPK/NF- κ B signaling pathway and an NLRP3 inflammasome activation. *PLoS One.*, 8(4), p. e62426.
- Rathinam VA, F. K. (2016). Inflammasome Complexes: Emerging Mechanisms and Effector Functions. *Cell.*, 165(4), pp. 792-800.
- Sharp J, P. B. (2012). Anaplastic large-cell lymphoma in a child with type I diabetes and unrecognised coeliac disease. *Case Rep Pediatr*, 2012, p. 269689.
- Stene LC, H. M. (2006). Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am J Gastroenterol.*, 101(10), pp. 2333-40.
- Sun YV, H. Y. (2016). Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv Genet.*, 93, pp. 147-90.
- Tjernberg AR, L. J. (2014). Children with celiac disease are more likely to have attended hospital for prior respiratory syncytial virus infection. *Dig Dis Sci.*, 59(7), pp. 1502-8.
- Vamathevan J, C. D. (2016). Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov.*, 18(6), pp. 463-477.
- Zmora N, L. M.-F. (2017). Inflammasomes and intestinal inflammation. *Mucosal Immunol.*, 10(4), . 865-883.

Περίληψη

Οι σύνθετοι κλινικοί φαινότυποι αποτελούν τη συνισταμένη γενετικών και περιβαλλοντικών επιρροών, που απαιτείται όχι μόνον να ταυτοποιηθούν, αλλά και να χαρτογραφηθούν οι συνοδές αλληλεπιδράσεις τους. Στο αυτό πλαίσιο, η προβληματική αυτορρύθμιση του φλεγμονοσώματος αποτελεί παράδειγμα καίριας σημασίας. Οι φλεγμονώδεις αποκρίσεις μετρατρέπουν έναν αμυντικό μηχανισμό σε μια διαωινιζόμενη φλεγμονώδη απόκριση, που οδηγεί σε δριμείες εκδηλώσεις και συννοσηρότητες.

Στην παρούσα μεταδιδακτορική έρευνα, αναπτύξαμε, βελτιστοποιήθηκε και εφαρμόστηκε μια καινοτόμος προγνωστική προσέγγιση για την αξιολόγηση κινδύνου συννοσηροτήτων, που υπογραμμίζει τις μοριακές υπογραφές της δριμείας νόσου. Ως πρότυπο-μοντέλο ασθένειας χρησιμοποιήθηκε μια πολυγονιδιακή και πολυπαραγοντική, σύνθετη, χρόνια ανοσομεσολαβούμενη εντεροπάθεια, η κοιλιοκάκη. Για την εν τω βάθει *in silico* ανάλυσή μας, βάσει ελεύθερα διαθέσιμων δεδομένων και κειμένων, χρησιμοποιήθηκαν μεθοδολογίες τεχνητής νοημοσύνης, μοριακής μοντελοποίησης και μηχανικής μάθησης (γλώσσες προγραμματισμού R και Python). Εφαρμόστηκαν δύο στρατηγικές: αγνωστική (μη στοχευμένη) ανάλυση και στοχευμένη ανάλυση. Τα πρωτεϊνικά δίκτυα αποκάλυψαν 1.156 παρανοηματικές παραλλαγές ενδιαφέροντος, αλλά και πρωτεϊνικά μεταλλάγματα και πρωτεομορφές, καθώς και δυνητικοί συνδέτες προς τη φαρμακευτική στόχευσή τους.

Μια προγνωστική αξιολόγηση του κινδύνου ανάπτυξης συννοσηρότητας ή/και δριμείας νόσου μπορεί να συνδράμει στον προσδιορισμό ασθενών υψηλού κινδύνου και επομένως, στη βέλτιστη διαστρωμάτωση των ασθενών.

Υπογραφή Επιβλέπουσας Καθηγήτριας



Κατσίλα Θεοδώρα

Υπογραφή Υποτρόφου



Μπαλασοπούλου Αγγελική