



**ΕΘΝΙΚΟ
ΜΕΤΣΟΒΙΟ
ΠΟΛΥΤΕΧΝΕΙΟ**

ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

Α.Π. : 31061
Αθήνα, 14/6/18

ΚΟΣΜΗΤΟΡΑΣ

**Προς τα Μέλη ΔΕΠ της
Σχολής Μηχ/γων
Μηχ/κών**

ΠΡΟΣΚΛΗΣΗ

Σας προσκαλούμε στην παρουσίαση της Διδακτορικής Διατριβής του Υ.Δ. κ. **ΣΑΚΕΛΛΑΡΟΠΟΥΛΟΥ Θεοδώρου** που εκπόνησε στον Τομέα Μηχανολογικών Κατασκευών & Αυτομάτου Ελέγχου, διπλωματούχος **ΜΗΧΑΝΟΛΟΓΟΣ ΜΗΧΑΝΙΚΟΣ** του ΕΜΠ, που θα πραγματοποιηθεί την *Τετάρτη 4 Ιουλίου 2018*, ώρα *15:00μ.μ.* στην αίθουσα Τηλεδιάσκεψης 1 στο κτήριο της Κεντρικής Βιβλιοθήκης του ΕΜΠ - Πολυτεχνειούπολη Ζωγράφου. Ο ελληνικός τίτλος της Διδακτορικής Διατριβής είναι ο εξής :

**«ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΚΥΤΑΡΙΚΗΣ ΑΠΟΚΡΙΣΗΣ ΜΕ ΜΕΘΟΔΟΥΣ
ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ»**

Και ο Αγγλικός τίτλος ως εξής:

« MODELLING CELL RESPONSE VIA OPTIMIZATION METHODS »

Ο ΚΟΣΜΗΤΟΡΑΣ ΤΗΣ ΣΧΟΛΗΣ


N. ΜΑΡΜΑΡΑΣ
Καθηγητής Ε.Μ.Π

Εκτεταμένη Περίληψη

Εισαγωγή

Συστημική Βιολογία είναι ένα μοντέλο βιολογικής έρευνας που προάγει την ολιστική περιγραφή των υπό μελέτη μοντέλων εν αντιθέσει με την παραδοσιακή Βιολογία που εστιάζει την περιγραφή των επιμέρους στοιχείων. Η φιλοδοξία για μία ολιστική κατανόηση των βιολογικών συστημάτων δεν είναι καινούργια, αλλά μόνο τις τελευταίες δύο δεκαετίες με την ραγδαία ανάπτυξη των τεχνολογιών μοριακής μέτρησης κατέστη δυνατή η συλλογή του απαραίτητου όγκου πληροφορίας που απαιτεί ένα τέτοιο εγχείρημα. Ο Δρ. Κιάνο, ένας εκ των “πατέρων” της σύγχρονης αναγέννησης του τομέα, υποδεικνύει 4 κατευθύνσεις αυτής της έρευνας:

- αναγνώριση της δομής του υπό μελέτη συστήματος
- προσδιορισμός των δυναμικών του ιδιοτήτων
- ανάπτυξη τεχνικών για τον έλεγχο του συστήματος
- ανάπτυξη μεθόδων για τον σχεδιασμό νέων συστημάτων

Η παρούσα διατριβή εστιάζει στην μελέτη της κυτταρικής συμπεριφοράς μέσω της αναγνώριση της δομής των σηματοδοτικών δικτύων του κυττάρου.

Το κύτταρο μοντελοποιείται ως ένα σύστημα πρωτεϊνών και γονιδίων. Η κατάσταση του συστήματος περιγράφεται από την συγκέντρωση και τη μορφή των πρωτεϊνών καθώς και την έκφραση των γονιδίων που το αποτελούν. Οι πρωτεΐνες είναι μακρομόρια, η μορφή και η συγκέντρωση των οποίων ρυθμίζει τις μακροσκοπικές ιδιότητες του κυττάρου, ενώ τα γονίδια είναι κομμάτια του γενετικού υλικού του κυττάρου, η έκφραση των οποίων ρυθμίζει την παραγωγή πρωτεϊνών. Το σύστημα αυτό βρίσκεται μέσα σε ένα περιβάλλον από το οποίο δέχεται σήματα τα οποία επεξεργάζεται και ανταποκρίνεται αλλάζοντας την κατάσταση του.

Τα σηματοδοτικά δίκτυα είναι ο μηχανισμός μέσω του οποίου το κύτταρο επεξεργάζεται και ανταποκρίνεται σε πληροφορίες από το περιβάλλον του. Συγκεκριμένα, το κύτταρο επικοινωνεί με το περιβάλλον του κυρίως μέσω ειδικών πρωτεϊνών που καλούνται υποδοχείς. Οι πρωτεΐνες αυτές, κατά πλειονότητα, προεξέχουν εκατέρωθεν της κυτταρικής μεμβράνης και όταν έρθουν σε επαφή με κάποιο περιβαλλοντικό ερέθισμα αλλάζουν μορφή. Η αλλαγή αυτή αναγνωρίζεται με τη σειρά της από τις γύρω πρωτεΐνες και διαδίδεται μέσω μίας αλληλουχίας βιοχημικών αντιδράσεων μέχρι των πυρήνα του κυττάρου όπου επηρεάζει την έκφραση των γονιδίων και εν τέλει αλλάζει την συνολική κατάσταση του κυττάρου.

Μία από τις σημαντικότερες σηματοδοτικές αντιδράσεις είναι αυτή της (απο)-φωσφορυλίωσης, κατά την οποία μία πρωτεΐνη αποκτά (χάνει) μία φωσφορική ομάδα με αποτέλεσμα να αλλάξει η μορφή της. Η αλλαγή αυτή, “ενεργοποιεί” (σηματοδοτικά) την πρωτεΐνη. Οι αντιδράσεις αυτές καταλύονται από ειδικά ένζυμα: τις κινάσες για την φωσφορυλίωση και τις φωσφατάσες για την αποφωσφορυλίωση. Για την παρούσα διατριβή, η μέτρηση των φωσφορυλιωμένων πρωτεϊνών έγινε με τη τεχνική bead-based sandwich ELISA. Η τεχνική αυτή επιτρέπει την μέτρηση ενός προκαθορισμένου πλήθους πρωτεϊνών σε πολλά δείγματα.

Ο όρος γονιδιακή έκφραση αναφέρεται στη διαδικασία κατά την οποία μέρος του γενετικού κώδικα (γονίδιο) μεταγράφεται σε RNA και βγαίνει από τον πυρήνα με σκοπό να μεταφραστεί σε πρωτεΐνη. Ως ένταση της έκφρασης ορίζεται το πλήθος των αντιγράφων RNA το οποίο είναι ανάλογο της συγκέντρωσης της παραγόμενης πρωτεΐνης. Για τις ανάγκες αυτής της διατριβής χρησιμοποιήθηκαν μετρήσεις τύπου μικροσυστοιχίας DNA από δημόσιες βάσεις δεδομένων. Οι τεχνικές αυτές έχουν την δυνατότητα να μετρούν την έκφραση ενός προκαθορισμένου αριθμού γονιδίων σε πολλά δείγματα.

Η αναγνώριση της δομής των σηματοδοτικών δικτύων συνίσταται στην περιγραφή των σηματοδοτικών αντιδράσεων μεταξύ πρωτεϊνών και γονιδίων. Προς αυτή την κατεύθυνση, έχουν αναπτυχθεί πολλές τεχνικές. Οι σημαντικότερες για την αναγνώριση πρωτεϊνικών αντιδράσεων είναι η yeast-two-hybrid (Y2H) και οι AP-MS, CoFrac-MS που βασίζονται στην τεχνολογία

φασματομετρία μάζας (Mass Spectrometry). Για τον προσδιορισμό των αντιδράσεων μεταξύ πρωτεϊνών (μεταγραφικοί παράγοντες) και γονιδίων επίσης έχουν αναπτυχθεί πολλές τεχνολογίες οι σημαντικότερες των οποίων βασίζονται στην τεχνική της ανοσοκατακρήμνισης χρωματίνης (chromatin immunoprecipitation - ChIP-X).

Οι τεχνικές αυτές έχουν δεχθεί κριτική για την ακρίβεια και την ευαισθησία τους, αλλά ο πιο σημαντικός περιορισμός τους δεν είναι τεχνικός αλλά θεμελιώδης. Τα σηματοδοτικά δίκτυα δεν είναι στατικές οντότητες οι οποίες μπορούν να μετρηθούν αν χρησιμοποιηθούν αρκετοί πόροι, αλλά μεταβάλλονται και προσαρμόζονται ανά κύτταρο και στις συνθήκες του περιβάλλοντος.

Μέθοδοι

Έχουν αναπτυχθεί πολλές μέθοδοι που χρησιμοποιούν τον κατάλογο των γνωστών αντιδράσεων σαν ένα δίκτυο πρότερης γνώσης (Prior Knowledge Network – PKN) και το προσαρμόζουν σε πειραματικά δεδομένα που περιγράφουν την απόκριση, ως αλλαγή στην κατάσταση, των υπό μελέτη κυττάρων σε εξωτερικά ερεθίσματα. Με αυτό το τρόπο επιτυγχάνουν όχι μόνο την αναγνώριση των ειδικών σηματοδοτικών δικτύων, αλλά και περιορίζουν τον θόρυβο των μετρήσεων μέσω σύμπλεξης τους.

Για την παρούσα διατριβή, αναπτύχθηκαν δύο τέτοιες μέθοδοι. Η μέθοδος Συνέπειας Επιρροών (Sign Consistency Method – SCM) και η μέθοδος Αντίστροφου Αιτιακού Λογισμού (Reverse Causal Reasoning – RCR). Οι μέθοδοι αυτές μοντελοποιούν τη διαδικασία σηματοδότησης σαν ένα σύστημα λογικών εξισώσεων. Όπως όλα τα μοντέλα που βασίζονται στην λογική, στοχεύουν σε μία ποιοτική περιγραφή του συστήματος και αντισταθμίζουν την “θυσία” στη διακριτική τους ικανότητα με το γεγονός ότι δεν χρειάζονται πολύ μεγάλο όγκο δεδομένων συγκριτικά με άλλες πιο ρεαλιστικές μεθόδους.

Οι μέθοδοι που αναπτύχθηκαν γενικεύουν και εξελίσσουν μία σειρά μεθόδων στις οποίες έχει συμβάλει το εργαστήριο Συστημικής Βιοϊατρικής της σχολής Μηχανολόγων Μηχανικών του ΕΜΠ. Συγκεκριμένα, και οι δύο μέθοδοι χρησιμοποιούν Γράφους Επιρροής (Interaction or Influence Graph - IG) για να μοντελοποιήσουν το δίκτυο και το μοντέλο της συνέπειας επιρροών για τον μηχανισμό διάδοσης σήματος. Τέλος, όπως και στις προηγούμενες δημοσιεύσεις, χρησιμοποιείται η τεχνική του ακέрайου γραμμικού προγραμματισμού για την επίλυση των προβλημάτων βελτιστοποίησης που σχεδιάζονται.

Οι Γράφοι Επιρροής είναι προσημασμένοι κατευθυνόμενοι γράφοι $G = (N, A, \sigma)$ όπου: N το σύνολο των κόμβων, A το σύνολο των ακμών, και $\sigma : A \rightarrow \{+1, -1\}$. Οι κόμβοι του γράφου αντιστοιχούν σε πρωτεΐνες ή γονίδια ενώ οι ακμές σε αντιδράσεις μεταξύ του. Το πρόσημο της ακμής υποδεικνύει την επιρροή που ασκεί ο μητρικός κόμβος στον θυγατρικό. Συγκεκριμένα, οι κόμβοι του δικτύου μπορούν να βρεθούν σε μία εκ των τριών πιθανών καταστάσεων: υπερδιέγερσης (+1) που σημαίνει αύξηση στην συγκέντρωση των φωσφοπρωτεϊνών ή στην έκφραση των γονιδίων, υποδιέγερσης (-1) που είναι η αντίστροφη κατάσταση, ή στην βασική κατάσταση (0) που υποδηλώνει καμία αλλαγή στην κατάσταση του κόμβου ως αποτέλεσμα της σηματοδότησης. Η επιρροή που ασκεί ένας μητρικός κόμβος στους θυγατρικούς του εξαρτάται από το πρόσημο της συνδέουσας ακμής $\sigma(a)$.

Για να είναι συνεπείς δύο κόμβοι που συνδέονται με ακμή πρέπει να βρίσκονται στην ίδια κατάσταση αν η ακμή είναι θετικά προσημασμένη και αντίθετα αν είναι αρνητικά. Για να είναι ένας γράφος ασθενώς συνεπής, τότε πρέπει κάθε κόμβος να είναι συνεπής με τουλάχιστον έναν από τους μητρικούς του κόμβους. Ένας συνεπής γράφος ερμηνεύει ποιοτικά τη διαδικασία διάδοσης σήματος καθώς για κάθε κόμβο υπάρχει ένα μονοπάτι αιτιακών σχέσεων που εξηγεί την κατάσταση του με βάση κάποιο αρχικό ερέθισμα. Επομένως, στόχος και των 2 προτεινόμενων μεθόδων είναι να βρουν τον συνεπή γράφο ο οποίος ερμηνεύει καλύτερα τα πειραματικά αποτελέσματα.

Η μέθοδος SCM διατυπώνει το πρόβλημα της ανεύρεσης του βέλτιστου συνεπή γράφου ως πρόβλημα ακέрайου γραμμικού προγραμματισμού. Συγκεκριμένα, ψάχνει όλα τα υποσύνολα του PKN για να βρει αυτό το οποίο μπορεί να προσομοιώσει βέλτιστα τις καταστάσεις των διαφόρων κόμβων όπως αυτές μετρήθηκαν πειραματικά. Οι περιορισμοί του προβλήματος κωδικοποιούν το

μοντέλο συνεπών επιρροών και εξασφαλίζουν ότι ο γράφος μπορεί να σημανθεί με τρόπο συνεπή ενώ η αντικειμενική συνάρτηση κωδικοποιεί την απόσταση μεταξύ της προσομοιωμένης και της πειραματικά προσδιορισμένης κατάστασης των κόμβων που μετρήθηκαν.

Σε προηγούμενες διατυπώσεις του προβλήματος, η αναζήτηση για το βέλτιστο υποδίκτυο γινόταν μόνο επί του συνόλου των ακμών. Στη διατύπωση που αναπτύχθηκε για αυτή τη διατριβή, η αναζήτηση γίνεται και επί του συνόλου των κόμβων. Η αλλαγή αυτή δεν επηρεάζει το σύνολο των λειτουργικά ισοδύναμων λύσεων. Τα βασικά οφέλη που αποφέρει είναι ότι πρώτον, επιτρέπει στον ερευνητή μεγαλύτερη ελευθερία να ενσωματώσει την πρότερη γνώση του στο σύστημα κατευθύνοντας την λύση, και δεύτερον, καθιστά τη διαδικασία της βελτιστοποίησης πιο ρεαλιστική καθώς υπάρχουν φυσικά γεγονότα που μπορούν να οδηγήσουν στην πλήρη αδρανοποίηση μίας πρωτεΐνης ή γονιδίου.

Ένας άλλος σημαντικός περιορισμός αυτής της μεθόδου ήταν η αδυναμία της να διερευνήσει δίκτυα που είχαν βρόχου ανατροφοδότησης (feedback loops), διότι βρόχοι μπορούν να μεταβούν σε μία μη βασική κατάσταση χωρίς να λάβουν σήμα από κάποιο ερέθισμα από το υπόλοιπο δίκτυο. Στη διατύπωση που αναπτύχθηκε για αυτή την διατριβή, επιπλέον περιορισμοί έχουν εισαχθεί οι οποίοι αποτρέπουν αυτή την παρενέργεια.

Η δεύτερη μέθοδος που αναπτύχθηκε RCR λύνει το πρόβλημα της ανεύρεσης του βέλτιστου συνεπή γράφου για την περίπτωση που το αρχικό ερέθισμα είναι άγνωστο. Για την μέθοδο SCM, απαιτούνται πειράματα κατά τα οποία τα κύτταρα διεγείρονται τεχνητά, συνήθως με κυτοκίνες, και εν συνεχεία μετράται η απόκριση τους, ως αλλαγή στην συγκέντρωση κάποιων φωσφοπρωτεϊνών ή την έκφραση κάποιων γονιδίων. Παρόλα αυτά, πολλές φορές η γενεσιουργός αιτία της κυτταρικής απόκρισης δεν είναι γνωστή αλλά η απόκριση μπορεί να μετρηθεί σαν διαφορά μεταξύ ενός πληθυσμού ελέγχου και ενός αποκρίνοντος, για παράδειγμα μεταξύ υγιών και καρκινικών κυττάρων. Η μέθοδος RCR γενικεύει περαιτέρω την μέθοδο SCM ώστε να μπορεί να μοντελοποιήσει και αυτά τα πειράματα. Συγκεκριμένα, η αρχική διέγερση, η οποία θεωρείτο γνωστή προηγουμένως, μετατρέπεται σε μεταβλητή η οποία συμμετέχει στην διαδικασία εύρεσης του βέλτιστου γράφου. Επειδή αυτή η αλλαγή αυξάνει πολύ τους βαθμούς ελευθερίας του προβλήματος, απαιτείται η χρήση επιπλέον περιορισμών για να διασφαλιστεί ότι οι λύσεις δεν θα είναι τετριμμένες. Οι περιορισμοί αυτοί έχουν στόχο να περιορίσουν τον αριθμό των πιθανών αιτιών και το μέγεθος του δικτύου και μπορούν να επιβληθούν ως αυστηρά όρια ή/και ως επιπλέον κόστη στη αντικειμενική συνάρτηση.

Εφαρμογές

Στη συνέχεια περιγράφονται 3 εφαρμογές στις οποίες έγινε χρήση των μεθόδων που περιγράφηκαν για την ανίχνευση σηματοδοτικών δικτύων. Τα δίκτυα αυτά χρησιμοποιήθηκαν στη συνέχεια για να διερευνηθούν οι σημαντικότεροι συντελεστές της κυτταρικής απόκρισης ή για να προταθούν πιθανές φαρμακευτικές αγωγές που θα επαναφέρουν τα κύτταρα σε υγιή κατάσταση.

Η πρώτη εφαρμογή που περιγράφεται αφορά τη διαδικασία της κυτταρικής γήρανσης. Για αυτή την εφαρμογή, ως βιολογικό μοντέλο χρησιμοποιήθηκαν ανθρώπινοι πρωτογενείς ινοβλάστες. Μετρήθηκε η συγκέντρωση 18 φωσφοπρωτεϊνών σε νεαρά και γηρασμένα κύτταρα κατόπιν διέγερσης των κυττάρων με 6 διαφορετικές κυτοκίνες. Τα νεαρά κύτταρα διπλασιάζονταν κάθε μέρα ενώ τα γηρασμένα δεν είχαν διπλασιαστεί για ένα μήνα. Για τις φωσφοπρωτεΐνες αυτές, προηγούμενες έρευνες είχαν δείξει ότι συμμετέχουν στη γήρανση και στον πολλαπλασιασμό. Η κατάσταση κάθε πρωτεΐνης προσδιορίστηκε κατόπιν σύγκρισης της συγκέντρωσης της με την αντίστοιχη κυττάρων που δεν είχαν διεγερθεί με κάποιο ερέθισμα. Σε συνδυασμό με το PKN από τη βάση αντιδράσεων MetaCore χρησιμοποιήθηκε μια παραλλαγή της μεθόδου SCM για να ανιχνευτεί το σηματοδοτικό δίκτυο των 2 κυτταρικών πληθυσμών. Η παραλλαγή συνίστατο στην σύζευξη των αντιδράσεων των 2 δικτύων. Συγκεκριμένα, θεωρήθηκε ότι οι διαφοροποιήσεις κατά την γήρανση οφείλονται σε μείωση της έκφρασης των γονιδίων οπότε το σύνολο το αντιδράσεων θα έπρεπε να είναι κοινό και τα δίκτυα να διαφέρουν μόνο ως προς το σύνολο των κόμβων. Οι διαφορές των δύο δικτύων επαλήθευσαν προηγούμενες μελέτες που κατέδειξαν υποτονική απόκριση στην παρουσία ινσουλίνης και ισχυρή ενεργοποίηση του NFκΒ και των συναφών

λειτουργιών κατά την διέγερση με IL1A και TNF. Επιπλέον, η δομή των δικτύων είχε διαφορές που δεν είχαν αναφερθεί στην σχετική βιβλιογραφία όπως ο κεντρικός ρόλος που παίζει η SRC, η οποία φαίνεται να είναι ανενεργή στα γηρασμένα κύτταρα, στη ρύθμιση πολλών πρωτεϊνών που σχετίζονται με τον πολλαπλασιασμό των κυττάρων.

Για τη δεύτερη εφαρμογή που περιγράφεται χρησιμοποιήθηκε η μέθοδος RCR για να αναλυθεί ο μηχανισμός λειτουργίας της φαρμακολογικής τοξικότητας στον πνεύμονα και να προταθούν πιθανές θεραπείες. Συγκεκριμένα, επιλέχθηκαν 200 φάρμακα από τη βάση δεδομένων Pneumotox τα οποία έχουν την δυνατότητα να προκαλέσουν ασθένεια του πνεύμονα. Στη συνέχεια οι πιθανοί στόχοι αυτών των φαρμάκων ανακτήθηκαν από την βάση δεδομένων STICH. Τέλος, χρησιμοποιήθηκε η βάση CMap, η οποία περιγράφει διαφορές στην γονιδιακή έκφραση 1000 γονιδίων σε 5 κυτταροσειρές πριν και μετά την επώαση τους με αυτά τα φάρμακα. Με βάση αυτές τις πληροφορίες και σε συνδυασμό με ένα PKN από τη βάση Reactome, δημιουργήθηκε ένα σηματοδοτικό δίκτυο για κάθε φάρμακο το οποίο ερμηνεύει την παρατηρούμενη αλλαγή στην γονιδιακή έκφραση. Επειδή και τα 200 φάρμακα μπορούν να προκαλέσουν ασθένεια του πνεύμονα, αναλύθηκε η τομή των 200 δικτύων και ανιχνεύτηκαν 2 υποδίκτυα τα οποία ήταν συστηματικά υπέρ ή υπό διεγερμένα. Για την επαλήθευση των αποτελεσμάτων, χρησιμοποιήθηκε ο αλγόριθμος GUIDE για να κατασκευάσει έναν ανεξάρτητο δίκτυο με βάση τα ίδια δεδομένα και παρατηρήθηκε ότι τα 2 αποτελέσματα είχαν στατιστικά σημαντική αλληλοκάλυψη. Στη συνέχεια, με παρόμοιο τρόπο κατασκευάστηκαν σηματοδοτικά δίκτυα και για τα υπόλοιπα, μη-τοξικά, φάρμακα που ήταν κοινά στις STICH και CMap έτσι ώστε να βρεθούν φάρμακα που να επάγουν την αντίστροφη κατάσταση στα υποδίκτυα που αναφέρθηκαν. Τα καλύτερα 40 προτεινόμενα φάρμακα, ελέγχθηκαν από φαρμακολόγο η οποία παρατήρησε ότι τα περισσότερα έχουν την δυνατότητα να βελτιώσουν την κατάσταση κυρίως λόγω της αντιφλεγμονώδους δράσης τους.

Για την τρίτη εφαρμογή, ακολουθήθηκε η ίδια διαδικασία όπως και στη δεύτερη αλλά αυτή τη φορά για να προταθούν φάρμακα για την αντιμετώπιση της μόλυνσης από το βακτήριο του άνθρακα. Χρησιμοποιήθηκαν τα ίδια μη-τοξικά φάρμακα όπως και στη δεύτερη εφαρμογή. Για την μελέτη των συνεπειών της μόλυνσης από άνθρακα στην γονιδιακή έκφραση χρησιμοποιήθηκαν δεδομένα από 4 έρευνες από την βάση GEO. Αφού κατασκευάστηκε το σηματοδοτικό δίκτυο που μοντελοποιεί την κυτταρική απόκριση στις τοξίνες που εκκρίνει το βακτήριο, με χρήση της μεθόδου RCR, αναζητήθηκαν φάρμακα που να αντιστρέφουν την απόκριση αυτή. Η αναζήτηση έγινε συγκρίνοντας τα δίκτυα των φαρμάκων και του άνθρακα, και τα φάρμακα ταξινομήθηκαν από το καλύτερο το χειρότερο ανάλογα με την απόσταση τους από την πλήρη αντιστροφή του δικτύου του άνθρακα. Για επαλήθευση έγινε και πάλι βιβλιογραφική έρευνα για τα φάρμακα που ήταν στη κορυφή και στον πάτο τη κατάταξης καταλληλότητας. Και για τις δύο κατηγορίες τα αποτελέσματα κρίθηκαν λογικά, καθώς πολλά από τα φάρμακα που ήταν ψηλά στη λίστα είχαν ήδη προταθεί και από άλλες μελέτες, ενώ αντίθετα φάρμακα που ήταν χαμηλά στη λίστα είχαν παρόμοιους στόχους με αυτούς του βακτηρίου οπότε αναμένεται να είχαν παρόμοιες συνέπειες στα κύτταρα.