



Αποκωδικοποιώντας τον πηγαίο κώδικα του εμβολίου κατά του Κορωνοϊού

Posted on Dec 26 2020

Ας ξεκινήσουμε με ένα μικρό μέρος του ίδιου του πηγαίου κώδικα του εμβολίου BioNTech/ Pfizer, επίσης γνωστό ως BNT162b2 ή Tozinameran, ή και ως Comirnaty.



WHO
International Nonproprietary Names Programme

9/2020

Sequence / Séquence / Secuencia

GAGAAΨAAAC	ΨAGΨAΨCΨΨ	CΨGGΨCCCCA	CAGACΨCAGA	GAGAACCCGC	50
CACCAΨGΨΨC	GΨGΨCΨΨGG	ΨGCΨGCΨGCC	ΨCΨGGΨΨC	AGCCAGΨGΨG	100
ΨGAACΨGAC	CACCAGAACA	CAGCΨGCCΨC	CAGCCΨACAC	CAACAGCΨΨΨ	150
ACCAGAGGGC	ΨGΨACΨACCC	CGACAAGGΨG	ΨΨCAGAΨCCA	GCGΨGCΨGCA	200
CΨCΨACCCAG	GACCΨGΨΨC	ΨGCCΨΨCΨΨ	CAGCAACGΨG	ACCΨGGΨΨC	250
ACGCCAΨCCA	CGΨGΨCCGGC	ACCAAΨGGCA	CCAAGAGAΨΨ	CGACAACCCC	300
GΨGCΨGCCΨΨ	ΨCAACGACGG	GGΨGΨACΨΨΨ	GCCAGCACCG	AGAAGΨCCAA	350
CAΨCAΨCAGA	GGCΨGGAΨCΨ	ΨCGGCACCAC	ACΨGGACAGC	AAGACCCAGA	400
GCCΨGCΨGAΨ	CGΨGAACAAC	GCCACCAACG	ΨGGΨCAΨCAA	AGΨGΨGCGAG	450
ΨΨCCAGΨΨCΨ	GCAACGACCC	CΨΨCCΨGGGC	GΨCΨACΨACC	ACAAGAACAA	500

[Πρώτοι 500 χαρακτήρες του BNT162b2 mRNA. **Πηγή: Παγκόσμιος Οργανισμός Υγείας**]

Το εμβόλιο mRNA BNT162b έχει αυτόν τον ψηφιακό κώδικα στην καρδιά του. Έχει μήκος 4284 χαρακτήρες, οπότε χωράει σε μερικά tweets. Στην αρχή της διαδικασίας παραγωγής εμβολίων, κάποιος ανέβασε αυτόν τον κώδικα σε έναν εκτυπωτή DNA (ναι!), ο οποίος στη συνέχεια μετέτρεψε τα byte στο δίσκο σε πραγματικά μόρια DNA.



Ένας εκτυπωτής DNA Codex DNA BioXp 3200

Από ένα τέτοιο μηχάνημα βγαίνουν μικροσκοπικές ποσότητες DNA, οι οποίες μετά από πολλές βιολογικές και χημικές επεξεργασίες καταλήγουν ως RNA (περισσότερα για τα αυτά αργότερα) στο φιαλίδιο εμβολίου. Μια δόση 30 μικρογραμμάρων αποδεικνύεται ότι περιέχει πραγματικά 30 μικρογραμμάρια RNA. Επιπλέον, υπάρχει ένα έξυπνο σύστημα συσκευασίας λιπιδίων που εισάγει το mRNA στα κύτταρα μας.

Το RNA είναι η πτητική έκδοση «μνήμης εργασίας» του DNA. Το DNA είναι σαν φλασάκι για την αποθήκευση της βιολογικής πληροφορίας. Το DNA είναι πολύ ανθεκτικό, εσωτερικά πλεοναστικό και πολύ αξιόπιστο. Αλλά όπως οι υπολογιστές δεν εκτελούν κώδικα απευθείας από μια μονάδα flash, πριν συμβεί κάτι, ο κώδικας αντιγράφεται σε ένα γρηγορότερο, πιο ευέλικτο αλλά συνάμα και πολύ πιο εύθραυστο σύστημα.

Για υπολογιστές, αυτή είναι η RAM, για τη βιολογία είναι το RNA. Η ομοιότητα είναι εντυπωσιακή. Σε αντίθεση με τη μνήμη flash, η μνήμη RAM υποβαθμίζεται πολύ γρήγορα, εκτός αν κάποιος την φροντίσει. Ο λόγος για τον οποίο το εμβόλιο Pfizer / BioNTech mRNA πρέπει να συντηρηθεί σε βαθιά κατάψυξη είναι ο ίδιος: το RNA είναι ένα εύθραυστο λουλούδι.

Κάθε χαρακτήρας RNA ζυγίζει στην τάξη των $0,53 \cdot 10^{-21}$ γραμμαρίων, που σημαίνει ότι υπάρχουν $6 \cdot 10^{16}$ χαρακτήρες σε μία δόση εμβολίου 30 μικρογραμμάρων. Εκφρασμένο σε byte, αυτό είναι περίπου 25 petabytes, αν και πρέπει να πούμε ότι αποτελείται από περίπου 2000 δισεκατομμύρια επαναλήψεις των ίδιων 4284 χαρακτήρων. Το σημαντικό (πληροφοριακό) περιεχόμενο του εμβολίου υπερβαίνει ελάχιστα το ένα kilobyte. Το ίδιο το SARS-CoV-2 ζυγίζει περίπου 7,5 kilobyte.

Ένα σύντομο υπόβαθρο

Το DNA είναι ένας ψηφιακός κωδικός. Σε αντίθεση με τους υπολογιστές, οι οποίοι χρησιμοποιούν 0 και 1, η ζωή χρησιμοποιεί A, C, G και U / T (τα «νουκλεοσίδια», «νουκλεοσίδια» ή «βάσεις»).

Στους υπολογιστές αποθηκεύουμε τα 0 και 1 ως ισχύς φόρτισης, ή ως ρεύμα, ως μαγνητική μετάβαση, ή ως τάση, ή ως διαμόρφωση σήματος ή ως αλλαγή

αντανακλαστικότητα. Η εν συντομία, τα 0 και 1 δεν είναι κάποιο είδος αφηρημένης έννοιας - ζουν ως ηλεκτρόνια και σε πολλές άλλες φυσικές ενσωματώσεις.

Στη φύση, τα A, C, G και U / T είναι μόρια, αποθηκευμένα ως αλυσίδες στο DNA (ή RNA).

Στους υπολογιστές, ομαδοποιούμε 8 bits σε ένα byte και το byte είναι η τυπική μονάδα δεδομένων που υποβάλλονται σε επεξεργασία.

Η φύση ομαδοποιεί 3 νουκλεοτίδια σε ένα κωδικόνιο, και αυτό το κωδικόνιο είναι η τυπική μονάδα επεξεργασίας. Ένα κωδικόνιο περιέχει 6 bits πληροφοριών ($2 * 3$).

Λοιπόν, τι κάνει αυτός ο κώδικας;

Η ιδέα ενός εμβολίου είναι να διδάξουμε στο ανοσοποιητικό μας σύστημα πώς να καταπολεμήσει έναν παθογόνο οργανισμό, χωρίς να αρρωστήσουμε πραγματικά. Ιστορικά αυτό γινόταν με την ένεση ενός εξασθενημένου ιού, καθώς και ενός «ανοσοεπιχυτικού» για να κινητοποιήσει το ανοσοποιητικό μας σύστημα σε δράση. Αυτή ήταν μια αποφασιστικά αναλογική τεχνική που περιλάμβανε δισεκατομμύρια αυγά (ή έντομα). Απαιτούσε επίσης πολλή τύχη και πολύ χρόνο. Μερικές φορές χρησιμοποιήθηκε επίσης ένας διαφορετικός (ασύνδετος) ιός.

Ένα εμβόλιο mRNA επιτυγχάνει το ίδιο πράγμα («εκπαίδευση του ανοσοποιητικού μας συστήματος»), αλλά με τρόπο μεγάλης ακρίβειας. Και αυτό το εννοώ και στις δύο έννοιες - πολύ στενά αλλά και πολύ δυνατά.

Λοιπόν, πώς λειτουργεί. Η ένεση περιέχει γενετικό υλικό που περιγράφει τη διάσημη πρωτεΐνη «Spike» SARS-CoV-2. Με έξυπνα χημικά μέσα, το εμβόλιο καταφέρνει να εισάγει αυτό το γενετικό υλικό σε μερικά από τα κύτταρα μας.

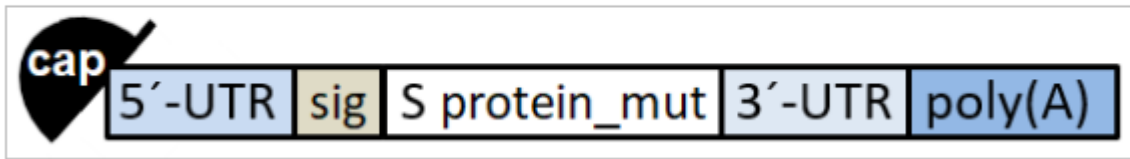
Αυτά στη συνέχεια αρχίζουν να παράγουν υπάκουα πρωτεΐνες SARS-CoV-2 Spike σε αρκετά μεγάλες ποσότητες ώστε το ανοσοποιητικό μας σύστημα να κινητοποιηθεί. Αντιμετωπίζοντας τις πρωτεΐνες Spike, ως ενδεικτικά σημάδια ότι τα κύτταρα έχουν καταληφθεί, το ανοσοποιητικό μας σύστημα αναπτύσσει μια ισχυρή απόκριση ενάντια σε πολλές πτυχές της πρωτεΐνης Spike ΚΑΙ ενάντια της διαδικασίας παραγωγής.

Και αυτό μας οδηγεί στο εμβόλιο με 95% αποτελεσματικότητα.

Ο πηγαίος κώδικας!

Το έγγραφο του ΠΟΥ έχει αυτήν την εικόνα:

Schematic



Αυτό είναι ένα είδος πίνακα περιεχομένων. Θα ξεκινήσουμε με το cap, που απεικονίζεται ως μικρό καπέλο.

Όπως δεν μπορείτε απλώς να ρίξετε orcodes σε ένα αρχείο υπολογιστή και να το εκτελέσετε, το βιολογικό λειτουργικό σύστημα απαιτεί κεφαλίδες, έχει συνδέσμους και πράγματα όπως συμβάσεις κλήσεων.

Ο κωδικός του εμβολίου ξεκινά με τα ακόλουθα δύο νουκλεοτίδια:

GA

Αυτό μπορεί να συγκριθεί σε μεγάλο βαθμό με κάθε DOS και Windows εκτελέσιμα που αρχίζουν με MZ, ή με UNIX scripts που αρχίζουν με #!. Τόσο στο σύστημα ζωής αλλά και στα λειτουργικά συστήματα, αυτοί οι δύο χαρακτήρες δεν εκτελούνται. Αλλά πρέπει να είναι εκεί γιατί διαφορετικά δεν συμβαίνει τίποτα.

Το mRNA 'cap' έχει πολλές λειτουργίες. Πρώτον, σημειώνει τον κώδικα ως προερχόμενο από τον πυρήνα. Στην περίπτωση μας φυσικά δεν ισχύει, ο κώδικας μας προέρχεται από το εμβόλιο. Αλλά δεν χρειάζεται να το πούμε αυτό στο κύτταρο. Το καπέλο κάνει τον κώδικα μας να φαίνεται «επικυρωμένο», το οποίο τον προστατεύει από την καταστροφή.

Τα δύο αρχικά GA νουκλεοτίδια είναι επίσης χημικώς ελαφρώς διαφορετικά από το υπόλοιπο του RNA. Υπό αυτήν την έννοια, το GA σήμα έχει κάποια σηματοδότηση «εκτός ζώνης».

Η «5' αμετάφραστη περιοχή»

Μερικές έννοιες εδώ. Τα μόρια RNA μπορούν να διαβαστούν μόνο σε μία κατεύθυνση. Παράξενα, το μέρος όπου ξεκινά η ανάγνωση ονομάζεται 5' ή «five-prime». Η ανάγνωση σταματά στο 3' ή στο «three-prime» άκρο.

Η ζωή αποτελείται από πρωτεΐνες (ή πράγματα που συνθέτονται από πρωτεΐνες). Και αυτές οι πρωτεΐνες περιγράφονται στο RNA. Όταν το RNA μετατρέπεται σε πρωτεΐνες, αυτό ονομάζεται μετάφραση.

Εδώ έχουμε την 5' αμετάφραστη περιοχή («UTR»), έτσι αυτό το κομμάτι δεν καταλήγει στην πρωτεΐνη:

GAAΨAAACΨAGΨAΨCΨΨCΨGGΨCCCCACAGACΨCAGAGAGAACCCGCCACC

Εδώ συναντάμε την πρώτη μας έκπληξη. Οι φυσιολογικοί χαρακτήρες RNA είναι A, C, G και U. Οι χαρακτήρες U είναι επίσης γνωστοί ως «T» στο DNA. Αλλά εδώ βρίσκουμε Ψ, τι συμβαίνει;

Αυτό είναι ένα από τα εξαιρετικά έξυπνα κομμάτια για το εμβόλιο. Το σώμα μας διαθέτει ένα ισχυρό σύστημα προστασίας από ιούς («το original»). Για αυτόν τον λόγο, τα κύτταρα δεν ενθουσιάζονται για το ξένο RNA και προσπαθούν πολύ σκληρά να το καταστρέψουν προτού κάνει κάτι.

Αυτό είναι κάπως πρόβλημα για το εμβόλιο μας - πρέπει να ξεφύγει από το ανοσοποιητικό μας σύστημα. Μετά από πολλά χρόνια πειραματισμού, διαπιστώθηκε ότι εάν το U στο RNA αντικατασταθεί από ένα ελαφρώς τροποποιημένο μόριο, το ανοσοποιητικό μας σύστημα χάνει το ενδιαφέρον. Στ' αλήθεια.

Έτσι, στο εμβόλιο BioNTech / Pfizer, κάθε U έχει αντικατασταθεί από 1-methyl-3'-pseudouridylyl, που υποδηλώνεται με Ψ. Το πραγματικά έξυπνο κομμάτι είναι ότι παρόλο που αυτή η αντικατάσταση Ψ καταπραΰνει (ηρεμεί) το ανοσοποιητικό μας σύστημα, είναι αποδεκτό ως κανονικό U από σχετικά μέρη του κυττάρου.

Στην ασφάλεια των υπολογιστών γνωρίζουμε επίσης αυτό το τέχνασμα - μερικές φορές είναι δυνατή η μετάδοση μιας ελαφρώς κατεστραμμένης έκδοσης ενός μηνύματος που μπερδεύει τα τείχη προστασίας και τις λύσεις ασφαλείας, αλλά αυτό εξακολουθεί να γίνεται αποδεκτό από τους διακομιστές backend - οι οποίοι στη συνέχεια μπορούν να παραβιαστούν.

Όπως και με άλλη θεμελιώδη επιστημονική έρευνα από την οποία τώρα επωφελούμαστε, οι ερευνητές της τεχνικής αυτής έπρεπε να παλέψουν για να χρηματοδοτηθεί το έργο τους και στη συνέχεια να εκδωθεί. Όλοι πρέπει να είμαστε πολύ ευγνώμονες και είμαι βέβαιος ότι τα βραβεία Νόμπελ θα φτάσουν εγκαίρως.

Εντάξει, ας επιστρέψουμε στο 5' UTR. Τι κάνουν αυτοί οι 51 χαρακτήρες; Όπως όλα στη φύση, σχεδόν τίποτα δεν έχει μια σαφή λειτουργία.

Όταν τα κύτταρα μας πρέπει να μεταφράσουν RNA σε πρωτεΐνες, αυτό γίνεται χρησιμοποιώντας μια μηχανή που ονομάζεται ριβόσωμα. Το ριβόσωμα είναι σαν τρισδιάστατος εκτυπωτής πρωτεϊνών. Καταλαμβάνει ένα σκέλος RNA και βασισμένο σε αυτό παράγει μια σειρά αμινοξέων, τα οποία στη συνέχεια διπλώνονται σε μια πρωτεΐνη.

Αυτό βλέπουμε να συμβαίνει παραπάνω. Η μαύρη κορδέλα στο κάτω μέρος είναι RNA. Η κορδέλα που εμφανίζεται στο πράσινο κομμάτι είναι η πρωτεΐνη που σχηματίζεται. Τα πράγματα που πετούν μέσα και έξω είναι αμινοξέα συν προσαρμογείς για να τα προσαρμόσουν στο RNA.

Αυτό το ριβόσωμα πρέπει να καθίσει πάνω στο σκέλος RNA για να μπορέσει να δουλέψει. Μόλις καθίσει, μπορεί να αρχίσει να σχηματίζει πρωτεΐνες με βάση το περαιτέρω RNA που καταναλώνει. Από αυτό, μπορείτε να φανταστείτε ότι δεν μπορεί ακόμη να διαβάσει τα μέρη όπου προσγειώνεται αρχικά. Αυτή είναι μόνο μία από τις λειτουργίες του UTR: η ζώνη προσγείωσης του ριβοσώματος. Το UTR παρέχει «lead-in».

Εκτός από αυτό, το UTR περιέχει επίσης μεταδεδομένα: Πότε πρέπει να γίνει μετάφραση; Πόσο; Για το εμβόλιο, πήραν το περισσότερο «τωρινό» UTR που μπορούσαν να βρουν, που ελήφθησαν από το γονίδιο άλφα σφαιρίνης. Αυτό το γονίδιο είναι γνωστό ότι παράγει σταθερά πολλές πρωτεΐνες. Τα προηγούμενα χρόνια, οι επιστήμονες είχαν ήδη βρει τρόπους να βελτιστοποιήσουν αυτό το UTR ακόμη περισσότερο (σύμφωνα με το έγγραφο του ΠΟΥ), επομένως αυτό δεν είναι ακριβώς το UTR άλφα σφαιρίνης. Είναι ακόμη καλύτερο.

Το πεπτίδιο σήματος S γλυκοπρωτεΐνης

Όπως σημειώθηκε, ο στόχος του εμβολίου είναι να αναγκάσει το κύτταρο να παράγει άφθονες ποσότητες της πρωτεΐνης Spike του SARS-CoV-2. Μέχρι αυτό το σημείο, συναντήσαμε ως επί το πλείστον στοιχεία μεταδεδομένων και “συμβάσεις κλήσεων” στον πηγαίο κώδικα εμβολίων. Αλλά τώρα μπαίνουμε στην πραγματική περιοχή των ιικών πρωτεϊνών.

Έχουμε ακόμη ένα επίπεδο μεταδεδομένων για να συνεχίσουμε. Μόλις το ριβόσωμα δημιουργήσει μια πρωτεΐνη, αυτή η πρωτεΐνη πρέπει να πάει κάπου. Αυτό κωδικοποιείται στο “πεπτίδιο σήματος S γλυκοπρωτεΐνης (εκτεταμένη αλληλουχία-οδηγός)”.

Ο τρόπος για να το δείτε είναι ότι στην αρχή της πρωτεΐνης υπάρχει ένα είδος ετικέτας διεύθυνσης - κωδικοποιημένο ως μέρος της ίδιας της πρωτεΐνης. Σε αυτή τη συγκεκριμένη περίπτωση, το πεπτίδιο σήματος λέει ότι αυτή η πρωτεΐνη πρέπει να εξέλθει από το κύτταρο μέσω του «ενδοπλασματικού δικτύου». Ακόμα και η διάλεκτος Star Trek δεν είναι τόσο φανταχτερή όσο αυτό!

Το «πεπτιδικό σήμα» δεν είναι πολύ μεγάλο, αλλά όταν κοιτάξουμε τον κώδικα, υπάρχουν διαφορές μεταξύ του ιικού RNA και του εμβολίου:

(Σημειώστε ότι για λόγους σύγκρισης, αντικατέστησα το τροποποιημένο Ψ με ένα κανονικό RNA U)

3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

Virus: AUG UUU GUU UUU CUU GUU UUA UUG CCA CUA GUC UCU AGU CAG UGU GUU

Vaccine: AUG UUC GUG UUC CUG GUG CUG CUG CCU CUG GUG UCC AGC CAG UGU GUU

! ! ! ! ! ! ! ! ! ! ! ! ! ! !

Τι συμβαίνει λοιπόν; Δεν έχω καταχωρήσει κατά λάθος το RNA σε ομάδες 3 γραμμάτων. Τρεις χαρακτήρες RNA αποτελούν ένα κωδικόνιο. Και κάθε κωδικόνιο κωδικοποιεί ένα συγκεκριμένο αμινοξύ. Το πεπτίδιο σήμα στο εμβόλιο αποτελείται από ακριβώς τα ίδια αμινοξέα όπως στον ίδιο τον ιό.

Γιατί, λοιπόν, είναι το RNA διαφορετικό;

Υπάρχουν $4^3 = 64$ διαφορετικά κωδικόνια, δεδομένου ότι υπάρχουν 4 χαρακτήρες RNA και υπάρχουν τρία από αυτά σε ένα κωδικόνιο. Ωστόσο, υπάρχουν μόνο 20 διαφορετικά αμινοξέα. Αυτό σημαίνει ότι πολλαπλά κωδικόνια κωδικοποιούν το ίδιο αμινοξύ.

Η ζωή χρησιμοποιεί τον ακόλουθο σχεδόν καθολικό πίνακα για τη χαρτογράφηση κωδικώνων RNA σε αμινοξέα:

1st base	2nd base				3rd base
	U	C	A	G	
U	UUU (Phe) Phenylalanine ↑	UCU	UAU (Tyr) Tyrosine ↑	UGU (Cys) Cysteine ↑	U
	UUC	UCC (Ser) Serine ↑	UAC	UGC	C
	UUA	UCA	UAA Stop (ochre) ↓	UGA Stop (opal) ↓	A
	UUG (Leu) Leucine ↑	UCG	UAG Stop (amber) ↓	UGG (Trp) Tryptophan ↑	G
C	CUU	CCU	CAU	CGU	U
	CUC	CCC	CAC (His) Histidine ↑	CGC	C
	CUA	CCA	CAA	CGA	A
	CUG	CCG	CAG (Gln) Glutamine ↑	CGG	G
A	AUU	ACU	AUU	AGU	U
	AUC (Ile) Isoleucine ↑	ACC	AAC (Asn) Asparagine ↑	AGC	C
	AUA	ACA	AAA	AGA	A
	AUG (Met) Methionine ↑	ACG	AAG	AGG	G
G	GUU	GCU	GAU	GGU	U
	GUC	GCC	GAC (Asp) Aspartic acid ↓	GGC	C
	GUA	GCA	GAA	GGA	A
	GUG	GCG	GAG (Glu) Glutamic acid ↓	GGG	G

Σε αυτόν τον πίνακα, μπορούμε να δούμε ότι οι τροποποιήσεις στο εμβόλιο (UUU -> UUC) είναι όλες *συνώνυμες*. Ο κωδικός RNA εμβολίου είναι διαφορετικός, αλλά παράγονται τα ίδια αμινοξέα και η ίδια πρωτεΐνη.

Αν κοιτάξουμε προσεκτικά, βλέπουμε ότι η πλειονότητα των αλλαγών συμβαίνει στην τρίτη θέση κωδικονίου, όπως σημειώθηκε με το «3» παραπάνω. Και αν ελέγξουμε τον γενικό πίνακα κωδικονίων, βλέπουμε ότι αυτή η τρίτη θέση συχνά δεν έχει σημασία για ποιο αμινοξύ θα παραχθεί.

Έτσι, οι αλλαγές είναι συνώνυμες, αλλά γιατί είναι εκεί; Κοιτάζοντας προσεκτικά, βλέπουμε ότι όλες οι αλλαγές εκτός από μία οδηγούν σε περισσότερους χαρακτήρες C και G.

Γιατί λοιπόν να το κάνεις αυτό; Όπως προαναφέρθηκε, το ανοσοποιητικό μας σύστημα έχει μια πολύ αμυδρή εικόνα του «εξωγενούς» κωδικού RNA, δηλαδή RNA που προέρχεται από το εξωτερικό του κυττάρου. Για να αποφύγει την ανίχνευση, το «U» στο RNA αντικαταστάθηκε ήδη από Ψ.

Ωστόσο, αποδεικνύεται ότι το RNA με υψηλότερη ποσότητα χαρακτήρων G και C μετατρέπεται πιο αποτελεσματικά σε πρωτεΐνες.

Και αυτό επιτεύχθηκε στο εμβόλιο RNA αντικαθιστώντας πολλούς χαρακτήρες με G και C, όπου αυτό ήταν δυνατό.

Είμαι λίγο γοητευμένος από τη μία αλλαγή που δεν οδήγησε σε επιπλέον C ή G, την τροποποίηση CCA -> CCU. Αν κάποιος γνωρίζει τον λόγο, ενημερώστε με!

Η πραγματική πρωτεΐνη Spike

Οι επόμενοι 3777 χαρακτήρες του εμβολίου RNA είναι ομοίως «βελτιστοποιημένο κωδικόνιο» για να προσθέσουν πολλά C και G. Για να μην επεκταθώ πάρα πολύ, δεν θα αναφέρω όλους τους κωδικούς εδώ, αλλά θα κάνουμε μεγέθυνση σε ένα εξαιρετικά ξεχωριστό κομμάτι. Αυτό είναι το κομμάτι που το κάνει να λειτουργεί, το κομμάτι που πραγματικά θα μας βοηθήσει να επιστρέψουμε στη ζωή ως συνήθως:

* *

L D K V E A E V Q I D R L I T G

Virus: CUU GAC AAA GUU GAG GCU GAA GUG CAA AUU GAU AGG UUG AUC ACA GGC

Vaccine: CUG GAC CCU CCU GAG GCC GAG GUG CAG AUC GAC AGA CUG AUC ACA GGC

L D P P E A E V Q I D R L I T G

! !!!!! ! ! ! ! ! !

Εδώ βλέπουμε τις συνήθεις συνώνυμες αλλαγές στο RNA. Για παράδειγμα, στο πρώτο κωδικόνιο βλέπουμε ότι το CUU έχει αλλάξει σε CUG. Αυτό προσθέτει ένα άλλο «G» στο εμβόλιο, το οποίο γνωρίζουμε ότι βοηθά στην ενίσχυση της

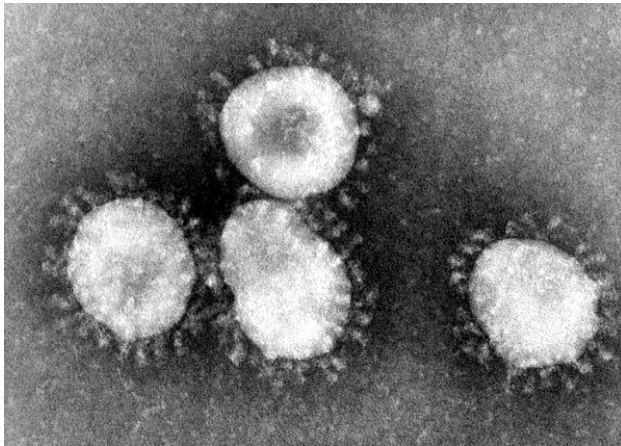
παραγωγής πρωτεϊνών. Τόσο το CUU όσο και το CUG κωδικοποιούν το αμινοξύ «L» ή τη λευκίνη, οπότε τίποτα δεν άλλαξε στην πρωτεΐνη.

Όταν συγκρίνουμε ολόκληρη την πρωτεΐνη Spike στο εμβόλιο, όλες οι αλλαγές είναι συνώνυμες όπως αυτή.. εκτός από δύο, και αυτό βλέπουμε εδώ.

Το τρίτο και το τέταρτο κωδικόνιο παραπάνω αντιπροσωπεύουν πραγματικές αλλαγές. Τα K και V αμινοξέα εκεί αντικαθίστανται και τα δύο από «P» ή Proline. Για το “K” απαιτούνται τρεις αλλαγές (“!!!”) και για το “V” απαιτούνται μόνο δύο (“!!”).

Αποδεικνύεται ότι αυτές οι δύο αλλαγές βελτιώνουν την αποτελεσματικότητα του εμβολίου !

Τι συμβαίνει λοιπόν εδώ; Εάν κοιτάξετε ένα πραγματικό σωματίδιο SARS-CoV-2, μπορείτε να δείτε την πρωτεΐνη Spike, καθώς και πολλές αιχμές:



Οι αιχμές τοποθετούνται στο σώμα του ιού («η πρωτεΐνη νουκλεοκαψιδίου»). Αλλά το ζήτημα είναι ότι, το εμβόλιο μας δημιουργεί μόνο τις αιχμές και δεν τοποθετούνται σε κανένα σώμα ιού.

Αποδεικνύεται ότι, οι μη τροποποιημένες, ανεξάρτητες πρωτεΐνες Spike δημιουργούν μια διαφορετική δομή. Εάν εγχυθεί ως εμβόλιο, αυτό θα έκανε πράγματι το σώμα μας να αναπτύξει ανοσία.. αλλά μόνο ενάντια στην διαφορετική δομή.

Και το πραγματικό SARS-CoV-2 εμφανίζεται με το ακιδωτό Spike. Το εμβόλιο δεν θα λειτουργούσε πολύ καλά σε αυτήν την περίπτωση.

Οπότε τί κάνουμε; Το 2017 περιγράφηκε πώς η τοποθέτηση διπλής υποκατάστασης Proline στο σωστό μέρος θα έκανε τις πρωτεΐνες SARS-CoV-1 και MERS S να πάρουν τη διαμόρφωση «προ-σύντηξης», ακόμη και χωρίς να είναι μέρος ολόκληρου του ιού. Αυτό λειτουργεί επειδή η προλίνη είναι ένα πολύ άκαμπτο αμινοξύ. Λειτουργεί

ως ένα είδος νάρθηκα, σταθεροποιώντας την πρωτεΐνη στην κατάσταση που πρέπει να δείξουμε στο ανοσοποιητικό σύστημα.

Το τέλος της πρωτεΐνης, τα επόμενα βήματα

Εάν κυλίσουμε στον υπόλοιπο πηγαίο κώδικα, θα συναντήσουμε μερικές μικρές τροποποιήσεις στο τέλος της πρωτεΐνης Spike:

```

V L K G V K L H Y T s
Virus: GUG CUC AAA GGA GUC AAA UUA CAU UAC ACA UAA
Vaccine: GUG CUG AAG GGC GUG AAA CUG CAC UAC ACA UGA UGA
V L K G V K L H Y T s s
! ! ! ! ! ! !

```

Στο τέλος μιας πρωτεΐνης συναντάμε ένα κωδικόνιο «στοπ», που υποδηλώνεται εδώ με πεζό s. Αυτός είναι ένας ευγενικός τρόπος για να πούμε ότι η πρωτεΐνη πρέπει να τελειώσει εδώ. Ο αρχικός ιός χρησιμοποιεί το κωδικόνιο διακοπής UAA, το εμβόλιο χρησιμοποιεί δύο κωδικόνια διακοπής UGA, ίσως μόνο απλά έτσι, για το καλό.

Η 3' αμετάφραστη περιοχή

Όπως όταν το ριβόσωμα χρειαζόταν κάποιο lead-in στο 5' άκρο, βρήκαμε την «5' αμετάφραστη περιοχή», στο τέλος μιας πρωτεΐνης βρίσκουμε ένα παρόμοιο κατασκεύασμα που ονομάζεται 3' UTR.

Πολλές λέξεις θα μπορούσαν να γραφτούν για το 3' UTR, αλλά εδώ παραθέτω αυτό που λέει η Wikipedia: «Η 3' αμετάφραστη περιοχή παίζει καθοριστικό ρόλο στην έκφραση γονιδίων επηρεάζοντας την απόδοση εντοπισμού, σταθερότητας, εξαγωγής και μετάφρασης ενός mRNA.. **παρά την τρέχουσα κατανόησή μας για τα 3' UTR, εξακολουθούν να είναι σχετικά μυστήρια** ».

Αυτό που γνωρίζουμε είναι ότι ορισμένα 3' UTR είναι πολύ αποδοτικά στην προώθηση της πρωτεϊνικής έκφρασης. Σύμφωνα με το έγγραφο του ΠΟΥ, το εμβόλιο BioNTech / Pfizer 3'-UTR επιλέχθηκε από τον «αμινο-τερματικό ενισχυτή του διαχωριζόμενου (AES) mRNA και το μιτοχονδριακό κωδικοποιημένο 12S ριβοσωμικό RNA για να προσδώσει σταθερότητα στο RNA και υψηλή ολική έκφραση πρωτεΐνης».

Το τέλος όλων το AAAAAAAAAAAAAAAAAAAAAAAAAA

Το άκρο του mRNA είναι πολυαδενυλιωμένο. Αυτός είναι ένας τρόπος για να πούμε ότι τελειώνει σε πολλά AAAAAAAAAAAAAAAAAA.

Το mRNA μπορεί να επαναχρησιμοποιηθεί πολλές φορές, αλλά καθώς συμβαίνει αυτό, χάνει επίσης μερικά από τα A στο τέλος. Μόλις τελειώσουν τα A, το mRNA δεν είναι πλέον λειτουργικό και απορρίπτεται. Με αυτόν τον τρόπο, η ουρά «poly-A» προστατεύει από την απόρριψη.

Έχουν γίνει μελέτες για να διαπιστωθεί ποιος είναι ο βέλτιστος αριθμός A στο τέλος για τα εμβόλια mRNA. Διάβασα στην ανοιχτή βιβλιογραφία ότι αυτό κορυφώθηκε στα 120 περίπου.

Το εμβόλιο BNT162b2 τελειώνει με:

***** **

UAGCAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGCAUAU GACUAAAAAA
AAAAAAAAAA

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA

Αυτά είναι 30 A, μετά ένας «σύνδεσμος 10 νουκλεοτιδίων» (GCAUAUGACU), ακολουθούμενος από άλλα 70 A.

Υποψιάζομαι ότι αυτό που βλέπουμε εδώ είναι το αποτέλεσμα της περαιτέρω ιδιοκτησιακής βελτιστοποίησης για την περαιτέρω ενίσχυση της έκφρασης πρωτεΐνης.

Συνοψίζοντας

Με αυτό, γνωρίζουμε τώρα το ακριβές περιεχόμενο mRNA του εμβολίου BNT162b2 και για τα περισσότερα μέρη καταλαβαίνουμε γιατί είναι εκεί:

- Το CAP για να βεβαιωθούμε ότι το RNA μοιάζει με κανονικό mRNA
- Μια γνωστά επιτυχημένη και βελτιστοποιημένη 5' αμετάφραστη περιοχή (UTR)
- Ένα πεπτίδιο σήματος βελτιστοποιημένο με κωδικόνιο για να στείλει την πρωτεΐνη Spike στη σωστή θέση (αντιγράφηκε 100% από τον αρχικό ιό)
- Με κωδικόνιο βελτιστοποιημένη έκδοση της αρχικής ακίδας, με δύο αντικαταστάσεις «Proline» για να βεβαιωθούμε ότι η πρωτεΐνη εμφανίζεται στη σωστή μορφή
- Μια γνωστά επιτυχημένη και βελτιστοποιημένη 3' αμετάφραστη περιοχή
- Μια ελαφρώς μυστηριώδης ουρά poly-A με έναν ανεξήγητο «σύνδεσμο» εκεί

Η βελτιστοποίηση κωδικονίων προσθέτει πολλά G και C στο mRNA. Εν τω μεταξύ, η χρήση Ψ (1-methyl-3'-pseudouridylyl) αντί του U βοηθά στην αποφυγή του ανοσοποιητικού μας συστήματος, οπότε το mRNA παραμένει αρκετό χρόνο ώστε να μπορεί πραγματικά να βοηθήσουμε στην εκπαίδευση του ανοσοποιητικού συστήματος.

Μετάφραση από:

Previous Previous post: Reverse Engineering the source code of the BioNTech/Pfizer SARS-CoV-2 Vaccine

Φροντιστήρια Βακάλη