

Genetic Variation in Eukaryotes

1. There is abundant variation in gene expression between individuals, populations, and species.
2. Genetic variation within a eukaryotic population/species is caused by meiosis and DNA sequence errors (single, indels), copy number alterations (structural variants), genome duplication, transposition
3. Genetic Richness is mostly fuelled by CNA, genome duplication, and meiosis in eukaryotes.
4. Most of the DNA sequence errors and CNVs are under neutral or balanced selection and usually happen in non-coding regions.
5. SNPs (*in cis* or *in trans* factors) variation is stabilizing/balancing the newly acquired genetic information (survival adaptation)
6. Stress-induced (epi)genetic variation (transposition, ncRNAs)

Η απαρχή της Ζωής βασίζεται στο RNA

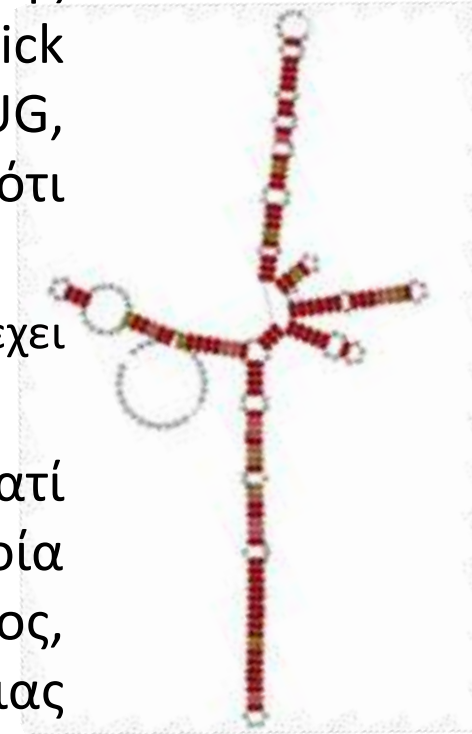
- Η Αδενίνη (η πρώτη νουκλεϊκή βάση που ενδεχομένως συντέθηκε στην προβιοτική σούπα), η μεθιονίνη (το πρώτο αμινοξύ που προστέθηκε κατά την μετάφραση) και η ριβόζη (το πρώτο σάκχαρο που δημιουργήθηκε κατά τον αναβολισμό νουκλεϊκών οξέων).
- Η προστασία των μορίων RNA σε ένα εξωφρενόφιλο περιβάλλον ήταν προαπαιτούμενη για τη δημιουργία βιοπολυμερών RNA
- Η Μεθυλίωση μέσω μικρών RNA ενδεχομένως να αποτέλεσαν τον πρώτο μηχανισμό προστασίας των RNA από υδρόλυση (στις μέρες μας αυτό γίνεται από τα ένζυμα μεθυλτρασφερασών)
- Αργότερα με την εμφάνιση στα RNA ικανοτήτων αντιγραφής και μετάφρασης πεπτιδίων, εμφανίστηκε ο καθολικά εφαρμοσμένος γενετικός κώδικας και εμφανίστηκαν οι πρωτεΐνες και το DNA.

Θερμοδυναμική και σταθερότητα/λειτουργία του δίκλωνου RNA

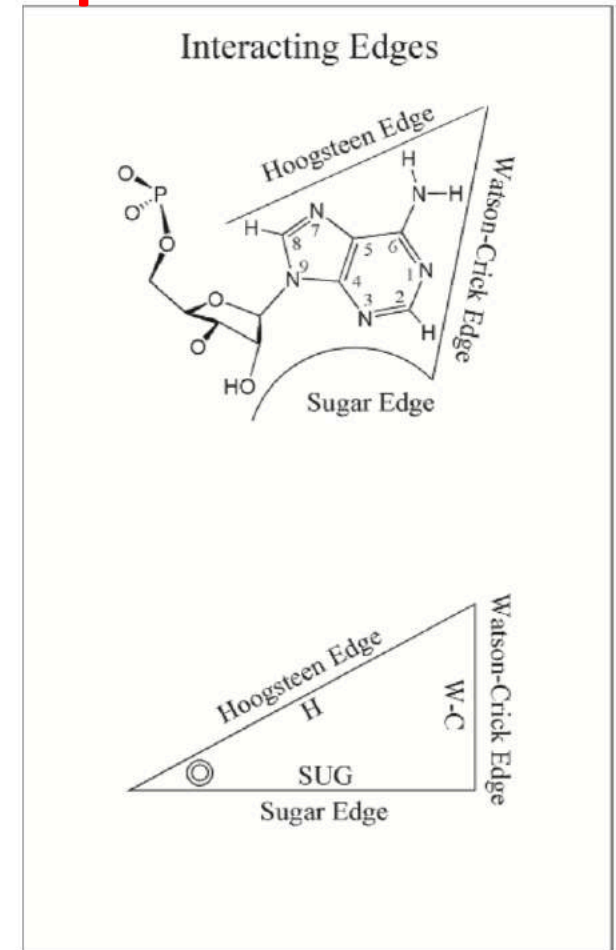
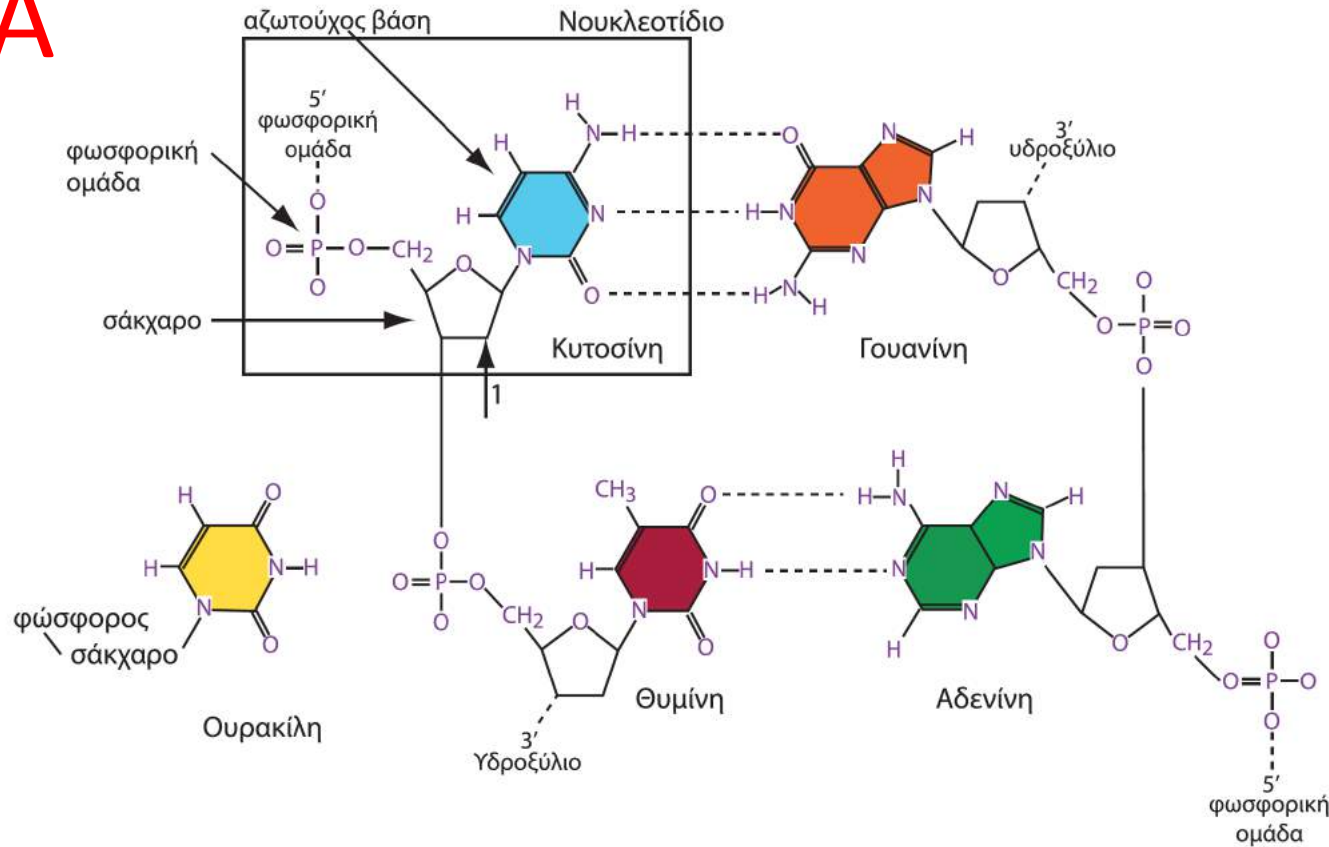
- Η αναδίπλωση του RNA είναι θερμοδυναμικά αναπόφευκτη-
- Οποιοδήποτε τυχαία αλληλουχία RNA λαμβάνει συμπαγείς δομές παρά γραμμικές αλυσίδες.
- Κάτω από φυσιολογικές συνθήκες ευνοείται ενεργειακά το ζευγάρωμα βάσεων WC και άρα η αναδίπλωση του RNA σε δευτεροταγείς και στη συνέχεια τριτοταγείς δομές
- Ένα RNA μπορεί να προσδένει άλλα RNAs δημιουργώντας τεταρτοταγείς δομές καθώς και μικρά μόρια, πρωτεΐνες και DNA προκειμένου να σχηματίζει από δίκλωνες μορφές έως τεραστια μοριακά εργοστάσια ή μηχανές όπως το ριβόσωμα και το σωματίο ματίσματος

Έλικες Watson–Crick και Δευτεροταγής Δομή

- Μονόκλινα μόρια RNA αναδιπλώνονται στους εαυτούς τους με αποτέλεσμα να αντιπαραβληθούν συμπληρωματικές βάσεις με αντι-παράλληλη φορά (νοηματική/αντι-νοηματική)
- Αυτό δημιουργεί έλικες Watson–Crick που είναι οι θεμελιώδεις μονάδες της δευτεροταγούς δομής του RNA. Οι έλικες αποτελούνται από κανονικά Watson–Crick (WC) ζεύγη βάσεων AU, UA, CG, and GC, και επίσης τα μη-κανονικά WC ζεύγη GU,UG, και άλλα. Αρχικά το ζεύγος G·U ονομαζόταν wobble γιατί ο Crick το 1966 υπέθεσε ότι ευθύνεται για το φαινόμενο ταλάντωσης (wobble effect).
 - Το ζεύγος βάσεων G·U είναι συνήθως εξαιρετικά συντηρημένο και η υποκατάστασή του έχει συχνά καταστροφικές συνέπειες για τη λειτουργία (και τη δομή) του RNA.
- Οι έλικες είναι συνήθως βραχείες (όχι μακρύτερες από 10-15 ζεύγη WC) γιατί παρεμβάλλονται ή τερματίζονται από αζευγάρωτα τμήματα αλληλουχιών, τα οποία ονομάζονται ανάλογα με τον τόπο που σχηματίζονται, φουρκέτα, εσωτερικός βρόχος, πολυ-ελικωτή ένωση «βρόχων» ή ψευδοβρόχος (μεταξύ ενός βρόχου και μιας εξωτερικής περιοχής).
- Μεταλλάξεις σε ζεύγη βάσεων που διαταράσσουν τη δομή/λειτουργία του RNA είναι λιγότερο πιθανό να κληρονομηθούν εξαιτίας της φυσικής επιλογής.



Κανονικά και μη-κανονικά ζευγαρώματα βάσεων του RNA

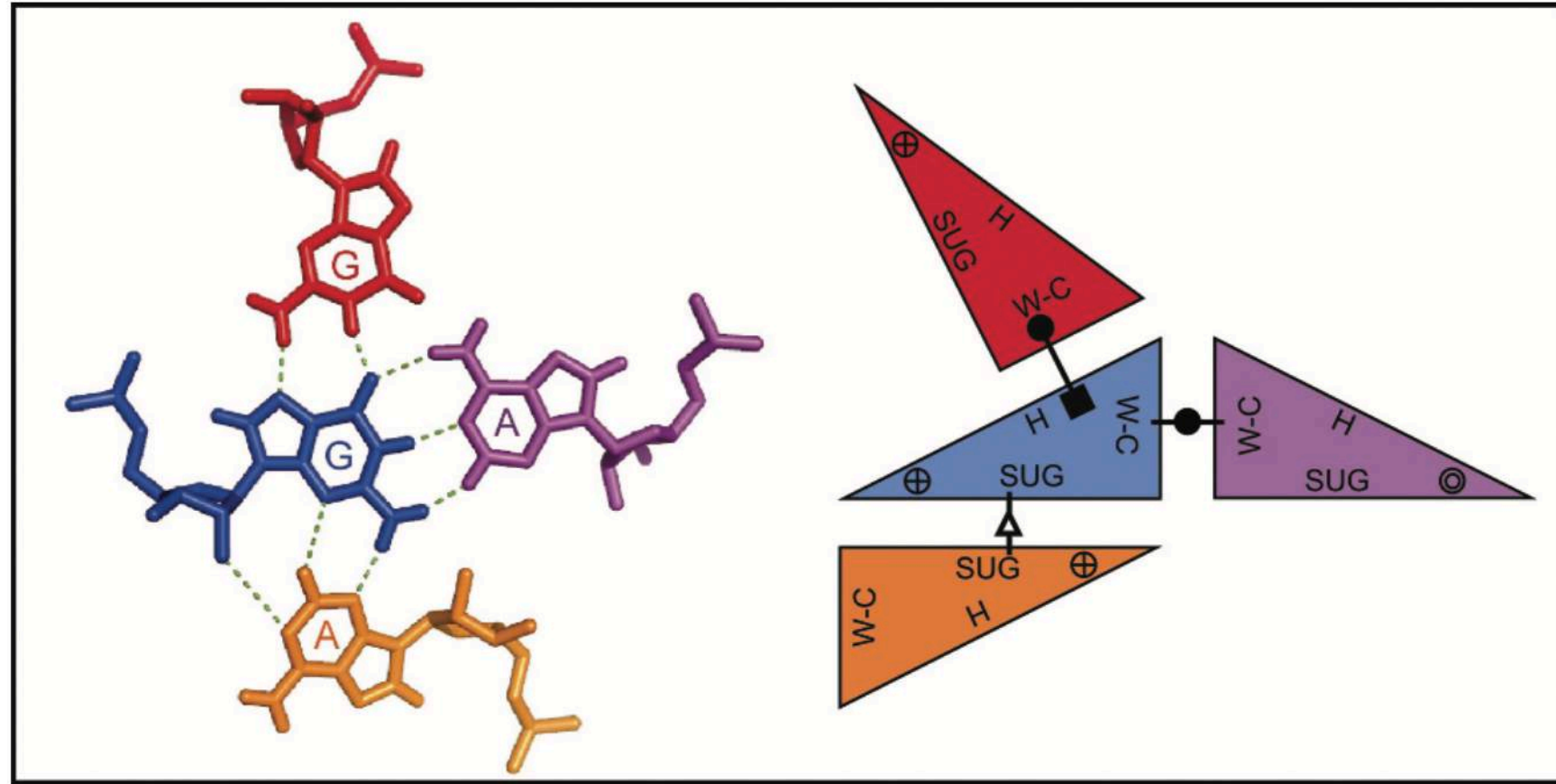


- Οι αζωτούχες βάσεις των νουκλεοτιδίων είναι οι πουρίνες γουανίνη και αδενίνη και οι πυριμιδίνες θυμίνη, ουρακίλη (η οποία αντικαθιστά τη θυμίνη στο RNA) και κυτοσίνη. Αυτές οι αζωτούχες βάσεις συνδέονται με το σάκχαρο ριβόζη, στο οποίο συνδέεται επίσης μία φωσφορική ομάδα. Στην περίπτωση του DNA, από τη ριβόζη λείπει μία ομάδα υδροξυλίου (κόκκινο βέλος) που υπάρχει στο RNA, η οποία είναι πολύ ευλύγιστη αλλά διαδραματίζει σημαντικό ρόλο στη σταθεροποίηση της δομής των RNA.
- Έτσι παρόλο που η πρωτογενής διαμόρφωση των νουκλεοτιδίων ή δι-νουκλεοτιδίων (GC, AU) του RNA δεν είναι τόσο σταθερή όσο του DNA μπορεί να κατηγοριοποιηθεί με βάση επαναλαμβανόμενα πρότυπα που συσχετίζονται με συγκεκριμένα μοτίβα

- Η δομή της Αδενοσίνης απεικονίζοντας τις τρεις άκρες των τριών επιπέδων αλληλεπιδράσεων του RNA (ως τριγώνου) - Watson-Crick, Hoogsteen and Sugar-edge

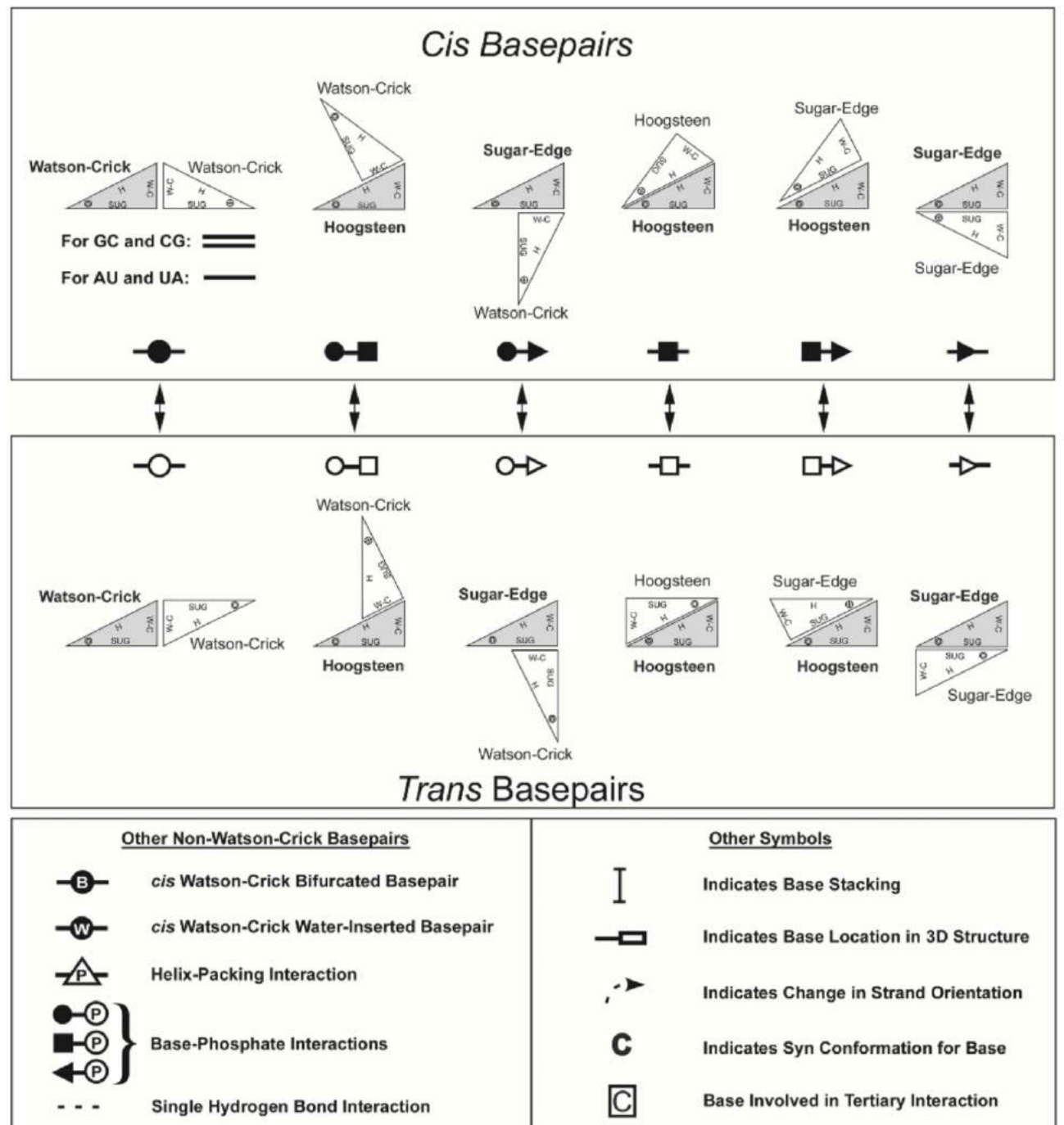
Οι Βάσεις του RNA μπορεί να ζευγαρώνουν με τρία διαφορετικά νουκλεοτίδια με τρεις διαφορετικούς τρόπους

- Οι βάσεις του RNA αλληλοεπιδρούν με τρεις διαφορετικές άκρες, μέσω δεσμών, Watson–Crick (WC), Hoogsteen (H) και Sugar (Sug), σχηματίζοντας «κορεσμένες» τετραπλές βάσεις



Οι κατηγορίες ζευγαρώματος βάσεων RNA

- Οι 12 οικογένειες γεωμετρικών ζευγαρώματων βάσεων απεικονίζονται με τρίγωνα. Οι κύκλοι αντιστοιχούν σε δεσμούς Watson-Crick, τα τετράγωνα δεσμούς Hoogsteen και τα τρίγωνα, δεσμούς σάκχαρων (Sugar edges). Οι κατηγορίες ζευγαρώματος βάσεων σχηματίζονται από τα σύμβολα των δεσμών και τα σύμβολα που υποδεικνύουν το επίπεδο ζευγαρώματος εάν είναι in cis (σκούρο χρώμα) in trans.
- Στο κάτω δεξιό πλαίσιο εξηγούνται κάποιιοι επιπλέον τρόποι ζευγαρώματος του RNA



Δομικές vs Λειτουργικές κατηγορίες μοτίβων RNA

- Τα μοτίβα RNA κατηγοριοποιούνται δομικά προκειμένου να αναγνωριστούν γεωμετρικά παρόμοια μοτίβα ή λειτουργικά για τη ταυτοποίηση παρόμοιων μοτίβων με την ίδια λειτουργία
- Επαναλαμβανόμενα δομικά μοτίβα RNA έχουν το ίδιο ή παρόμοιο λειτουργικό ρόλο σε διαφορετικά μόρια RNA ή σε διαφορετικά μέρη του ίδιου μορίου RNA και επομένως παρέχουν πληροφορία για το τρόπο που το RNA αναδιπλώνεται στο χώρο και λειτουργεί.
- Τα δομικά μοτίβα μπορεί να έχουν την ίδια λειτουργία ακόμα και εάν τα πρότυπα αλληλουχιών από τα οποία αποτελείται είναι διαφορετικά σε σύσταση ή αριθμό νουκλεοτιδίων. -> Διαφορετικά μοτίβα μπορεί να έχουν την ίδια λειτουργία.
- **Οι λειτουργικοί ρόλοι των μοτίβων RNA ομαδοποιούνται κυρίως σε αρχιτεκτονικά, σταθεροποιητικά της δομής, πρόσδεσης προσδετών ή καταλυτικά μοτίβα.**

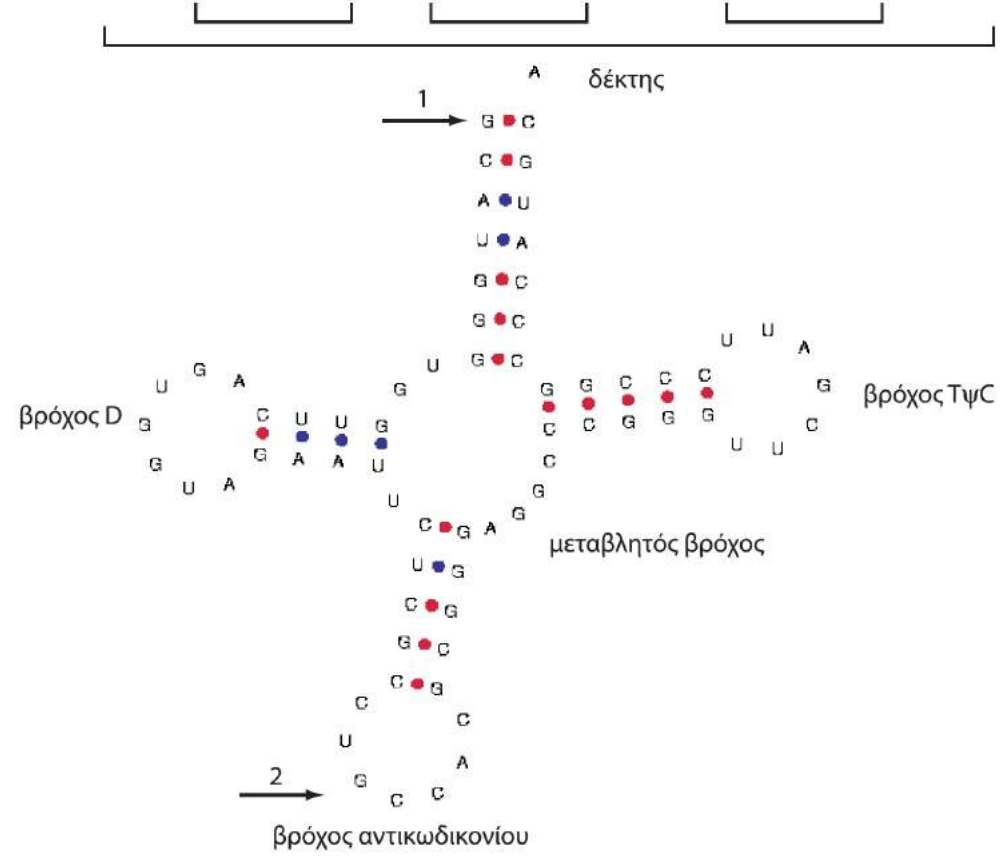
Μεταφορικό RNA

- Στο ανθρώπινο γονιδίωμα υπάρχουν πάνω από 600 γονίδια tRNA
- Τα μεταφορικά RNAs (tRNAs) είναι μήκους 73-90 nts και κουβαλάνε ένα συγκεκριμένο αμινοξύ που ταιριάζει με το αντίστοιχο κωδικόνιο σε ένα mRNA κατά την πρωτεϊνοσύνθεση.
- Το tRNA δημιουργεί μια δομή 70–90 νουκλεοτιδίων αναδιπλωμένων σε μια χαρακτηριστική δομή «τριφύλλι». Τα κύρια χαρακτηριστικά αυτής της δομής, περιλαμβάνουν ένα βρόγχο D, ένα βρόγχο αντικωδικονίου που είναι υπεύθυνο για την αναγνώριση των κωδικονίων στο mRNA, ένα βρόγχο T, και ένα 3' άκρο στο οποίο προσκολλάται το κατάλληλο αμινοξύ για το καθένα tRNA.

(β)

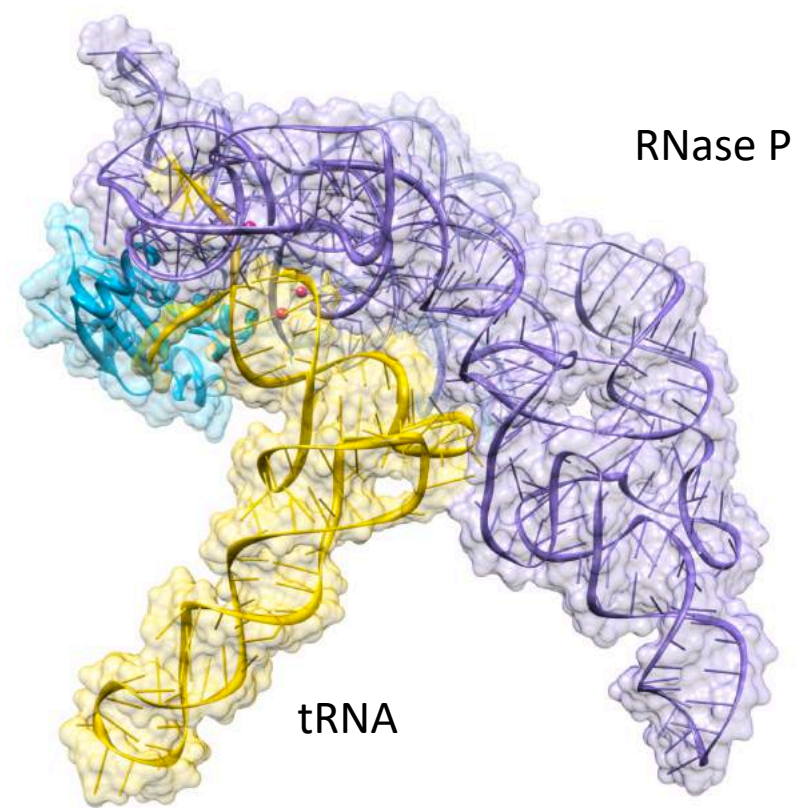
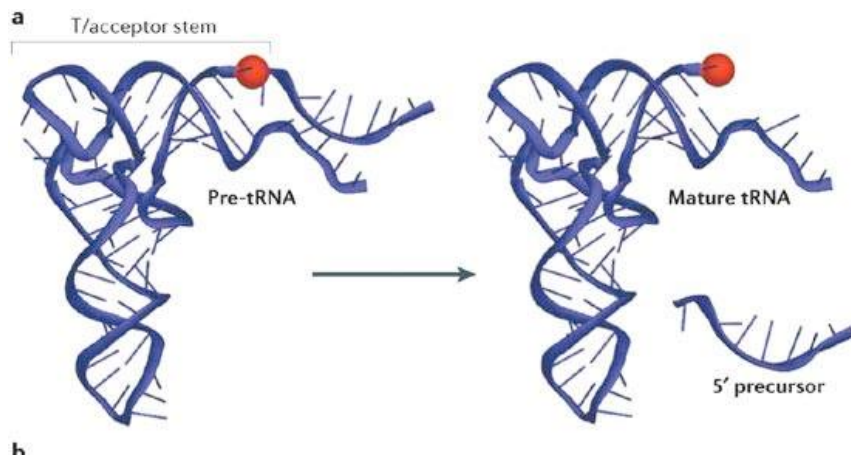
```

Your-seq.trna1 (1-71)   Length: 71 bp
Type: Gly              Anticodon: GCC at 33-35 (33-35) Score: 71.03
                       *     *     *     *     *     *     *     *
Seq: GCATGGGTGGTTCAGTGGTAGAATTCTCGCCTGCCACGCGGGAGGCCCGGGTTCGATTCCCGGCCCATGCA
Str: >>>>>>...>>>>.....<<<<<<>>>>>>.....<<<<<<.....>>>>>>.....<<<<<<<<<<<<.
  
```



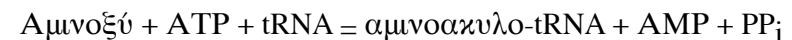
Το Καταλυτικό RNA

- Πολλοί τύποι αντιδράσεων καταλύονται από μόρια RNA
- Η λέξη ριβοένζυμο (ribozyme) χρησιμοποιείται ως ένας γενικός όρος που περιγράφει ένα μόριο RNA.
- Η ενζυμική ενεργότητα των ριβοενζύμων χαρακτηρίζεται με τον ίδιο τρόπο που χαρακτηρίζεται και στα «συμβατικά» πρωτεϊνικά ένζυμα.
- Το υπόστρωμα μια αντίδρασης που καταλύεται από ένα ριβοένζυμο μπορεί να είναι το ίδιο το ριβοένζυμο ή άλλα μόρια RNA.

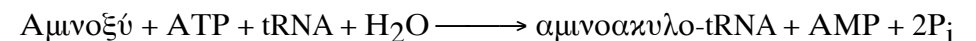


Το επόμενο βήμα είναι η μεταφορά της αμινοακυλικής ομάδας του αμινοακυλο-AMP σε ένα ειδικό μόριο tRNA για να σχηματιστεί ένα αμινοακυλο-tRNA.

Αμινοακυλο-AMP + tRNA = αμινοακυλο-tRNA + AMP Το άθροισμα αυτών των βημάτων ενεργοποίησης και μεταφοράς είναι:



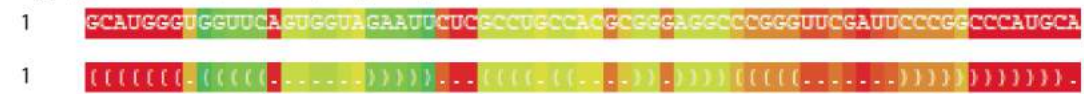
Η ΔG^0 της αντίδρασης αυτής πλησιάζει το 0, διότι η ελεύθερη ενέργεια της υδρόλυσης του εστερικού δεσμού του αμινοακυλο-tRNA είναι παρόμοια με εκείνη της υδρόλυσης της ATP σε AMP και PP_i . Όπως έχουμε δει πολλές φορές, μια αντίδραση ωθείται από την υδρόλυση πυροφωσφορικού. Το άθροισμα αυτών των τριών αντιδράσεων είναι εξαιρετικά εξώεργο:



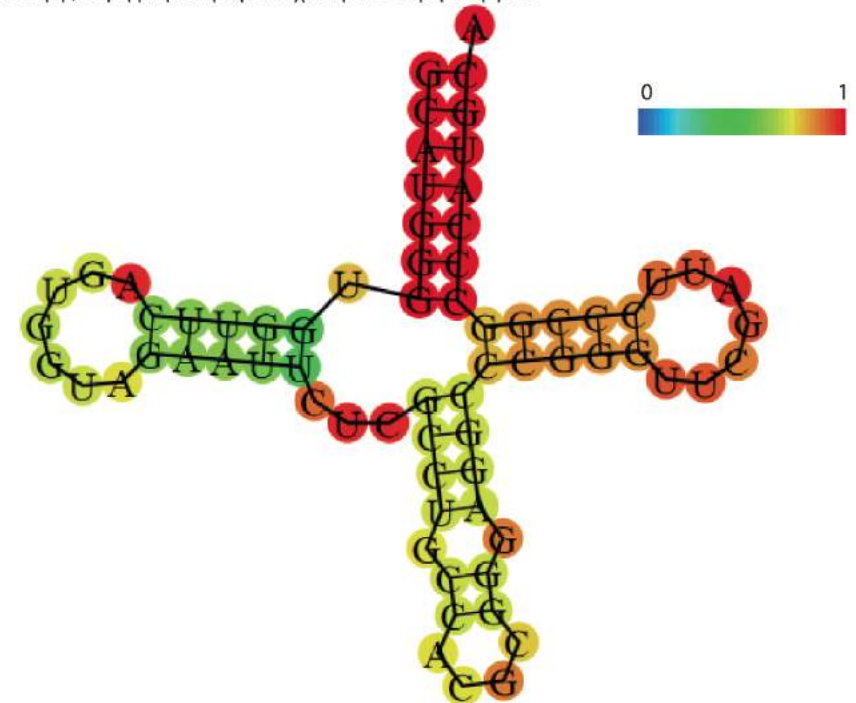
Πρόβλεψη δομής RNA βασισμένη στην ελάχιστη ελεύθερη ενέργεια αναδίπλωσης

- Πρόβλεψη της δομής του RNA με βάση τον προσδιορισμό της θερμοδυναμικά ευνοούμενης αναδίπλωσης με την ελάχιστη ελεύθερη ενέργεια. Παρουσιάζονται τα αποτελέσματα της ανάλυσης του tRNA του χρωμοσώματος 21 με το πακέτο RNA της Βιέννης. (α) Συμβολική αποτύπωση της βέλτιστης προβλεπόμενης δευτεροταγούς δομής του tRNA με το σύστημα κωδικοποίησης παρένθεσης-κουκκίδας. Τα μη ζευγαρωμένα νουκλεοτίδια αναπαρίστανται με κουκκίδες, ενώ τα ζευγαρωμένα νουκλεοτίδια αναπαρίστανται με ένα ζεύγος παρενθέσεων. Η ελάχιστη ελεύθερη ενέργεια υπολογίστηκε ότι είναι $-35,96$ kcal/mol. (β) Μοντέλο της δομής του tRNA στο οποίο φαίνονται τα στελέχη (δίκλωνες περιοχές με ζεύγος βάσεων) και οι βρόχοι (μονόκλωνες περιοχές).

(α) Πρόβλεψη με βάση την ελάχιστη ελεύθερη ενέργεια (η πιθανότητα ζευγαρώματος υποδεικνύεται με έναν χρωματικό κώδικα): ανάλυση ενός tRNA με το πακέτο Vienna 2.0

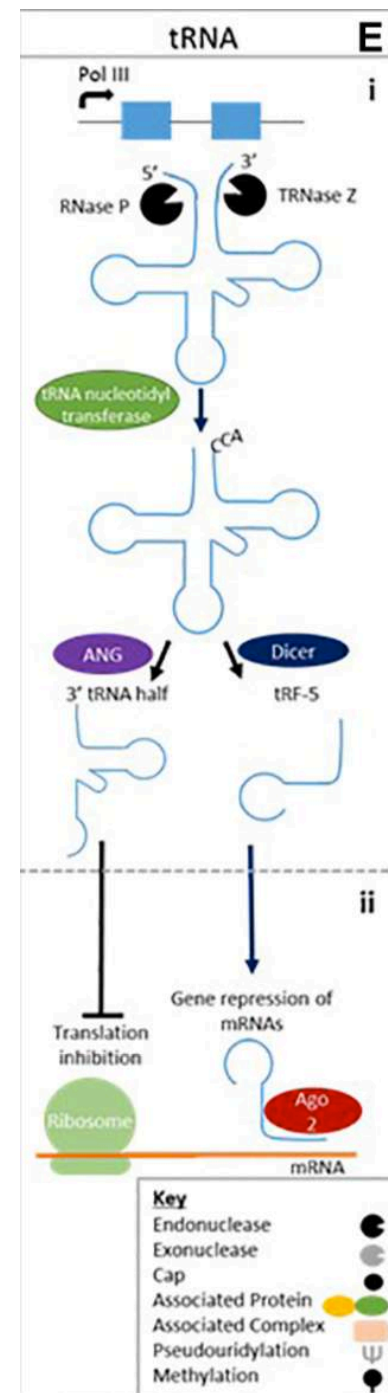


(β) Δευτεροταγής δομή με βάση την ελάχιστη ελεύθερη ενέργεια



Μεταφορικό RNA

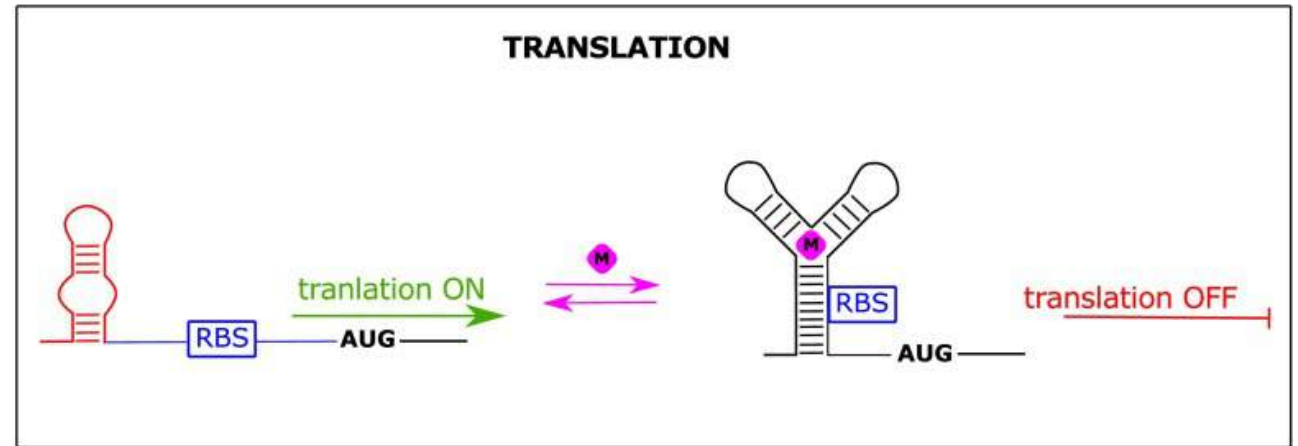
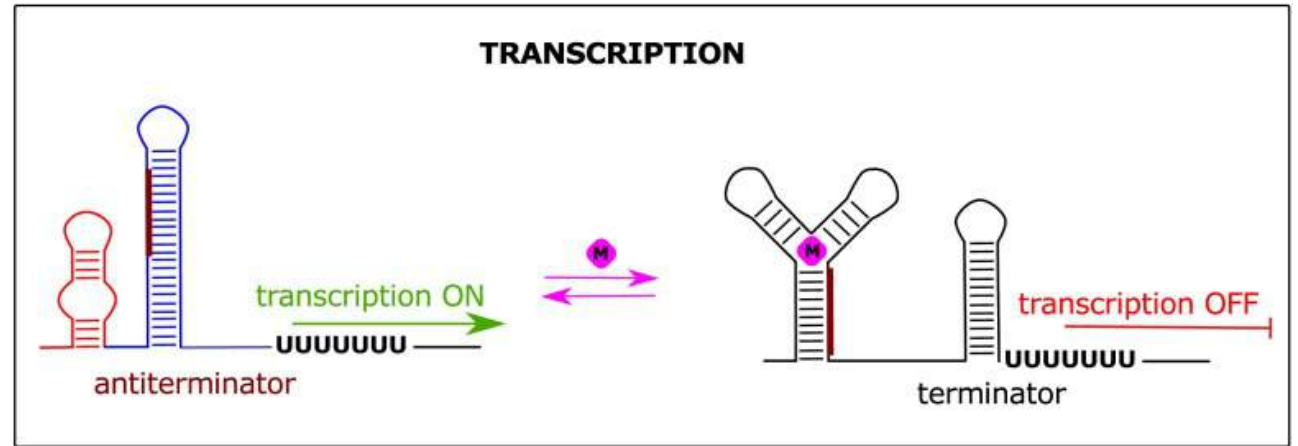
- Τα tRNAs είναι η πιο άφθονη μορφή των μικρών RNAs, αποτελώντας το 4–10% του ολικού RNA.
- Στο παρελθόν θεωρούνταν να έχουν στατικό ρόλο στη γονιδιακή έκφραση, δρώντας ως μόρια προσαρμογείς στη μετάφραση.
- Πρόσφατα έχει βρεθεί ότι tRNAs έχουν μοναδικές λειτουργίες ευνοώντας τη σηματοδότηση και τη δυναμική ρύθμιση ποικίλων κυτταρικών λειτουργιών.
- Το ώριμο tRNA δημιουργείται από μεταγραφή του πρόδρομου tRNA (pre-tRNA) μέσω της RNA πολυμεράσης III. Οι ενδονουκλεάσες ριβονουκλεάση P (ribonuclease P, RNase P) και ριβονουκλεάση Z (ribonuclease Z) κόβουν το πρόδρομο pre-tRNA στην 5' αλληλουχία οδηγό και την 3' ουρά πολύ-ουρακίλης (poly-U tail), αντίστοιχα πριν το tRNA ενεργοποιηθεί με την προσθήκη της ουράς 3'CCA από την νουκλεοτιδική τρανσφεράση.
- Τα ώριμα tRNA μπορούν να επεξεργαστούν σε ειδικά μικρά πριόντα RNA; (1) tRNA-μισά, (2) tRNA προερχόμενα τμήματα.



Ριβοδιακόπτες

Ένας ριβοδιακόπτης (**riboswitch**) είναι μια ρυθμιστική μονάδα του mRNA που προσδένει ένα μικρό μόριο (ιόντα μαγνησίου, νουκλεοσίδια, συνένζυμα, αμινοξέα) με αποτέλεσμα την μεταβολή των πρωτεϊνών που μεταφράζονται από το mRNA.

Γενικός μηχανισμός της δράσης των ριβοδιακοπών. Η μεταβολή της δομής της δομικής περιοχής με κόκκινο που προκαλείται από την πρόσδεση ενός μεταβολίτη (ροζ διαμάντι) μπορεί να προκαλέσει μεταβολή της γονιδιακής έκφρασης (πάνω πάνελ) ή της μετάφρασης (κάτω πάνελ).



Ριβοσωμικό RNA

- Τα ριβοσωμικά μόρια RNAs (rRNAs) δημιουργούν δομικά και λειτουργικά συστατικά ριβοσωμάτων, που είναι οι βιοχημικές μονάδες που ευθύνονται για την πρωτεϊνοσύνθεση. Το rRNA αποτελεί περίπου το 80–85% του συνολικού RNA σε ένα κύτταρο.
- Το rRNA προέρχεται από μία οικογένεια ριβοσωμικών γονιδίων που βρίσκεται σε πολλαπλά αντίγραφα στο γονιδίωμα (~400 αντίγραφα στον άνθρωπο).
- Στον άνθρωπο αυτές οι οικογένειες εντοπίζονται στα p άκρα των πέντε ακροκεντρικών χρωμοσωμάτων (13, 14, 15, 21, και 22) του ανθρώπου.
- Οι γενετικοί τόποι του rRNA οργανώνονται σε συστοιχίες πολλαπλών αντιγράφων που ονομάζονται ως rDNA και παρουσιάζουν υψηλό βαθμό συντήρησης. Η «ομοιογενοποίηση» της αλληλουχίας τους είναι αποτέλεσμα εναρμονισμένης εξέλιξης (concerted evolution) μέσω ανασυνδυασμού και γονιδιακής μετατροπής (gene conversion)
- Ωστόσο το μέγεθος των γενετικών τόπων του ριβοσωμικού RNA (ο αριθμός των επαναλήψεων που περιέχουν) εμφανίζει μεγάλη μεταβλητότητα μεταξύ των διαφορετικών ατόμων)

Ριβοσωμικό RNA



- Δομή της επαναλαμβανόμενης μονάδας του ευκαρυωτικού ριβοσωμικού DNA σε ένα ακροκεντρικό χρωμόσωμα. Η κάθετη διακεκομμένη γραμμή υποδεικνύει το προς το τελομερές όριο της επαναλαμβανόμενης μονάδας, ανοδικά από το οποίο υπάρχουν δύο ψευδογονίδια (επισημαίνονται με το ελληνικό γράμμα ψ στην αρχή της ονομασίας τους) και μία ακόμα αλληλουχία η οποία αναφέρεται ως DJU. Αυτή η μονάδα επαναλαμβάνεται πολλές φορές και οι επαναλήψεις της έχουν ταυτόσημη ή σχεδόν ταυτόσημη αλληλουχία DNA. Οι αριθμοί καταχώρισης αυτής της περιοχής στην GenBank είναι U67616 (8.353 ζεύγη βάσεων στα οποία περιλαμβάνονται μια ποικιλία επαναλαμβανόμενων στοιχείων και 28S rDNA ψευδογονίδια) και U13369 (42.999 ζεύγη βάσεων στα οποία περιλαμβάνονται τα γονίδια του 18S, του 5,8S και του 28S rRNA, μεταγραφόμενα ενδιάμεσα τμήματα και διάφορα επαναλαμβανόμενα στοιχεία). IGS (Intergenic Spacer): μεσογονιδιακό ενδιάμεσο τμήμα, αναφέρεται και ως NTS (Nontranscribed Spacer): μη μεταγραφόμενο ενδιάμεσο τμήμα. ETS (External Transcribed Sequence): εξωτερική μεταγραφόμενη αλληλουχία. t-p53: τερματικές (terminal) θέσεις πρόσδεσης της p53. i-p53: εσωτερικές (internal) θέσεις πρόσδεσης της p53.

Κυριότητες μορφές rRNA στα βακτήρια και στους ευκαρυώτες

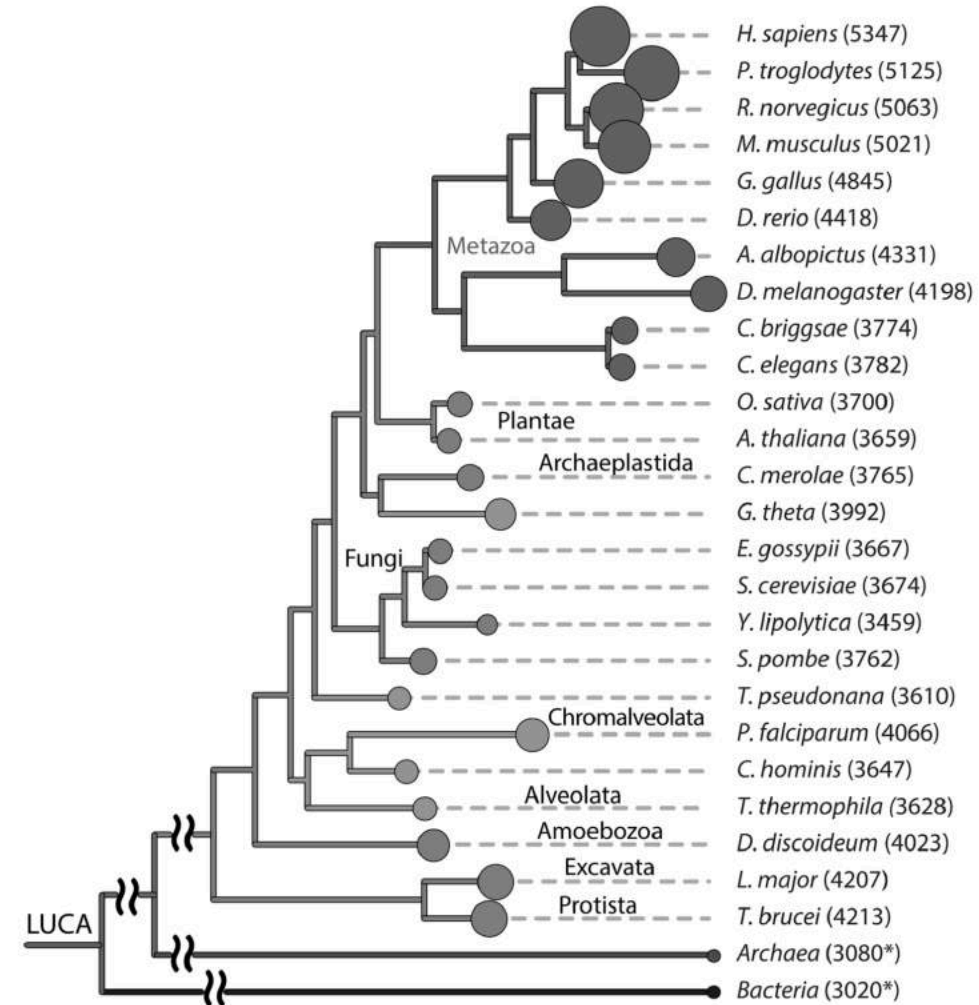
Πίνακας 10.3 Βασικοί τύποι rRNA στα βακτήρια και στους ευκαρυώτες. S: συντελεστής καθίζησης, MB: μοριακό βάρος. Οι αριθμοί καταχώρισης αφορούν τα rRNA της *E. coli* και του ανθρώπου. Προσαρμοσμένος από το NCBI και τη δημοσίευση Dayhoff et al. (1972, σελ. D352).

Κλάδος	Ριβόσωμα	MB	Ριβοσωμικές υπομονάδες	Τύπος μορίων rRNA	Λειτουργία	Αριθμός καταχώρισης	Αριθμός ζευγών βάσεων	Καταχώριση στην Rfam
Βακτήρια	70S	$2,6 \times 10^6$	30S (μικρή)	16S	Πρόσδεση στο mRNA	M25588.1	1.504	RF00177
			50S (μεγάλη)	23S	Σχηματισμός πεπτιδικού δεσμού	M25458.1	542	RF02541
				5S		M24300.1	120	RF00001
Ευκαρυώτες	80S	$4,3 \times 10^6$	40S (μικρή)	18S	Πρόσδεση στο mRNA	NR_003286.2	1.869	RF01960
			60S (μεγάλη)	28S	Σχηματισμός πεπτιδικού δεσμού	NR_003287.2	5.070	RF02543
				5.8S		NR_003285.2	156	RF00002
				5S	Small RNAs	NR_023363.1	121	RF00001

- Οι μικρές ριβοσωμικές υπομονάδες 16S (στα βακτήρια) και 18S (στους ευκαρυώτες) χρησιμοποιούνται συχνά στις φυλογενετικές αναλύσεις.

Ριβοσωμικό μέγεθος

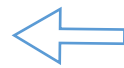
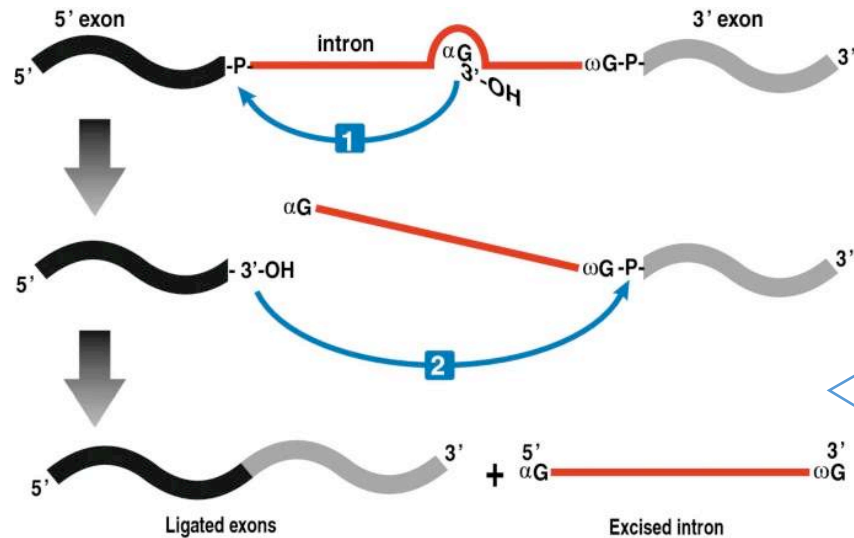
- Το ριβοσωμικό μέγεθος αλλά όχι το γονιδιωματικό μέγεθος συσχετίζεται με την πολυπλοκότητα των ειδών. Αυτό το φυλογενετικό δέντρο απεικονίζει την εκρηκτική αύξηση του ριβοσωμικού μεγέθους στους ανώτερους οργανισμούς και την απουσία συσχετισμού με το γονιδιωματικό μέγεθος.
- Οι ακτίνες των κύκλων είναι ανάλογοι με το ριβοσωμικό μέγεθος (συνολικό μήκος των μεγάλων υπομονάδων του rRNA). Τα μεγέθη των μεγάλων υπομονάδων rRNA στα αρχαία και τα βακτήρια είναι πολύ περιορισμένα και οπότε αντιπροσωπεύονται με ένα είδος το καθένα.



Το ιντρονικό RNA

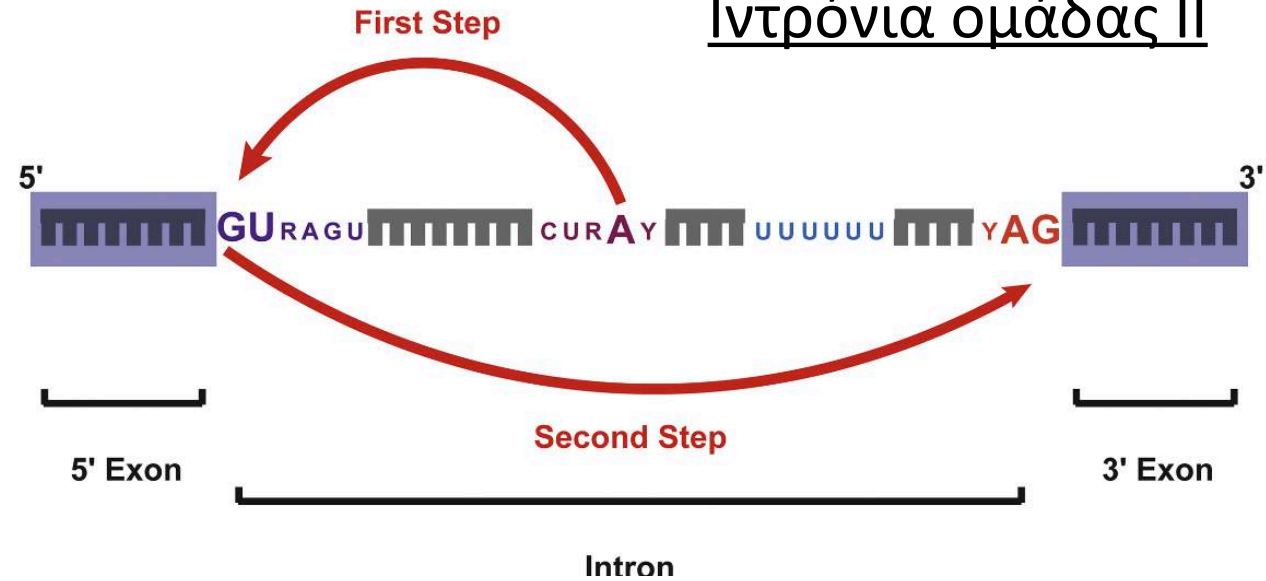
- Η δομή του ιντρονικού RNA
- Ιντρόνια ομάδας I και II (ριβοένζυμα με ή χωρίς τη επικουρία πρωτεϊνών)
- Λειτουργικά ιντρόνια
 - Κυκλικά RNAs (circular RNAs)

Ιντρόνια ομάδας I



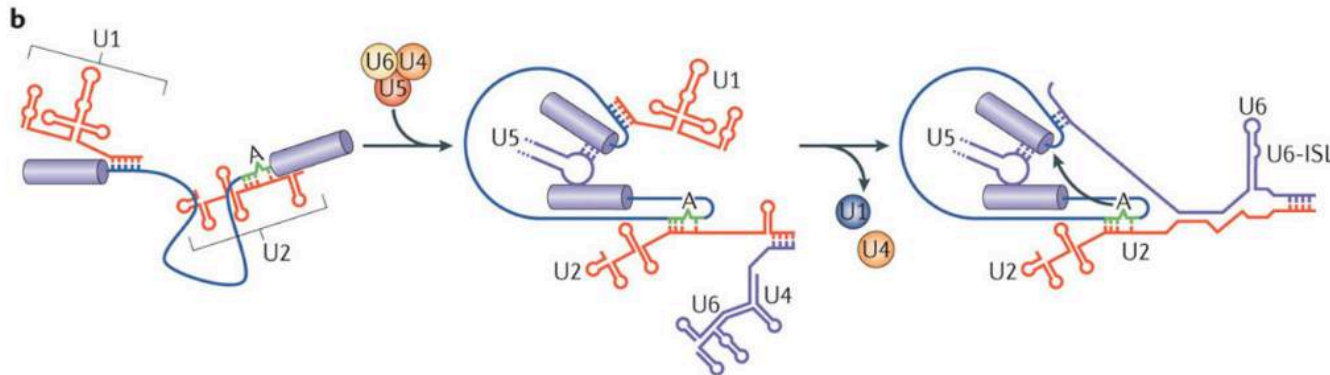
Schematic representation of group I intron splicing. The splicing pathway consists of two sequential transesterification reactions. The first reaction is initiated by the 3'-OH group of an exogenous GTP (αG) that docks into the G-binding pocket located in the P7 region and the 3'-OH group attacks the 5' splice site. In the second reaction, the 3'-OH of the released 5' exon attacks the phosphodiester bond between the intronic terminal G (ωG) and the 3' exon, resulting in the liberation of the intron and the ligation of the exons.





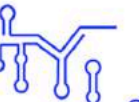
Ιντρόνια ομάδας II



Μικρό πυρηνικό RNA (Small nuclear RNA, snRNA)

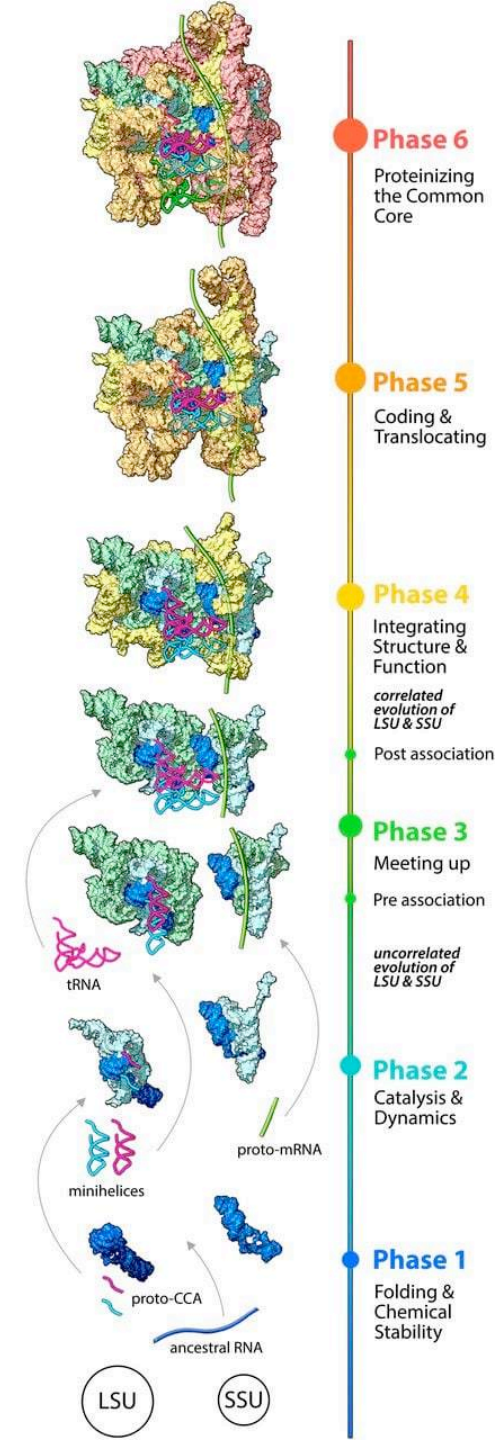
- Τα snRNAs εντοπίζονται στον πυρήνα και αποτελούνται από μια οικογένεια RNAs που μεταγράφονται από την πολυμεράση II και ευθύνονται για το (εναλλακτικό) μάτισμα και για τη διατήρηση των τελομερών
- Αποτελούνται από πολύ συντηρημένες και εμπλουτισμένες αλληλουχίες ουριδίνης (γλυκοσυλιωμένη μορφή ουρακίλης) με πέντε snRNAs να συνιστούν τον πυρήνα τους U1, U2, U4, U5 and U6. Αυτά τα snRNAs προσδέονται με πρωτεΐνες για να σχηματίζουν μικρά πυρηνικά ριβονουκλεοπρωτεϊνικά σύμπλοκα (small nuclear ribonucleoprotein, snRNPs) τα οποία αποτελούν κεντρικά μέρη του σωματίου μάτισματος
- 100 αντίγραφα U1, 1,200 αντίγραφα U6



snRNP	RNA secondary structure (length ^{**})	Sm proteins	Other core proteins associated with snRNA	Associated proteins
U1 snRNP	 (568 nt)	Sm proteins (B, D3, G, E, F, D2 and D1)	Snp1 (U1-70K) Mud1 (U1A) Yhc (U1C)	Prp39, Prp40, Prp42 Snu71 Nam8 Snu56 Urn1
U2 snRNP	 (1175 nt)	Sm proteins (B, D3, G, E, F, D2 and D1)	Lea1 (U2A'), Msl1 (U2B'), Prp9 (SF3a60) Prp11 (SF3a66) Prp21 (SF3a120) Rds3 (SF3b14b) Snu17 (SF3b14a/p14) Hsh155 (SF3b155) Cus1 (SF3b145) Rse1 (SF3b130) Hsh49 (SF3b49) Ysf3 (SF3b10)	U2AF35 Mud2 (U2AF65) Msl5 (SF1/BBP)
U4/U6 snRNP	 (160/112 nt)	Sm proteins for U4 snRNA (B, D3, G, E, F, D2 and D1) LSm proteins for U6 snRNA (Lsm2-8)	Prp3 Prp31 Prp4 Snu13/15.5K	
U5 snRNP	 (179 or 214 nt, for short or long forms)	Sm proteins (B, D3, G, E, F, D2 and D1)	Prp8 Prp6 Prp28 Brr2 Snu114 U5-40K Dib1	Snu23 Prp38 Prp2 Spp2 Yju2 Cbc2 (52K)
U4/U5/U6 snRNP		2 set of Sm proteins (B, D3, G, E, F, D2 and D1) for U4 and U5 1 set of LSm proteins (Lsm2-8) for U6	Prp3 Prp31 Prp4 Snu13/15.5K Prp8 Prp6 Prp28 Brr2 Snu114 U5-40K SnRNP27/27K Dib1	Snu23 Prp38 Prp2 Spp2 Yju2 Snu66 Sad1

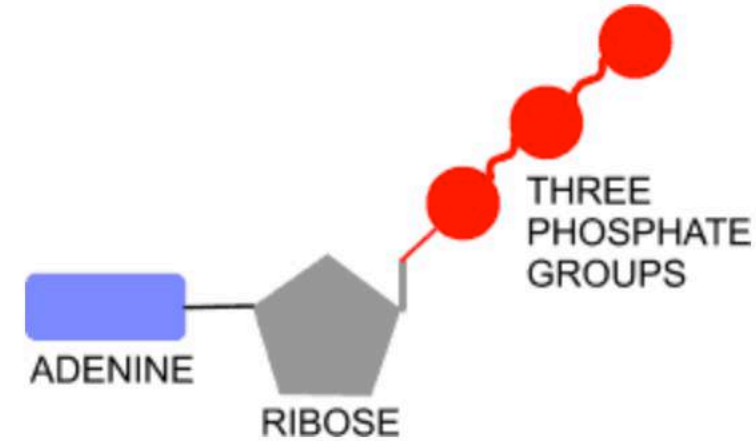
Η Εξέλιξη του Γενετικού Κώδικα

- Οι απαρχές του Γενετικού κώδικα στο πρωτο-κύτταρο απαρτίζονταν κατά πάσα περίπτωση από ένα προδρομικό ριβοσωμικό RNA, ένα tRNA άγκυρα, πυρήνα πρόδρομου RNA που μπορούσε να προσδέσει γρήγορα πεπτίδια ή λιπίδια ώστε να καθαρίζει/μαζεύει πληροφορία και ενέργεια και να αυτο-οργανώνει αυθόρμητα την ομοιόσταση και την όσο το δυνατόν, αξιόπιστη αντιγραφή του.
- Το RNA είναι το πρώτο «ζωντανό» μόριο χωρίς να περιέχει σημαντικό κωδικοποιημένο μήνυμα (π.χ. ριβοδιακόπτες, ριβόσωμα).
- Η μετάφραση δεν ήταν αξιόπιστη και η γενετική πληροφορία αν και μη αξιόπιστη από γενιά σε γενιά κατάφερε να διατηρεί την έκφραση πρωτεϊνών με ένα στατιστικό τρόπο κι όχι απόλυτο.
- Το DNA επέτρεψε την αξιόπιστη αποθήκευση κι αντιγραφή της γενετικής πληροφορίας ενώ τα χρωμοσώματα αργότερα διαμερισματοποίησαν το DNA στο χώρο και λειτούργησαν ως διαμεσολαβητές εξειδίκευσης (πυρήνας/κυτταρόπλασμα), αναπαραγωγής και διαίωξης της καταλληλότερης γενετικής πληροφορίας.
- Η εξέλιξη του ριβοσώματος/ριβοσωμικών πρωτεϊνών και η εξέλιξη των συνθετασών, ήταν δύο αλληλοεξαρτώμενες διαδικασίες και τα δύο ευνοήθηκαν αμοιβαία, γιατί η πρώτη περιόρισε τα μεταφραστικά λάθη και η δεύτερη την μείωση της ασάφειας του γενετικού κώδικα.
- Στην πραγματικότητα, σχεδόν τα πάντα έχουν τροποποιηθεί στη συσκευή της πρωτεϊνοσύνθεσης στην πορεία της εξέλιξης και η μοναδική εξαιρετική εξαίρεση αποτελούν οι κανόνες του γενετικού κώδικα. Είναι οι μοναδικές οντότητες που έχουν διατηρηθεί για δισεκατομμύρια χρόνια



ΑΤΡ- Η μπαταρία χημικής ενέργειας

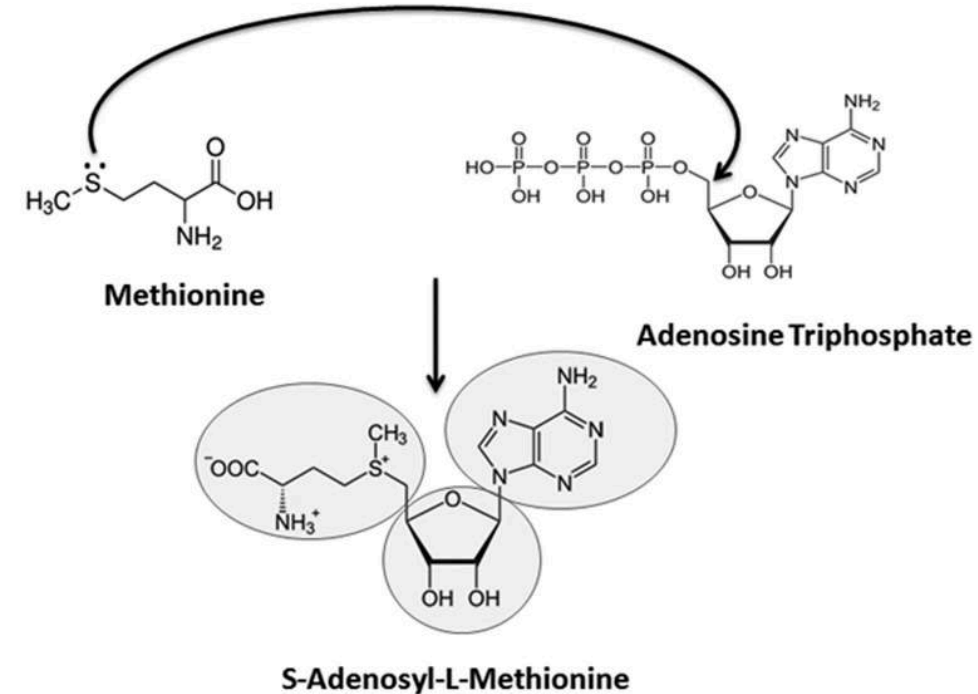
- Τριφωσφορική αδενοσίνη, (adenosine triphosphate), ονομάζεται το μόριο που αποτελείται από αδενίνη και το σάκχαρο ριβόζη, η ένωση των οποίων δημιουργεί την αδενοσίνη, στο οποίο και έχουν προσκολληθεί τρεις φωσφορικές ομάδες (φωσφορυλομάδες), που ενώνονται με δεσμούς υψηλής ενέργειας
- Το ΑΤΡ αποτελεί ένα τροποποιημένο RNA μονομερές (νουκλεοτίδιο), το οποίο σε αντίθεση με τη μονοφωσφορική αδενοσίνη του RNA, α) δεν πολυμερίζεται με άλλα νουκλεοτίδια, β) δε δημιουργεί πολυμερή και γ) περιέχει τρεις φωσφορυλομάδες.
- Το μόνο και παγκόσμιο νόμισμα ενέργειας όλων των μορφών ζωής έχει βάση το RNA.



Η αζωτούχα βάση της Αδενίνης μπορεί να δημιουργηθεί από μια ηλεκτρική εκκένωση σε ένα υδατικό διάλυμα με αμμωνία (NH₃) και υδροκυάνιο (HCN). Η ριβόζη μπορεί να δημιουργηθεί από φορμαλδεύδη μέσω της αντίδρασης φορμόζης και βασικών συστατικών, άργιλο, θερμότητα και ακτινοβολία

Πιθανός τρόπος δημιουργίας του συνενζύμου S-adenosyl-L-methionine (SAM)

- Η νουκλεοφιλική προσθήκη μεθιονίνης στο ATP, δημιουργεί το παράγοντα SAM, που αποτελεί το δεύτερο πιο διαδεδομένο μόριο (μετά το ATP) όλων των κυττάρων και συμμετέχει στις αντιδράσεις μεθυλίωσης.
- Οι μεθυλιωμένες μορφές των μικρών RNA, tRNA, είναι καθοριστικές για τον συντονισμό του γενετικού κώδικα καθώς καθορίζουν την ποιότητα της αλληλοεπίδρασης κωδικονίου-αντικωδικονίου και την ταχύτητα μετάφρασής τους στο ριβόσωμα. Δηλαδή στον εκφυλιστικό γενετικό κώδικα, ορισμένα κωδικόνια μεταφράζονται κατά προτεραιότητα και αυτός ο μηχανισμός είναι σημαντικός για τη ρύθμιση της γονιδιακής έκφρασης- καθορίζοντας το πρωτέομα του κυττάρου.
- Η μεθυλίωση του RNA είναι πολύ σημαντική στην προστασία του από υδρόλυση



Πίνακας 10.1 Οι 13 μεγαλύτερες οικογένειες RNA της Rfam. Μέλη: δίνεται (στρογγυλοποιημένος στην κοντινότερη χιλιάδα) ο αριθμός μελών της οικογένειας στο σύνολο των αλληλουχιών της βάσης δεδομένων (όχι ο αριθμός αντιπροσωπευτικών αλληλουχιών – για όλες τις ανεξάρτητα προσδιορισμένες αλληλουχίες που θεωρείται πως προέρχονται από το ίδιο γονίδιο υπάρχει μία αντιπροσωπευτική αλληλουχία). Id: ο μέσος όρος του ποσοστού ταύτισης των κατά ζεύγη στοιχίσεων.

Όνομασία	Αριθμός καταχώρισης	Μέλη	Μήκος (μέσος όρος αριθμού ζευγών βάσεων)	Id	Τύπος	Περιγραφή
5_8S_rRNA	RF00002	376.000	152	69	Γονίδιο, rRNA	5,8S ριβοσωμικό RNA
tRNA	RF00005	298.000	73	46	Γονίδιο, tRNA	tRNA
5S_rRNA	RF00001	229.000	116	60	Γονίδιο, rRNA	5S ριβοσωμικό RNA
UnaL2	RF00436	101.000	54	78	<i>cis</i> -ρυθμιστικό	UnaL2 LINE 3' στοιχείο
HIV_POL-1_SL	RF01418	83.000	113	77	<i>cis</i> -ρυθμιστικό	HIV pol-1 στέλεχος-βρόχος
U6	RF00026	72.000	105	77	Γονίδιο, snRNA, μάτισμα	U6 RNA του σωματίου ματίσματος
mtDNA ssA	RF01853	62.000	104	67	Γονίδιο, αντισημαίνον	Περιοχή ελέγχου με δευτεροταγή δομή A του μιτοχondριακού DNA
Intron_gpI	RF00028	60.000	365	36	Ιντρόνιο	Καταλυτικά ιντρόνια ομάδας I
Intron_gpII	RF00029	51.000	87	54	Ιντρόνιο	Καταλυτικά ιντρόνια ομάδας II
Hammerhead_1	RF00163	49.000	59	70	Γονίδιο, ριβοένζυμο	Ριβοένζυμο Hammerhead (τύπος I)
RRE	RF00036	44.000	337	97	<i>cis</i> -ρυθμιστικό	Στοιχείο απόκρισης HIV Rev
HIV_GSL3	RF00376	39.000	84	82	<i>cis</i> -ρυθμιστικό	HIV gag στέλεχος-βρόχος 3 (GSL3)
SNORA7	RF00409	26.000	140	79	Γονίδιο, snRNA, snoRNA, πλαίσιο HACA	Μικρό πυρηνισκικό RNA SNORA7

Τα μη-κωδικοποιά RNA της βάσης δεδομένων Rfam

- (α) Τύποι μη κωδικών RNA. Ο αριθμός των οικογενειών είναι ανάλογος της έντασης του χρώματος. Ο αριθμός των υπομνηματισμένων περιοχών είναι ανάλογος του μεγέθους των ορθογωνίων.

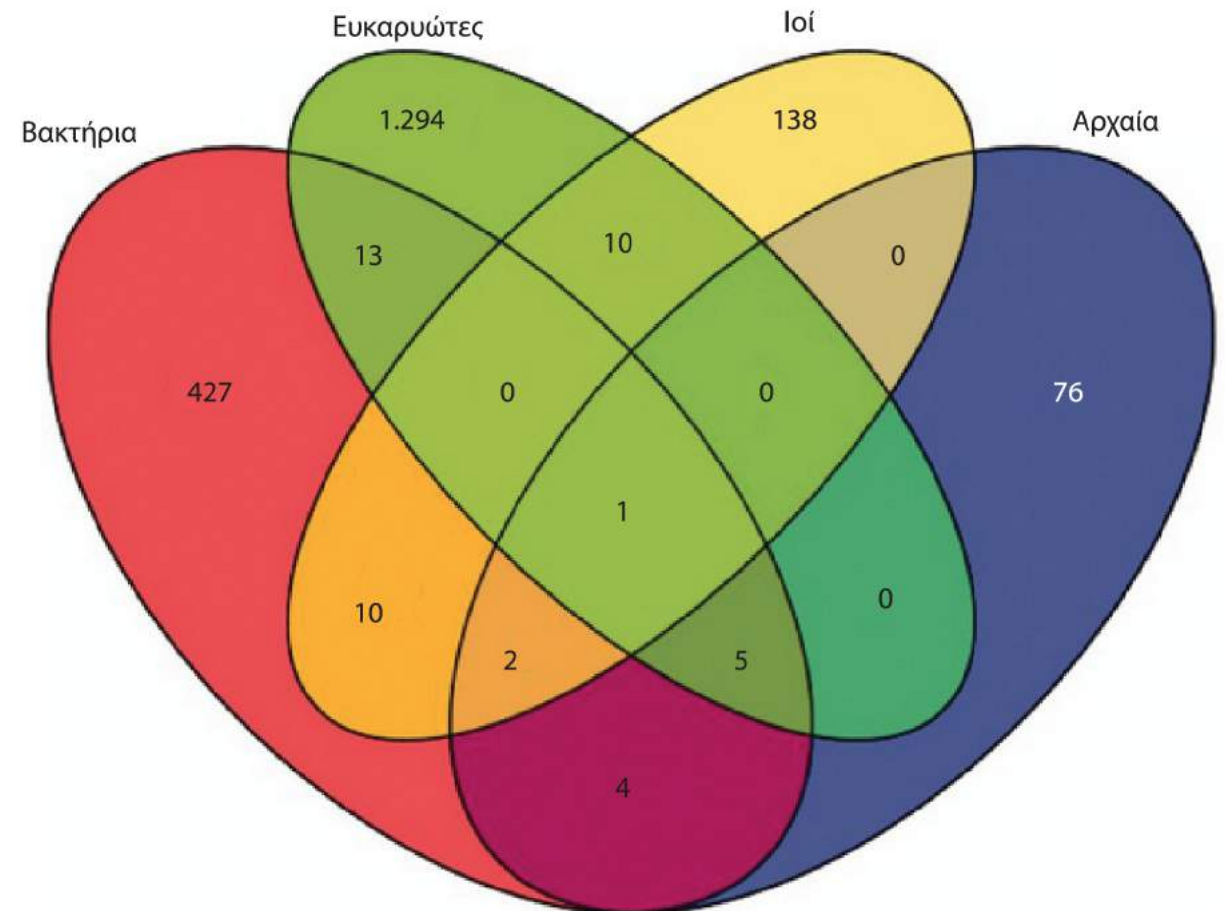
(α) Οικογένειες μη κωδικών RNA της Rfam



Ταξινόμηση των οικογενειών της Rfam στους τρεις κλάδους του δέντρου της ζωής (βακτήρια, αρχαία και ευκαρυώτες) και στους ιούς.

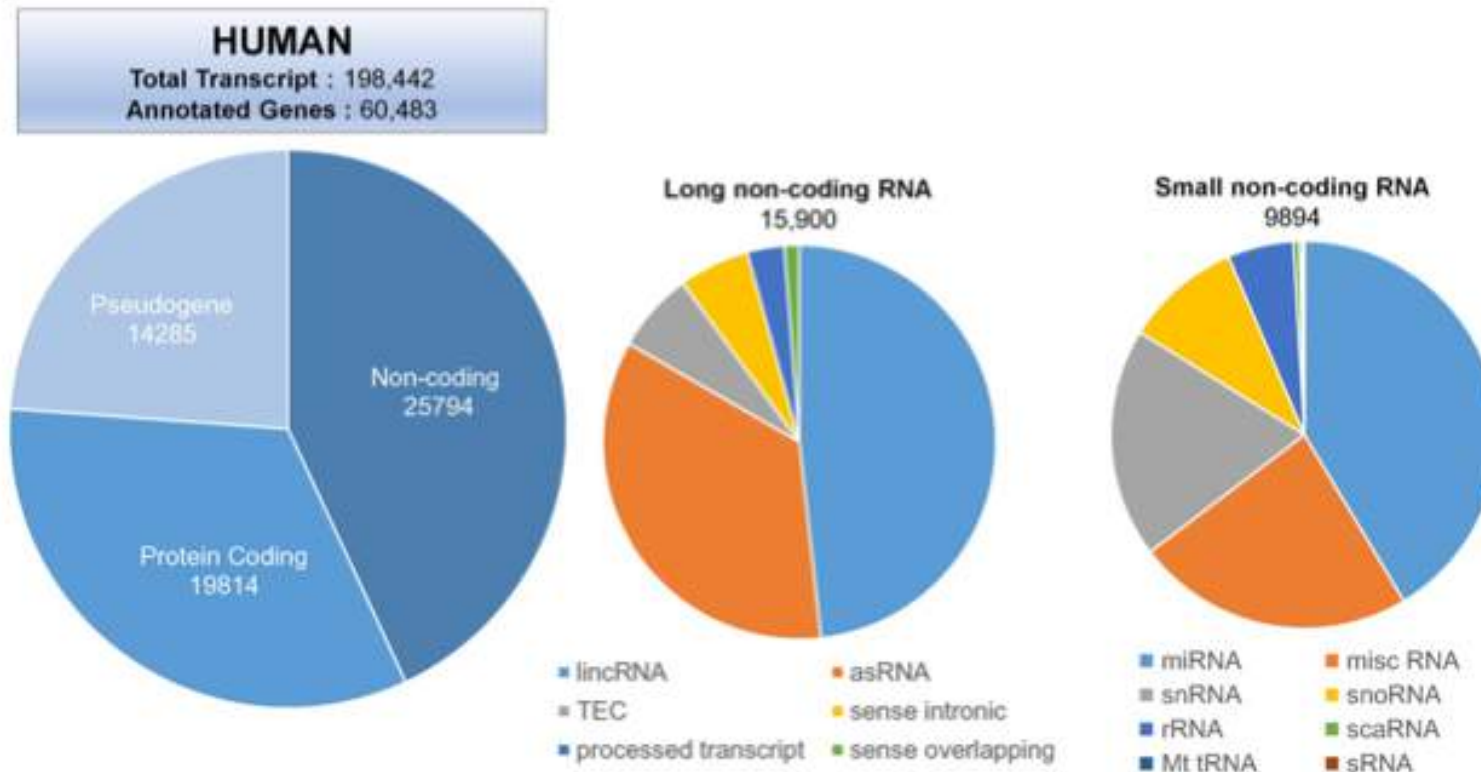
- Η ταξινόμηση βασίζεται σε αλληλουχίες αναφοράς για κάθε οικογένεια οι οποίες ονομάζονται σπερματικές αλληλουχίες (seed sequences) και όχι στο σύνολο των καταχωρίσεων της Rfam. [ΣΜΑΠ, Στοιχείο Μετατόπισης Αναγνωστικού Πλαισίου (frameshift element).]

(β) Ταξινομικές ομάδες στην Rfam

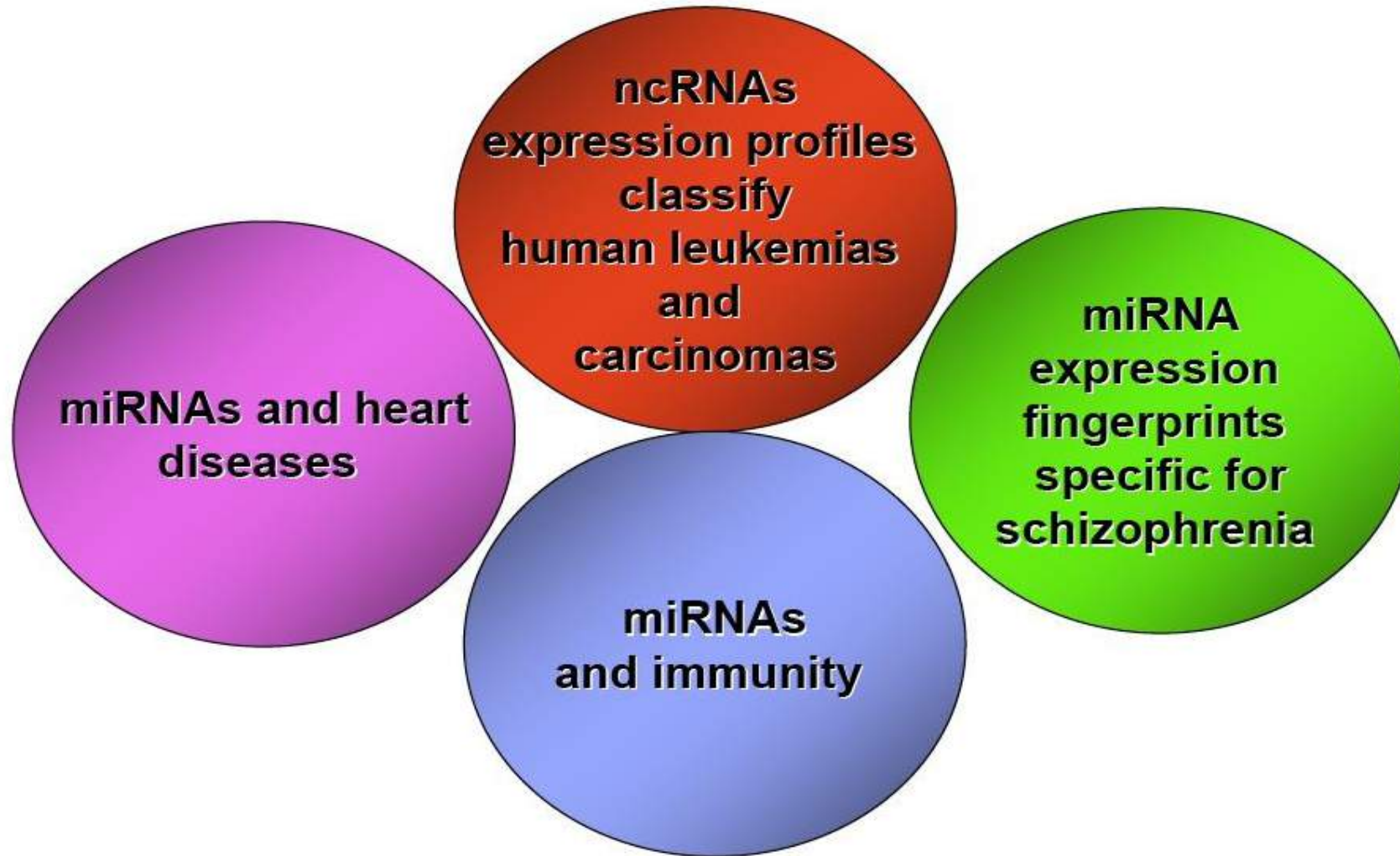


Γονίδια και Λειτουργία

- Το ανθρώπινο γονιδίωμα περιέχει ~20,000 γονίδια που κωδικοποιούν πρωτεΐνες (protein coding genes) και γύρω στα ~15,000-22,000 γονίδια που δεν κωδικοποιούν πρωτεΐνες (μακρά μη κωδικοποιά γονίδια, long non-coding RNAs; lncRNAs)
- Περίπου για το 40 % των κωδικοποιών γονιδίων δεν έχει αποδοθεί ακόμη κάποια λειτουργία!
- Τα μακρά μη κωδικοποιά γονίδια είναι φυσικά ελάχιστα μελετημένα – Περίπου 100 από αυτά έχουν μελετηθεί πειραματικά προς το παρόν



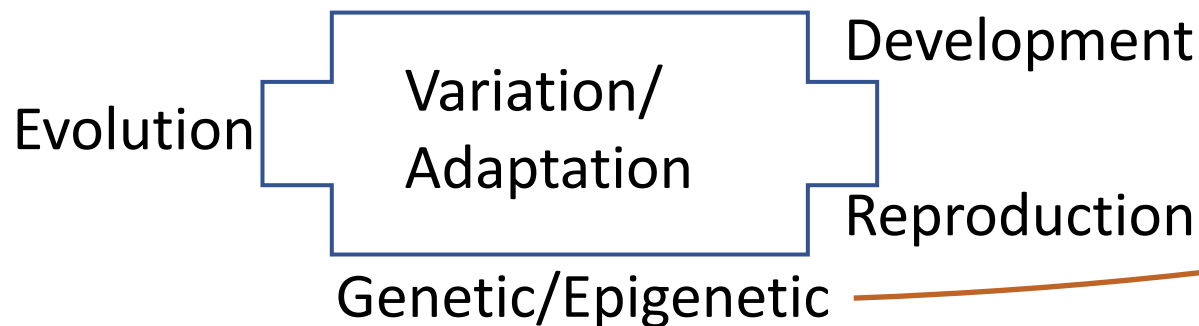
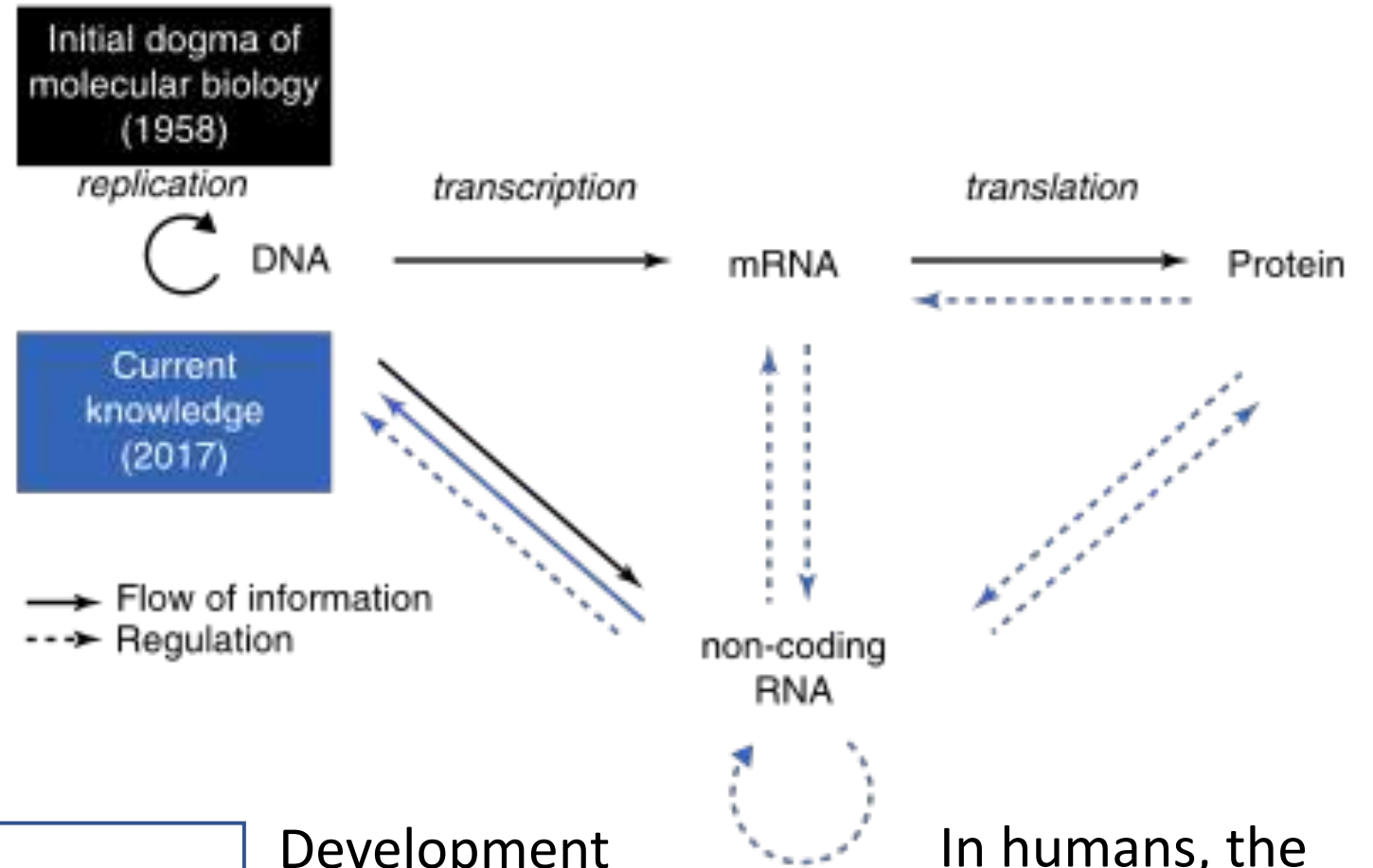
***ALTERATIONS OF NONCODING RNAS ARE FOUND IN
EVERY TYPE OF HUMAN DISEASE***



(Calin et al, PNAS 2002; Lu et al, Nature, 2005; Landgraf et al, Cell 2007; Perkins et al Genome Biol 2007; Hansen et al PLoS ONE, 2007; Beveridge et al, Hum Molec Genet 2008, Baltimore D, Nat Immunol 2008; van Rooij, Trends Genet, 2008)

Το νέο δόγμα ροής της γενετικής πληροφορίας

- Αλλάζει τον ορισμό του γονιδίου και της γονιδιακής έκφρασης
- Αλλάζει τη γενετική βάση των ασθενειών
- Το RNA δεν είναι απλώς ένα ενδιάμεσο μόριο



In humans, the majority of which lies in ncDNA/ncRNA regions

Το ρυθμιστικό μη-κωδικοποιό RNA

- Οι κύριες κατηγορίες μη-κωδικοποιών RNA είναι το tRNA και το rRNA τα οποία μαζί αντιπροσωπεύουν το 95% του συνολικού RNA σε ένα ευκαρυωτικό κύτταρο
- Άλλα μη-κωδικοποιά RNAs είναι:
 - Το μικρό πυρηνικό RNA (snRNA, small nuclear RNA)
 - Το μικρό πυρηνισκικό RNA (snoRNA, small nucleolar RNA)
 - Το μικρόRNA (miRNA, microRNA)
 - Το μικρό παρεμβαλλόμενο RNA (siRNA, small interfering RNA), και το
 - Το μακρύ μη-κωδικοποιό RNA (lncRNA, long non-coding RNA)

Εν συντομία τα ρυθμιστικά μη-κωδικοποιά RNA

- Μικρά RNAs με μεγάλες λειτουργίες.
- Δεκάδες έως και εκατοντάδες lncRNAs με μια αφθονία από πολυδιάστατες λειτουργίες.
- Μεταγραφομική, Επιγενετική, Επιμεταγραφομική, Γενετική.
- Επεξεργασία του RNA (RNA editing) και πλαστικότητα της λειτουργίας του εγκεφάλου.
- «Λάθη» στην επεξεργασία του RNA (RNA processing) τόσο στα ίδια τα RNA όσο και στις πρωτεΐνες που προσδένουν RNA αλλά και στους αντίστοιχους γενετικούς τους τόπους) κυριαρχούν στις νευρο-εκφυλιστικές, νευρο-ψυχιατρικές, νευρο-ανοσολογικές και αυτο-άνοσες ασθένειες.

Οικογένειες μη κωδικών RNA του χρωμοσώματος 21 στη βάση δεδομένων Rfam. Εμφανίζεται μόνο ένα υποσύνολο των καταχωρίσεων.

- Όταν ψάχνεται στη βάση δεδομένων Rfam με μια αλληλουχία αναζήτησης, τα αποτελέσματα περιλαμβάνουν ένα bit σκορ στη γνωστή μορφή λογάριθμου του λόγου πιθανοτήτων
- Δηλαδή ένα θετικό bit σκορ υποδεικνύει σημαντική ταύτιση της αλληλουχίας αναζήτησης με μια αλληλουχία της Rfam, καθώς η πιθανότητα με βάση το μοντέλο συνδιακύμανσης είναι αυξημένη συγκριτικά με αυτή της μηδενικής υπόθεσης, που υποδηλώνει τυχαιότητα του αποτελέσματος.

Family	Start	End	Bits score
<u>trNA</u>	9,734,391	9,734,325	31.22
<u>RSV RNA</u>	9,990,192	9,989,909	36.74
<u>RSV RNA</u>	10,142,311	10,142,595	36.83
<u>Metazoa SRP</u>	10,380,661	10,380,378	122.56
<u>SNORA70</u>	10,385,953	10,386,047	42.25
<u>trNA</u>	10,492,972	10,492,907	26.05
<u>trNA</u>	10,493,037	10,492,973	37.46
<u>mir-548</u>	11,052,015	11,051,932	82.85
<u>U6</u>	14,419,904	14,420,010	66.41
<u>U6</u>	14,993,898	14,994,004	76.41
<u>U6</u>	15,340,916	15,340,810	63.69
<u>5S rRNA</u>	15,443,192	15,443,307	42.69
<u>pRNA</u>	15,448,359	15,448,271	68.52
<u>U6</u>	16,986,602	16,986,708	75.19
<u>U6</u>	17,407,829	17,407,733	41.70

<u>SNORD74</u>	17,657,089	17,657,017	59.88
<u>mir-10</u>	17,911,414	17,911,485	69.08
<u>let-7</u>	17,912,152	17,912,227	62.65
<u>lin-4</u>	17,962,567	17,962,636	76.22
<u>U1</u>	18,091,317	18,091,476	91.09
<u>U6</u>	18,803,865	18,803,965	62.92
<u>trNA</u>	18,827,177	18,827,107	63.87
<u>Metazoa SRP</u>	18,878,771	18,879,046	64.51
<u>Y RNA</u>	18,899,565	18,899,458	41.00
<u>Y RNA</u>	18,949,116	18,949,224	40.22
<u>RSV RNA</u>	19,938,102	19,937,818	72.79
<u>U1</u>	20,717,465	20,717,629	93.53
<u>U6</u>	21,728,164	21,728,060	49.35
<u>7SK</u>	21,728,965	21,729,208	75.15
<u>mir-492</u>	21,798,181	21,798,066	40.04
<u>U4</u>	23,577,511	23,577,651	73.90
<u>U2</u>	24,654,231	24,654,058	62.66