

# Seguint les empremtes de la traducció genètica

El paper de les modificacions dels RNA de transferència en l'evolució dels genomes

Premi SCB al Jove Investigador

Eva Maria Novoa

La traducció genètica és un procés central que ocorre en els tres dominis de la vida, en el qual el codi de quatre lletres de l'RNA missatger (mRNA) és descodificat a l'alfabet proteic de vint-i-dues lletres per produir un polipèptid específic, seguint les normes del codi genètic. Per traduir aquest codi, els aminoàcids són prèviament units als seus corresponents RNA de transferència (tRNA) a través de la reacció d'aminoacilació, catalitzada per la família d'aminoacil-tRNA-sintetases (aaRS). Els tRNA aminoacilats són llavors portats al ribosoma pels factors d'elongació, on l'anticodó del tRNA és «encaixat» amb el codó de l'mRNA, i es produeix la transferència de l'aminoàcid carregat en el tRNA a la cadena polipeptídica que s'està sintetitzant, i dona lloc a la traducció del missatge.

El codi genètic estàndard, amb algunes excepcions, és el mateix en totes les espècies, i està constituït per seixanta-quatre triplets diferents (codons), dels quals seixanta-un codifiquen aminoàcids. Per descodificar tots els codons, en teoria caldrien seixanta-quatre tRNA diferents, però en realitat es pot fer la feina amb menys tRNA. El mecanisme mitjançant el qual un organisme pot llegir tots els codons sense tenir seixanta-quatre gens de tRNA va ser hipotetitzat per primer cop per Francis Crick en la seva *hipòtesi bambolejant* (*wobble hypothesis*).

Com que hi ha més codons (seixanta-quatre) que aminoàcids (vint), la majoria d'aminoàcids estan codificats per diversos codons, fenomen conegut com a *degeneració* del codi genètic. Aquests codons redundants, malgrat que són «sinònims» i codifiquen un mateix aminoàcid, no són utilitzats amb la mateixa freqüència a través dels genomes. Diversos estudis han demostrat que l'ús de certs codons determina la velocitat de traducció, i s'estableixen, en conseqüència, categories de codons «favorables» i «desfavorables». Per tant, els codons sinònims no són equivalents. De fet, quan volem incrementar l'expressió heteròloga de proteïnes modifiquem la freqüència de codons favorables i

desfavorables —sense modificar la seqüència aminoacídica— de tal manera que l'ús de codons s'assembla al de l'hoste. Però, la pregunta llavors és: per què els codons «favorables» no són els mateixos en totes les espècies?

Per intentar respondre aquesta pregunta, diversos grups han fet estudis comparatius de l'ús de codons en diferents espècies, però la presència de múltiples factors que governen l'ús dels codons a través dels genomes, com ara el percentatge de GC, dificulten aquest tipus d'anàlisi. En canvi, l'aproximació complementària, comprendre la distribució de còpies de tRNA, pràcticament no s'ha explorat. Tenint en compte que les freqüències de tRNA i codons han de coevolucionar, vam hipotetitzar que comprenent les forces que governen l'evolució dels gens de tRNA podríem comprendre les forces que han modulat l'ús diferencial de codons.

Malgrat el paper central dels tRNA en la traducció de proteïnes, la connexió entre la dinàmica de la població de gens de tRNA i l'evolució dels genomes encara no es comprèn del tot. No comprenem per què existeixen variacions entre els *pools* de tRNA entre les diferents espècies, ni els principis que determinen les abundàncies de tRNA. Vam, per tant, decidir analitzar centenars de genomes de tots els dominis de la vida, no des del punt de vista del seu ús de codons, sinó des del punt de vista del seu contingut en gens de tRNA. Els gens de tRNA solen estar presents en múltiples còpies, amb més còpies de tRNA corresponents a aquells codons més altament utilitzats en el genoma. En molts casos, hi ha diversos isoacceptors de tRNA per a un mateix aminoàcid, i cadascun d'aquests isoacceptors reconeix grups de codons diferents o solapats per al mateix aminoàcid.

A través de l'anàlisi comparativa del contingut de gens de tRNA en les diferents espècies observem que el contingut de gens de tRNA ha evolucionat de manera diferent en cada domini de la vida. Però, per què certs isoacceptors

tRNA han incrementat selectivament el seu nombre de còpies, i per què aquests increments no han estat els mateixos en totes les espècies?

Per poder respondre aquesta pregunta hem d'observar les diferències que hi ha entre els diferents isoacceptors de tRNA. Els transcrits de tRNA sofreixen extensives modificacions posttranscripcionals per donar lloc a un tRNA completament funcional i madur que pugui ser utilitzat en la traducció genètica. Hi ha més de cent modificacions posttranscripcionals identificades en els tRNA, moltes essencials per a la supervivència de les cèl·lules. Malgrat que són essencials, aquestes modificacions s'han considerat «decoracions» dels tRNA, que en varien lleugerament l'estabilitat i la capacitat d'aparellament. Però, i si algunes d'aquestes modificacions no són meres decoracions, i donen avantatges traduccionals envers els altres isoacceptors?

Entre la llista de modificacions de tRNA, n'hi ha dues que mereixen una atenció especial: la *inosina* (*I*) en posició 34, catalitzada per les adenosina-desaminases (ADAT), i els derivats de 5'-*hidroxiuridina* (xo5U), també en posició 34, catalitzats per les uridina-metiltransferases (UM). Aquestes modificacions, a diferència de la resta de modificacions del tRNA, expandeixen la capacitat de fer *wobble pairing* (aparellament bambolejant) dels tRNA modificats, i els fan així més eficients en la traducció proteica. Si tornem a mirar la representació dels gens de tRNA que han incrementat en nombre de còpies, observem que els gens que estan augmentats principalment corresponen als substrats d'ADAT i UM, i això suggereix que l'eficiència de traducció ha estat un factor essencial a l'hora de determinar l'evolució del contingut de gens de tRNA en les diferents espècies. De fet, l'abundància dels codons llegits per aquests tRNA modificats en un mRNA es correlaciona directament amb els seus nivells d'expressió. Això suggereix no només que la desviació de l'ús dels codons és una estratègia per regular els nivells d'expressió gènica, sinó també que

la modulació de l'eficiència de traducció té lloc, entre altres factors, a través de l'ús de modificacions de tRNA específiques. Els nostres resultats suggereixen que l'aparició de diferents enzims de modificació del tRNA va modelar la distribució de gens de tRNA, i en conseqüència, l'ús de codons en les diferents espècies, a través de l'aparició de les ADAT en *Eukarya*, i les UM en *Bacteria*. En aquest context les modificacions de tRNA no són meres «decoracions». Són, de fet, una força motriu que té la capacitat de modelar l'evolució dels tRNA, i en conseqüència, de l'ús dels codons en els genomes que són descodificats.

En els darrers anys el camp de les modificacions d'RNA ha sofert una revolució. Amb el

descobriments de l'enzim FTO, es va veure que les modificacions d'RNA poden ser reversibles, i per tant, constitueixen una inexplorada capa funcional de regulació dels nivells d'expressió gènica, involucrades en processos essencials com el desenvolupament o la diferenciació cel·lular. D'altra banda, cal remarcar que certes mutacions o variacions dels nivells de diferents modificacions de tRNA estan associades a diverses malalties en humans, com ara el càncer, la diabetis de tipus 2, malalties neurològiques o malalties mitocondrials, i per tant, entendre la biologia i funció d'aquestes modificacions és essencial per entendre els fenotips d'aquestes malalties. De ben segur que els propers anys ens portaran moltes sorpreses i excitants descobriments

relacionats amb el funcionament d'aquesta nova gran desconeguda capa reguladora: l'*epitranscriptòmica*.

## Agraïments

Vull agrair al meu supervisor de tesi, el doctor Lluís Ribas de Pouplana, la direcció, l'ajuda i les infinites hores de pluja d'idees que vam tenir durant els anys de la meua tesi, que van fer possible aquest projecte. També vull agrair a l'Institut de Recerca Biomèdica (IRB) i a La Caixa el finançament rebut durant el meu doctorat, i a l'European Molecular Biology Organization (EMBO) i al Human Frontier Science Program (HFSP) el finançament actual. •

## Bibliografia

- BATISTA, P. J. [et al.] (2014). «m6A RNA modification controls cell fate transition in mammalian embryonic stem cells». *Cell Stem Cell*, 15: 1-13.
- CRICK, F. (1958). «On protein synthesis». *Symp. Soc. Exp. Biol.*, 12: 138-163.
- JIA, G. [et al.] (2011). «N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO». *Nat. Chem. Biol.*, 7: 885-887.
- NOVOA, E. M. [et al.] (2012). «A role for tRNA modifications in genome structure and codon usage». *Cell*, 149: 202-213.
- NOVOA, E. M.; RIBAS DE POUPLANA, L. (2012). «Speeding with control: codon usage, tRNAs and ribosomes». *Trends Genet.*, 28 (11): 574-581.
- TORRES, A. G. [et al.] (2014). «Role of tRNA modifications in human diseases». *Trend. Mol. Med.*, 20 (6): 306-314.



**Eva Maria Novoa** (Barcelona, 1983) es va llicenciar en bioquímica (2007) i doctorar en biomedicina (2012) per la Universitat de Barcelona, i va rebre els premis extraordinaris de llicenciatura i de doctorat, respectivament. Va fer el seu doctorat com a becària de La Caixa al laboratori de traducció genètica de

l'Institut de Recerca Biomèdica (IRB Barcelona), sota la direcció del professor Lluís Ribas de Pouplana. Finançada per les beques postdoctorals EMBO i Human Frontier, va unir-se al grup de biologia computacional del professor Manolis Kellis al Massachusetts Institute of Technology (Cambridge, EUA) (2013-2016). Després va unir-se al Laboratori de Biologia i Plasticitat de l'RNA, dirigit pel professor John Mattick, al Garvan Institute of Medical Research (Sydney, Austràlia) (2016-2018). Actualment, és cap de grup del laboratori d'Epitranscriptòmica i Dinàmica de l'RNA al Centre de Regulació Genòmica (CRG), on desenvolupa nous mètodes per a l'estudi de l'epitranscriptoma utilitzant una seqüenciació de tercera generació, per tal de comprendre el paper que tenen les modificacions d'RNA en la regulació de l'embrionogènesi i herència transgeneracional.