



*Uno studio del Cnr-Imm e dell'Università di Bologna evidenzia un collegamento tra i cosiddetti codici circolari, che permettono di leggere sequenze di parole senza bisogno di segnali di inizio e fine parola, e i livelli di sintesi delle proteine. Questi risultati possono condurre all'ottimizzazione di sequenze genomiche in applicazioni di bioinformatica e biotecnologie. Il lavoro è pubblicato su Scientific Reports, del gruppo Nature*



Roma, 31 maggio 2021 - Un gruppo di studiosi dell'Istituto per la microelettronica e i microsistemi del Consiglio nazionale delle ricerche (Cnr-Imm) e dei dipartimenti di Scienze Statistiche e di Farmacia e Biotecnologie dell'Università di Bologna ha trovato un collegamento tra i cosiddetti codici circolari (codici che permettono l'identificazione del corretto frame di lettura, ovvero, consentono di leggere correttamente sequenze di parole concatenate fra loro senza bisogno di segnali di inizio e fine parola) e i livelli di sintesi delle proteine.

Questi risultati possono condurre alla definizione di nuovi strumenti universali per l'ottimizzazione di sequenze genomiche in applicazioni di bioinformatica e biotecnologie. Il lavoro è stato pubblicato su *Scientific Reports*, rivista del gruppo Nature.

“Questa ricerca, frutto della collaborazione più che decennale tra il Cnr-Imm e l’Università di Bologna, ha identificato per la prima volta proprietà universali (indipendenti dall’organismo considerato) suggerite dalla teoria dei gruppi di simmetria su cui si basano i codici circolari. Sorprendentemente, le proprietà universali identificate sono correlate con l’efficienza della sintesi delle proteine, misurata attraverso tecniche sofisticate di biologia molecolare”, dichiara Diego Gonzalez del Cnr-Imm.

La presenza di codici circolari in sequenze genomiche codificanti (la parte di mRNA che viene effettivamente letta dal ribosoma e tradotta in proteine) è stata ipotizzata per la prima volta nel 1996.

“Questi codici consentono di determinare il corretto frame di lettura della sequenza codificante e potrebbero essere il risultato evolutivo di codici ancestrali, i cosiddetti codici ‘comma-free’, proposti per la prima volta da Sir Francis Crick nel 1957”, aggiunge Simone Giannerini, professore al Dipartimento di Scienze Statistiche "Paolo Fortunati" dell’Università di Bologna e primo autore dello studio.

“Inoltre, con questo approccio, vengono messe in evidenza proprietà universali delle sequenze codificanti in connessione alle trasformazioni chimiche delle basi azotate che compongono tali sequenze: in questo modo è possibile mettere in luce alcuni processi chiave nel complesso meccanismo che regola la sintesi delle proteine”, conclude Giannerini.

Queste conoscenze, applicate a SARS-CoV-2, aiuteranno a comprendere se il genoma virale presenta una ottimizzazione dei codoni (i gruppi di tre basi azotate che identificano l’aminoacido che viene incorporato alla proteina nascente) che favorisca la sintesi proteica virale, a discapito di quella dell’ospite.