

The wonder of DNA

EN - C1 - <https://italianoconmartin.com/en/readings/the-wonder-of-dna.html>



La meraviglia del DNA - livello C1

Advanced scientific dissemination

Descrivere il DNA come “codice della vita” è efficace, ma incompleto. La sequenza nucleotidica costituisce certamente un archivio ereditabile; tuttavia, quell'archivio acquista significato soltanto dentro una cellula capace di ripiegargli, conservarlo, copiarlo, ripararlo e consultarlo selettivamente. La vera meraviglia non è dunque una stringa di lettere isolata, bensì il sistema dinamico formato da DNA, proteine, RNA e architettura nucleare.

Un genoma umano aploide completo comprende poco più di tre miliardi di coppie di basi. Moltiplicando questa quantità per l'avanzamento assiale della doppia elica, circa 0,34 nanometri per coppia di basi, si ottiene una lunghezza prossima al metro. Una cellula diploide contiene due serie cromosomiche e raggiunge quindi all'incirca due metri di DNA nucleare, distribuiti in 46 cromosomi.

Il confronto con una pallina da tennis, spesso presentato con un filo di 40 chilometri, dipende dalle dimensioni scelte. Se un nucleo di 5-10 micrometri viene ingrandito fino ai 6,7 centimetri della pallina, due metri diventano circa 13-27 chilometri. Un nucleo di 3,4 micrometri produrrebbe invece la cifra di 40 chilometri. Non si tratta quindi di una falsità assoluta, ma di un'analogia che dovrebbe dichiarare le proprie ipotesi. La conclusione solida è che il filo, alla scala della pallina, misurerebbe decine di chilometri.

La compattazione comincia quando 145-147 coppie di basi si avvolgono per quasi due giri attorno a un ottamero istonico. I nucleosomi accorciano il DNA e ne modulano l'accessibilità. La cromatina non procede però attraverso una singola gerarchia rigida di fibre sempre uguali. Le osservazioni moderne descrivono un polimero irregolare, organizzato in domini, contatti e anse che cambiano con il tipo cellulare e lo stato funzionale.

Questa architettura risolve un problema logistico fondamentale. Il nucleo deve proteggere il genoma senza sottrarlo agli enzimi. Le regioni più compatte tendono a essere meno accessibili; quelle attive

ospitano nucleosomi riposizionati, varianti istoniche e modificazioni chimiche associate alla trascrizione. Le coesine possono far crescere anse di cromatina, mentre proteine di confine limitano determinati contatti. Di conseguenza, sequenze lontane lungo il cromosoma possono incontrarsi nello spazio tridimensionale.

La metafora informatica consente altri confronti sorprendenti. Con quattro basi possibili, ogni posizione può codificare idealmente $\log_2 4 = 2$ bit. I circa 3,05 miliardi di simboli di un genoma aploide equivalgono dunque a circa 6,1 miliardi di bit, cioè 760 megabyte in notazione decimale. Le due copie nucleari di una cellula diploide portano la sequenza grezza a circa 1,5 gigabyte.

Tale equivalenza va maneggiata con cautela. Un file genomico può essere compresso perché contiene molte ripetizioni e perché gli esseri umani condividono gran parte della sequenza. Ma la "quantità di informazione biologica" non coincide con la dimensione del file: una mutazione può avere effetti diversi in contesti diversi; un enhancer può agire a distanza; lo stesso gene può generare più RNA; l'attività dipende dalla storia e dallo stato della cellula. Né è corretto identificare automaticamente il DNA non codificante con spazzatura: solo una frazione ridotta codifica proteine, mentre numerose regioni non codificanti hanno funzioni regolatrici, strutturali o trascrizionali. Altre possono essere neutrali o residui evolutivi.

La capacità gravimetrica ideale del DNA è ancora più vertiginosa. Assumendo circa 660 grammi per mole di coppie di basi e due bit per coppia, un grammo può rappresentare approssimativamente $2,3 \times 10^{20}$ byte: circa 230 exabyte, ossia 228 miliardi di gigabyte. Si tratta di un limite teorico. Un archivio pratico deve evitare sequenze problematiche, identificare i frammenti, tollerare molecole mancanti e correggere errori di sintesi o lettura.

Per questo il risultato sperimentale di DNA Fountain - 215 petabyte, cioè 215 milioni di gigabyte per grammo - è inferiore di circa tre ordini di grandezza al limite ideale, pur rimanendo eccezionale. L'archiviazione sintetica sfrutta lo stesso alfabeto della biologia, ma non replica l'intero funzionamento del genoma: converte bit in oligonucleotidi, conserva le molecole e recupera i dati mediante sequenziamento e codici di correzione.

L'eredità introduce un'altra forma di compressione creativa. Durante la meiosi, i cromosomi omologhi di origine materna e paterna si appaiano. Il crossing-over scambia segmenti corrispondenti e l'assortimento indipendente distribuisce i cromosomi nei gameti. Ovulo e spermatozoo portano ciascuno 23 cromosomi; la fecondazione ricostituisce il corredo diploide. Per la maggior parte dei loci autosomici, ogni cellula nucleata eredita quindi un allele dalla madre e uno dal padre, ma le combinazioni sono mosaici delle generazioni precedenti.

La regola richiede diverse note a margine. Il genoma mitocondriale è trasmesso quasi esclusivamente dalla madre. Nei maschi, il cromosoma X proviene dalla madre e il cromosoma Y dal padre. L'imprinting genomico fa sì che alcuni alleli vengano espressi diversamente a seconda dell'origine parentale. I globuli rossi maturi non conservano il nucleo; i gameti sono aploidi; i linfociti riorganizzano specifiche regioni per produrre una grande varietà di anticorpi; mutazioni somatiche creano nel tempo un mosaico di genomi leggermente differenti.

La duplicazione del DNA, definita semiconservativa, conserva in ogni molecola figlia un filamento parentale. L'elicasi apre la doppia elica e genera una forcina di replicazione. Poiché le DNA polimerasi sintetizzano soltanto in direzione $5' \rightarrow 3'$, il filamento guida viene copiato in modo continuo, mentre il filamento ritardato viene assemblato in frammenti di Okazaki. Primer, polimerasi, nucleasi e ligasi coordinano una catena di operazioni che deve procedere su miliardi di basi.

La correzione degli errori è stratificata. La selettività chimica e geometrica della polimerasi costituisce il primo filtro; il proofreading esonucleasico rimuove molti nucleotidi appena inseriti in modo errato; il mismatch repair riconosce successivamente distorsioni sfuggite al controllo. Vie distinte affrontano basi ossidate, addotti voluminosi, rotture a singolo o doppio filamento. Nel complesso, la frequenza finale di errore della replicazione è spesso stimata nell'ordine di una base ogni miliardo o dieci

miliardi copiate, benché il valore vari con enzima, sequenza e tipo di cellula.

Un controllo assoluto sarebbe, paradossalmente, incompatibile con la storia della vita. La stabilità permette l'eredità; la rara mutazione introduce variazione. Se avviene nelle cellule germinali, una mutazione può passare alla generazione successiva; se avviene in una cellula somatica, resta nel corpo e può essere innocua, utile a una particolare funzione immunitaria oppure contribuire a una malattia.

Per leggere un gene, la cellula deve trasformare localmente la geometria della cromatina. Fattori di trascrizione riconoscono motivi di sequenza e reclutano cofattori; complessi ATP-dipendenti fanno scorrere, espellono o ristrutturano nucleosomi; acetilazioni e altre modificazioni alterano le interazioni; enhancer e promotori possono essere avvicinati da anse. La RNA polimerasi forma infine una bolla di trascrizione di dimensioni limitate e produce un RNA complementare al filamento stampo. L'intero cromosoma non viene srotolato: si apre una finestra molecolare nel punto e nel momento opportuni.

La selettività della lettura spiega perché un neurone e un epatocita, pur condividendo quasi lo stesso genoma, possiedano morfologia e metabolismo radicalmente diversi. Ogni tipo cellulare mantiene un repertorio di regioni accessibili e reti regolatrici. La differenza fondamentale non sta soltanto nei geni posseduti, ma nei geni utilizzati, nella quantità di RNA prodotto e nella cronologia dell'attivazione.

Il DNA appare dunque come una materia piena di tensioni feconde: chilometrico alla scala del nucleo eppure ordinato; densissimo eppure interrogabile; copiato con fedeltà estrema eppure capace di evolvere; quasi identico tra le cellule eppure interpretabile in centinaia di modi. La meraviglia non consiste nell'assenza di errori o nella perfezione statica, ma in un equilibrio attivo tra conservazione, accesso, riparazione e cambiamento.

Useful words

assiale = axial ottamero istonico = histone octamer accessibilità = accessibility enhancer = enhancers capacità gravimetrica = gravimetric capacity oligonucleotide = oligonucleotide crossing-over = crossing over proofreading = proofreading addotto = adduced filamento stampo = mold filament

Understanding and reflection

- What limits do the code and library metaphors have?
- Why aren't compaction and accessibility opposite goals?
- How does non-absolute correction benefit life?

La tua risposta

Argomenta questa tesi: “La meraviglia del DNA non è la perfezione, ma l'equilibrio dinamico tra stabilità e cambiamento”.

La risposta resta salvata solo su questo dispositivo.



Italiano con Martin

italianoconmartin.com