

## PIU DI 70 ANNI DOPO: IL DNA E' OGGI ARCHIVIO DEL SAPERE

Anna Valerio



Quel 28 febbraio 1953 nell'Eagle Pub di Cambridge quando Watson e Crick annunciarono agli amici di aver scoperto il segreto della struttura del DNA, forse solo la fantascienza più spinta poteva immaginare che quella stessa molecola, cuore della vita sulla terra, a poco più di 70 anni di distanza sarebbe stata usata come il più potente mezzo per immagazzinare la conoscenza di un intero pianeta.

Invece è proprio questo il futuro che ci attende grazie alla straordinaria scoperta di un gruppo di ricercatori diretti

da Nick Goldman, che lavorano in Gran Bretagna all'European Bioinformatics Institute (EBI) di Hixton, il centro più avanzato in Europa per la ricerca in campo biologico, qualche cosa come il Cern di Ginevra ma dedicato agli studi di biologia anziché alla fisica delle particelle.

E' stato realizzato un sogno ma ciò che colpisce è il significato profondo di questa scoperta: passare dal mondo inorganico del silicio al mondo organico della molecola della vita per conservare tutto il nostro sapere.

Da che cosa è partita questa ricerca?

Non ci si pensa mai ma persino quando banalmente scriviamo al computer qualche cosa che poi vogliamo salvare aggiungiamo byte su byte alla già enorme mole di informazione digitale, stimata essere oggi intorno a 3 mila miliardi di miliardi di byte, che grava sulla nostra terra. E' una quantità destinata a crescere e l'archiviazione di tutti questi dati è la sfida del futuro. Nonostante in questo campo negli ultimi anni la tecnologia abbia fatto passi da gigante (passando dal floppy all'hard disk e poi a CD e al DVD) superando ogni volta se stessa, questa massa di dati sempre meno potrà essere conservata con i metodi tradizionali. Gli hard disk infatti sono troppo costosi e soprattutto dipendono da una fonte costante di energia, mentre i sistemi magnetici hanno il limite di degradarsi nel tempo.

Ecco allora l'idea: immagazzinare le informazioni in un frammento di DNA.

Ma perché proprio il DNA?

Perché in un grammo della doppia elica possono essere contenuti 2 petabyte ( $10^{15}$ ) di dati (oggi per questo ci servirebbero 3 milioni di CD), perché non c'è bisogno di energia elettrica (basterà conservare il "DNA-drive" in un luogo buio, fresco e asciutto) e perché la tipologia di immagazzinamento garantisce un'affidabilità per migliaia o forse centinaia di migliaia di anni (la vita sul nostro pianeta lo testimonia meglio di qualsiasi altro criterio di attendibilità) senza che di debba pagare il prezzo di cambi di tecnologia che oggi rendono difficile o impossibile la lettura di dati stoccati con sistemi obsoleti. Infatti le stesse strumentazioni che verranno usate per leggere un codice genetico andranno bene per accedere a tutte le informazioni di una civiltà, anche se stivate migliaia di anni prima.

Ah, *last but not least*, lo spazio richiesto è davvero ridottissimo: nel pugno di un bimbo il sapere della terra.

Loro, quelli dell'EBI, proprio lì vicino alla Cambridge di Watson e Crick, hanno trovato il modo. Ci lavoravano da tempo in collaborazione con altri centri di ricerca perché, se leggere le informazioni contenute nel DNA era già possibile e relativamente semplice, scrivere invece sulla doppia elica era ancora troppo complicato.

Lui, Nick Goldman, ha una grande esperienza nel campo poiché fin dagli esordi della sua carriera di ricercatore ha usato proprio metodi per l'analisi del DNA e delle sequenze di amminoacidi per studiare l'evoluzione.

Ciò che sono riusciti a fare insieme è ideare un codice con il quale archiviare i dati sul DNA.

Ma che cos'è il DNA? E' una macromolecola, cioè una struttura di grandi dimensioni, formata da una successione di unità più piccole monomeriche, legate le une alle altre come le perle in una collana. I monomeri sono chiamati nucleotidi e sono di quattro tipi (come se le perle della nostra collana fossero di 4 colori diversi). Ciò che dà il "colore" alle perle è una specifica componente, la base azotata, che può essere di 4 tipi differenti: Adenina, Guanina, Timina e Citosina (A, G, T, C). Due lunghi filamenti di "perle" fatti in questo modo si uniscono insieme a formare la doppia elica del DNA, una molecola ISAC: Informazionale (porta un'informazione), Spiralizzata (è avvolta a spirale per autoprotettersi), Antiparallela (i due filamenti si guardano testa-coda), Complementare (le coppie che si formano tra le basi "dirimpettaie" sono fisse e fedeli).

Come apprendiamo dal loro articolo apparso su Nature, proprio la stessa rivista che nel 1953 riportò le conclusioni di Watson e Crick, i ricercatori si sono mossi così: sono partiti da un'informazione codificata in formato digitale, normalmente espressa in bit, e l'hanno convertita in base 3 (trit), cioè nell'unità di base di un sistema numerico che utilizza tre cifre (generalmente 0, 1 e 2) al posto delle solite due (0 e 1). A ogni *trit* hanno poi fatto corrispondere un nucleotide dei quattro (A, T, G o C). E' stato possibile in questo modo formare lunghi filamenti di DNA nei quali, grazie alle molteplici combinazioni numeriche ottenibili con un codice a tre cifre, era quasi impossibile che due nucleotidi dello stesso tipo fossero vicini e si generassero polimeri strutturalmente identici (questo è il maggior problema che si incontra con questo tipo di tecnologie!).

Hanno poi diviso il lungo filamento in frammenti più corti di circa 100 basi, certamente più agevoli da manipolare e, come sistema di controllo a garanzia della correttezza della codifica, hanno approntato altri segmenti nei quali gli stessi dati erano stati archiviati in ordine inverso (come scrivere una parola al contrario); infine vi hanno legato sequenze *segnale* che permettessero di indicarne la posizione precisa all'interno del filamento originale e quindi di ricostruire l'intero file.

In questo "DNA pioniere" sono stati memorizzati: una clip audio mp3 con il discorso di Martin Luther King "I Have a Dream", i 154 sonetti di Shakespeare in un file .txt, un .pdf con la memoria scientifica di James Watson e Francis Crick nella quale annunciavano al mondo la struttura a doppia elica del codice della vita e infine un file di immagine che illustra l'intero processo di codifica.

Questi diversi elementi del messaggio scelto per l'esperimento sono stati trasmessi dall'EBI all'Agilent Technologies di Santa Clara (California) che ha provveduto a trasformarli in una stringa di basi nucleotidiche, consegnate poi a un laboratorio di Heidelberg, che le ha lette con un sistema laser e le ha inviate al professor Goldman che ha potuto fare il confronto con i dati iniziali. Non sono stati riscontrati errori: nel DNA erano state stoccate correttamente le informazioni che si volevano conservare.



Nel passato si erano già fatti tentativi di registrare informazioni sul DNA ma c'erano stati dei problemi in quanto si era riusciti a produrre solo brevi sequenze per di più con molti errori. Ora Goldman e collaboratori hanno saputo superare questi ostacoli non solo grazie all'invenzione del nuovo codice che impedisce le ripetizioni ma anche all'idea di rompere la sequenza in una serie di frammenti che si sovrapponevano alle estremità e che contenessero un segnale per essere poi ricollocati al posto giusto. In questo modo c'era la possibilità di verificare la bontà del sistema: solo ritrovando lo stesso errore in 4 diversi frammenti, evento statisticamente quasi impossibile, si sarebbe mancato il bersaglio con il conseguente fallimento. Certamente ora ci sono molti aspetti pratici da risolvere ma la scelta del DNA è sicuramente vincente in quanto la sua densità e la straordinaria durevolezza ne fanno una molecola ideale per l'archiviazione.

Gli autori scrivono che "la conservazione su DNA è già economicamente valida per archivi a lunga scadenza e con una bassa aspettativa di accesso frequente, come gli archivi storici e governativi", per i quali si può prevedere un orizzonte di conservazione compreso fra i 600 e i 5000 anni. Non si esclude però che in futuro il metodo possa diventare conveniente anche per grandi insiemi di dati che si vogliono conservare "solo" per 50-100 anni o per tempi inferiori.

Qualcuno ha parlato di primo passo verso il computer biologico, capace di conservare quantità di dati inimmaginabili come e più della mente umana, del resto il DNA contiene tutte le informazioni che definiscono le caratteristiche di ognuno di noi viventi. Qualcun altro ha ricordato i sogni mai sopiti di elaborare un'intelligenza artificiale e di creare macchine in grado di pensare e agire come gli esseri umani, ma con potenzialità enormemente maggiori.

Certo ne sono stati fatti di passi avanti da quel lontano 1964 quando, a tutto il mondo, era parsa fantascienza la realizzazione da parte di Joseph Weizenbaum, allora giovane ricercatore del MIT, di Eliza (nome ispirato al personaggio del Pigmaleone di G. B. Shaw) uno dei primi programmi di intelligenza artificiale in grado di sostenere una conversazione in inglese su "copioni" prestabiliti di diverso argomento.

Ci chiediamo: in un futuro che in questo momento ci appare ancora lontano qualcuno non avrà forse l'idea di costruire un robot-archivio dotandolo di un "proprio" DNA che contenga tutto il sapere della terra?

A questo punto è quasi impossibile non ripensare a "2001: Odissea nello spazio" dove la mente di Stanley Kubrick aveva creato Hal 9000, computer intelligente, sensibile all'arte, in grado di governare la Discovery e ... di prendere decisioni autonomamente.

## RIFLESSI ON LINE

Iscrizione presso il Tribunale di Padova  
n.2187 del 17/08/2009

Direttore Responsabile  
Luigi la Gloria  
luigi.lagloria@riflessionline.it

Vice Direttore  
Anna Valerio  
anna.valerio@riflessionline.it

Coordinatore Editoriale  
Gianfranco Coccia