

E ora il Dna conserva la memoria digitale

Nick Goldman ha immagazzinato un file in una molecola

PIETRO GRECO

PARE CHE LUI, NICK GOLDMAN, L'IDEA L'ABBA MATURETA INSIEME AL SUO AMICO E COLLEGA EWAN BIRNEY IN UN PUB, DAVANTI A UNA PINTA DI BIRRA. Fatto è che a inizio dello scorso mese di gennaio il biologo molecolare in forze all'European Bioinformatics Institute, di Hinxton, nei pressi di Cambridge, in Gran Bretagna annuncia sulla rivista *Nature* di aver immagazzinato un file di 739 kilobytes - contenente, tra l'altro, un sonetto di Shakespeare e il più famoso discorso di Martin Luther King, quello iniziato con la frase *I have a dream* (ho un sogno) - in un filamento di Dna opportunamente sintetizzato. E che, a richiesta, l'insolita memoria digitale ha restituito il file integro, senza un solo errore.

Nick Goldman è venuto a BergamoScienza 2013, il festival scientifico che è iniziato nella città lombarda lo scorso 4 ottobre e che si concluderà il prossimo 20 ottobre. Usare il Dna come memoria dei computer del futuro si annuncia come la (possibile) rivoluzione nelle tecnologie informatiche prossima ventura. Chiediamo allo scienziato inglese come e perché.

Dottor Goldman, lei sta tentato di sviluppare un nuovo metodo per conservare l'informazione digitale in una molecola biologica, il Dna. Su cosa si basa il suo metodo? Perché Dna sintetico e non naturale? Perché il Dna e non un altro biopolimero?

«Molte persone sanno che l'informazione digitale contenuta nei nostri computer è conservata in modalità binaria, ovvero in un codice che usa due sole unità: uno o zero. L'informazione è conservata come una lunga stringa di zero e di uno. La nostra idea è semplice: perché non usare un codice a quattro lettere. Ecco il Dna, la molecola che conserva l'informazione genetica, è per l'appunto una lunga stringa di quattro lettere che si ripetono: A, C, G e T. Queste lettere sono le iniziali delle quattro basi nucleotidiche che costituiscono il Dna. Abbiamo pensato che in linea di principio anche noi, come la natura, possiamo conservare l'informazione in una lunga stringa con le quattro lettere del Dna. Naturalmente c'è bisogno di un codice appropriato e la nostra idea è stata proprio quella di iniziare a metterlo a punto. Usiamo

Dna sintetico e non naturale perché abbiamo bisogno di conservare informazione digitale con un codice che è diverso da quello usato per trasmettere l'informazione biologica. È un po' come l'italiano e l'inglese. Le due lingue usano le medesime lettere, ma le parole e le frasi sono diverse. Noi abbiamo bisogno di un codice diverso rispetto a quello genetico. Ma sfruttiamo il medesimo fantastico sistema di base del Dna.

Potremmo utilizzare altri biopolimeri? Questo è un problema importante. Certo, potremmo utilizzare il sistema di immagazzinare informazione di altre grandi molecole biologiche. Ma la storia ha segnato la strada. Oggi noi siamo in grado di studiare, manipolare e copiare il Dna e, quindi, utilizziamo questa molecola. Ma nulla vieta che in futuro potremo utilizzare il sistema di base di altre molecole biologiche».

Il vostro metodo si basa su un processo con correzione degli errori. Come funziona?

«Quando abbiamo progettato il nostro sistema ci siamo resi conto che potevano esserci molte fonti di errore sia in fase di scrittura, nel modo in cui l'informazione veniva registrata, sia in fase di lettura, nel modo in cui veniva restituita. Sappiamo, per esempio, che nella sintesi artificiale del Dna anche il miglior sistema attuale presenta errori dell'ordine dell'1 o 2%. Inoltre gli studi di genomica degli ultimi decenni ci consentono di conoscere l'errore che commettiamo quando leggiamo il Dna. Se ci sono tante fonti di errore, abbiamo pensato che dobbiamo inventare un sistema di correzione di errori. D'altra parte sistemi di correzione degli errori esistono anche nelle tecnologie digitali attuali. I nostri computer, i nostri cellulari, le nostre televisioni li utilizzano, anche se non ce ne accorgiamo. Ecco, noi abbiamo messo a punto un sistema di correzione degli errori che, in principio, non è diverso da quello utilizzato nei nostri telefoni cellulari. In pratica, nei nostri file da registrare inseriamo informazioni aggiuntive e inoltre "scriviamo" la medesima informazione in di-



versi modi. Ciò consente di correggere gli eventuali errori. Abbiamo dimostrato che il nostro sistema funziona. L'informazione scritta nel Dna sintetico ci è stata restituita integra».

Quanta informazione può essere conservata in memorie a Dna, oggi e in futuro? Quanto a lungo? È il Dna l'archivio ideale dei «Big Data»?

«Nel nostro esperimento iniziale abbiamo immagazzinato poco meno di un megabit. Non è moltissimo per gli standard attuali. E il costo è stato altissimo. Ma la velocità con la quale i biologi hanno imparato a leggere le sequenze di Dna è aumentata di un milione di volte negli ultimi 5 o 6 anni. La sintesi del Dna non è così avanzata, ma confidiamo che aumenti a ritmi notevoli nei prossimi anni, soprattutto se le industrie investiranno abbastanza nel settore. Cioè pensiamo che nel giro di una decina di anni conservare i dati nel Dna diventerà una tecnologia competitiva per capacità e costi. I biologi sanno che ci sono organismi fossili che hanno conservato intatto il loro

Il biologo oggi a Bergamoscienze parlerà della sua ricerca «fantascientifica»: usare come database un filamento (sintetico) del codice genetico

Dna per decine e persino centinaia di migliaia di anni. È ragionevole ritenere che è possibile trovare le condizioni adatte per conservare integro il Dna sintetico per un tempo analogo. E non ci sarà bisogno di elettricità o di altri sistemi che richiedono manutenzione. Insomma, avremo un sistema assolutamente sicuro in grado di conservare una quantità limitata di informazione preziosa per un lunghissimo periodo. Quale tipo di informazione verrà conservata in un tale scrigno non lo possiamo dire. Ci piacerebbe che venisse conservata informazione che l'umanità assolutamente non può perdere, come i sonetti di Shakespeare o i discorsi di Martin Luther King».

La tecnologia del Dna è migliore di quella basata sul silicio o sulle nanotecnologie per conservare informazione digitale? Pensa che il Dna sarà l'archivio del futuro? Scrivere e leggere col Dna sarà più facile e veloce che con il silicio?

«Al momento le tecnologie al silicio sono certamente migliori per facilità, rapidità e costi. Ma il Dna ha un vantaggio di lungo periodo: la stabilità. Presto riusciremo a leggere le sequenze di Dna con notevole rapidità. Inoltre le tecnologie al silicio invecchiano presto e domani avremo difficoltà a leggere quello che abbiamo scritto ieri. Questo col Dna non è possibile. Il Dna e il suo linguaggio non possono diventare vecchi. Cioè, ripeto, nel caso di relativamente piccole quantità di informazioni di alto valore non ho dubbi: il Dna a un bel po' di vantaggi rispetto al silicio».

Lei sta studiando anche l'elaborazione delle informazioni? Pensa che in futuro avremo computer a Dna?

«Io e il mio gruppo abbiamo scelto di non interessarci di questo aspetto. Ma quello del computer a Dna è un settore di ricerca ormai ben stabilito. Noi ci interessiamo solo di una parte del problema, che è quello della conservazione dei dati. Penso tuttavia che l'ela-



borazione dell'informazione via Dna invece che via silicio sia un obiettivo molto più difficile della sola conservazione dell'informazione. Penso sinceramente che il Dna non sia il futuro del "data processing", perché il Dna è troppo difficile da controllare, troppo lento nel processare e, al momento, troppo costoso».

Le ricerche che lei porta avanti col suo gruppo richiedono competenze molto diverse: in matematica, in informatica, in biochimica. Richiede interdisciplinarietà. Come superate i muri che dividono le varie discipline?

«Questo è un punto molto importante. Personalmente ho esperienza nel campo dell'evoluzione molecolare. Ma col mio gruppo siamo stati in grado di portare avanti il nostro progetto perché abbiamo messo insieme un sacco di gente esperta in computer technology, matematica, teoria dell'informazione, microbiologia, lavori di laboratorio e tante altre competenze.

Nel mettere su questo gruppo ho sempre cercato il migliore equilibrio tra gli esperti di matematica anche molto tecnica e di biologia applicata, imparando che è molto importante fare in modo che tutte queste persone con diverso background scientifico e un diverso linguaggio tecnico trovassero un modo di parlarsi l'un l'altro, che comprendessero le esigenze di ciascuno e iniziassero ad apprendere l'uno dall'altro per trovare una soluzione inusuale e sorprendente a un problema che altre persone non avevano neppure pensato, non avendo un crogiolo di esperienze così diverse».

Lei è a Bergamo per prendere parte a un festival della scienza. Che importanza attribuisce alla comunicazione della scienza al grande pubblico?

«Devo confessare che è diventata molto più importante da quando, circa un anno fa, abbiamo pubblicato il nostro articolo sulla conservazione dell'informazione in un Dna di sintesi. Il mio settore di interesse primario, l'evoluzione molecolare, non desta grande attenzione nel grande pubblico. Io e i miei col-

laboratori eravamo felici di vivere nella famosa torre d'avorio e comunicare quasi esclusivamente con i nostri colleghi. Ma trovo semplicemente fantastico, ora, essere coinvolto in un progetto che cattura così tanto l'immaginazione della gente comune. Mi considero molto fortunato per gli inviti che da un anno mi vengono da ogni parte d'Europa per parlare al pubblico del mio lavoro. Per l'occasione che mi viene data di mostrare il valore della scienza in generale e delle mie ricerche in particolare. Spero che tutto questo risulti interessante anche per le persone che frequenteranno il Festival di Bergamo e che contribuisca ad aumentare l'interesse per l'impresa scientifica».