



**C'è un
computer in
quel faggio**

Un'idea nuova
e brillante per
immagazzinare
un'infinità di
canzoni e dati
di ogni tipo.
In una foglia.

BIT BIOLOGICI. A lato e in basso, le piantine di *Nicotiana benthamiana* usate negli esperimenti.

Il sogno di Karin Ljubič è entrare in una foresta e ascoltare le canzoni dei Rolling Stones. Oppure leggere le informazioni di Wikipedia che riguardano gli alberi... e scaricarle direttamente da loro, dalle piante. Per realizzare il suo desiderio, Karin, esperta di genetica e informatica all'Università di Maribor, in Slovenia, ha dovuto inventarsi un metodo avanzato e applicarlo al codice genetico di vegetali.

VALANGHE DI NUMERI. Tutto è partito da un'osservazione che fanno in molti: le informazioni si accumulano sui computer, riempiono i dischi fissi, strabordano sui cloud e sui cellulari. Ogni giorno produciamo 2.500 miliardi di gigabyte e il trend sembra in salita. È vero che la maggior parte di questi dati sono ridondanti, inutili o addirittura dannosi, ma per molti liberarsene è un problema. E ci sono anche i dati scientifici indispensabili: il solo Cern di Ginevra, nel 2013, ha immagazzinato circa 90 petabyte di dati (un petabyte è 10^{15} , cioè un 1 seguito da 15 zeri, e corrisponde a circa 1.000 hard disk da un terabyte).

Anche Karin Ljubič e il marito, Iztok Fister, nel loro lavoro producono moltissimi dati, perché si occupano di intelligenza computazionale e algoritmi applicati alla gestione degli allenamenti sportivi (oltre a essere ricercatori, Karin e Iztok si dedicano al triathlon). «È un problema reale», commenta Maurizio Casiraghi, professore associato di zoologia all'Università di Milano Bicocca, biotecnologo ed esperto di *Dna barcoding* (v. riquadro

La prima frase scritta nel Dna di una pianta (e poi riletta) è stata questa: "Hello world"



Getty Images/SPL

nella pagina accanto). «Anche soltanto l'estrazione del Dna che studio produce molti gigabyte, che devono essere messi da qualche parte per essere usati».

FUORI DAGLI SCHEMI. Guardando fuori dalla finestra verso un bosco, in un giorno piovoso, Karin e Iztok pensarono che in natura esistesse già un metodo per accumulare informazioni: i genomi delle specie viventi, in particolare dei vegetali. In pochi grammi di Dna c'è spazio per miliardi di gigabyte. «Il Dna è perfetto per accumulare informazioni; la struttura a doppio filamento affiancato consente di scrivere un messaggio su entrambi. Anche quando si dividono (nel corso della riproduzione sessuale, ndr), i due filamenti contengono le stesse informazioni», spiega Ljubič. Tutto il processo non è però così facile: «Vorremmo proporre nuove idee per immagazzinare i dati, basate su tecnologie esistenti o in via di sviluppo. E un po' di pensieri *out of the box*, cioè fuori dai soliti schemi», conclude. Lo spunto arrivò da alcuni esperimenti che avevano già provato la possibilità di creare Dna artificiale con una sequenza particolare di basi (le unità di informazione), in modo da "scrivere" sulla

molecola. Per esempio nel 2012 George Church, chimico e genetista statunitense, aveva convertito un documento di oltre 53.000 parole, 121 immagini e un programma in una sequenza di Dna. Dopo la scrittura era però necessario riuscire a leggere il testo; lo si poteva fare con uno strumento per la decodifica. Purtroppo la molecola era molto lunga e, al tempo, leggerla portava a numerosi errori. Altri esperimenti l'anno successivo, condotti da Nick Goldman dell'European Bioinformatics Institute, dimostrarono che era possibile scrivere messaggi e leggerli fedelmente. Bastava creare testi più brevi, lunghi al massimo 117 caratteri, che risultavano più leggibili dalla macchina decodificatrice. In questo modo riuscirono a memorizzare 5 milioni di bit; per esempio tutti i 154 sonetti di Shakespeare e 26 secondi del celebre discorso *I have a dream* di Martin Luther King. Il sistema di codifica inventato da Goldman si basava su un concetto semplice: le parole di base del linguaggio dei computer (i cosiddetti byte, ognuno dei quali è formato da 8 bit di 0 e 1) erano trasformate in parole di cinque lettere, una combinazione formata a partire dalle quattro basi del Dna (v. schema nella



Getty Images

prossima pagina). In linea di principio tutto funzionava, ma nella pratica c'erano ancora alcuni problemi da risolvere. Il primo era: dove mettere questo Dna ricco di informazioni? Fino ad allora veniva immagazzinato in luoghi freddi e secchi, come accade per il genoma dei mammut morti in Siberia, il cui Dna rimane leggibile anche dopo 60.000 anni.

Un altro ostacolo all'uso della tecnologia era costituito dal fatto che scrivere nel Dna è un processo costoso, e ripeterlo

per molte volte può andare bene durante un esperimento, ma non è una procedura che si possa utilizzare tutti i giorni.

Ecco, quindi, l'idea di Karin e Iztok. Visto che il Dna è il materiale genetico di (quasi) tutti gli esseri viventi sul pianeta, perché non usare le stesse specie, vegetali in questo caso, per immagazzinare i dati?

CON I BATTERI. Dopo aver creato la sequenza di Dna attraverso un sistema di codifica simile a quello di Goldman, il ▶

IL CODICE A BARRE DELLA VITA

CHE SPECIE È? Una tecnica per leggere il Dna dei viventi esiste. E viene usata per determinare la specie studiata, partendo da un frammento minuscolo del patrimonio genetico. Come dice Maurizio Casiraghi, professore all'Università di Milano Bicocca, il cosiddetto *Dna barcoding* è l'analisi di un frammento di Dna mitocondriale (appartenente cioè al mitocondrio, un organello definito la "centrale energetica" della cellula). Basandosi sulle minuscole differenze che esistono tra le specie proprio in questo pezzetto di gene, la tecnica serve a scoprire da dove proviene un campione. Il frammento, lungo circa 600 "basi" (l'unità elementare), è confrontato con un database globale che racchiude il Dna di riferimento di molte specie. Il metodo serve, tra l'altro, a scoprire se il palombo o la salsa di pomodoro sono veramente quello che dichiarano i venditori. La sua semplicità ha indotto la Fda (l'ente che si occupa di cibo e medicine negli Stati Uniti) a richiedere il *barcoding* per molti prodotti importati in America.



Alamy/ippa

SIMBOLO VEGETALE. Raffigurazione metaforica di una pianta a forma di Dna.

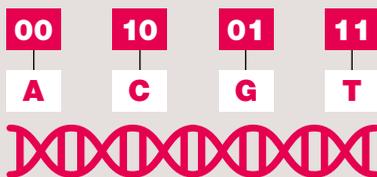
E I DATI SBOCCIANO.
A destra, i due ricercatori con le loro piantine. Sotto, un germoglio in cui sono stati inseriti i dati.



Karin Ljubič e Iztok Fister

SCRIVERE NEL DNA

CODIFICA. Per convertire la sequenza di un file, costituita da 0 e 1, in Dna, i ricercatori hanno trasformato il codice binario così:



A, C, T, G sono le “basi” (adenina, citosina, timina e guanina) che, in sequenza, costituiscono il Dna.

passo seguente è stato inserire questo Dna in un vegetale. Per farlo si usa di solito il batterio *Agrobacterium tumefaciens*, che a sua volta infetta la pianta inserendo del materiale genetico nelle sue cellule.

COME UNA LIBRERIA. Karin Ljubič e Iztok Fister sono riusciti a usare con successo questa tecnica con un parente del tabacco chiamato *Nicotiana benthamiana*. I frammenti di Dna sono stati introdotti in una specie di “gnocco” di tessuto spugnoso indifferenziato, che si trasforma pian piano in una pianta come le altre, da cui spuntano cioè nuove foglie e ramoscelli. Questa nuova pianticina conteneva quindi, oltre al suo patrimonio genetico, anche del Dna estraneo, ma innocuo; una specie di libreria microscopica ma ricchissima.

I due ricercatori sloveni promettono che il passo successivo sarà codificare l'intera Wikipedia inglese, che contiene 4,4 milioni di articoli per un totale di 42 gigabyte; che però è un'inezia, per esempio, rispetto ai 90 petabyte che, come abbiamo detto, il Cern produce in un anno. Per contenere questi ultimi ci vorrebbero 41 grammi di Dna. Sembra poco, ma non è così. «A dire la verità», chiosa Casiraghi, «non è proprio una quantità minuscola. Sarebbero necessarie parecchie cellule». Il sistema, comunque, risolve automa-

ticamente un problema importante, quello della conservazione e della riproduzione dei dati: ci pensano i vegetali a replicare, a ogni divisione cellulare, l'enciclopedia o quello che è stato introdotto nel materiale genetico.

Oltre alla moltiplicazione dei filamenti, la pianta si incarica anche di proteggerli dagli eventuali disturbi ambientali, meglio di quanto possa fare il freddo della tundra per le mummie di mammut.

Una volta immagazzinate nelle foglie, però, per essere utili, le informazioni devono essere anche leggibili. I metodi per estrarle dalla sequenza del Dna sono

tantissimi, ma prima o poi, secondo i ricercatori sloveni, avremo strumenti non dissimili da quelli dei film di fantascienza che ci permetteranno di farlo; realizzando così il sogno di Karin, di leggere il Dna vegetale anche mentre siamo serenamente immersi nella natura.

ATTENTI AI RISCHI. I coniugi Fister ammettono che il processo non è facile, e i particolari tecnici non sono ancora del tutto definiti. I due cominciano però a porsi domande più complesse, che riguardano il potenziale impatto di questa tecnologia. «Ci sono molti vantaggi, ma anche aspetti negativi, che riguardano per esempio la diffusione globale del processo», dice Ljubič. Il pericolo è che possa essere usato per scopi non proprio scientifici o ricreativi. Per esempio, per immagazzinare informazioni sensibili e a questo punto impossibili o quasi da ritrovare (immaginate di cercare, senza sapere dov'è, la cartella clinica di un determinato paziente in mezzo a una foresta). Oppure per diffondere informazioni segrete su obiettivi militari o altro ancora. Insomma, conclude Karin Ljubič, una tecnologia potentissima come questa avrebbe bisogno anche di una società equilibrata che la gestisca in modo responsabile. **F**

Marco Ferrari

Se funzionasse, tutti gli archivi del mondo potrebbero essere contenuti in una scatola di semi