



UNIVERSITÀ
DI SIENA
1240

Inaugurazione del 784° anno accademico

22 novembre 2024

Prolusione della Professoressa Luisa Bracci

Ordinario di Biochimica nel Dipartimento di Biotecnologie Mediche

Decodificare il linguaggio delle molecole per poter interagire con loro (senza averne paura)

I nostri sensi, pur sviluppati, hanno dei limiti. La vista ci consente di percepire immagini entro definiti limiti di dimensioni, spazio e luce, ma con strumenti tecnologici adeguati il campo delle possibilità di visione si amplia. La tecnologia ci permette di vedere ad esempio le cellule usando un microscopio, ed addirittura seguire le cellule vive che si muovono, con microscopi più raffinati. L'utilizzo di strumenti e tecnologie permette inoltre di ampliare le nostre capacità di osservazione, oltre ciò che possiamo percepire con i nostri sensi.

Siamo abituati a considerare osservare e vedere come sinonimi, ma in realtà non lo sono. Nella scienza è possibile osservare e monitorare fenomeni biologici, senza vederli.

La vita di tutti gli organismi si basa su interazioni fra molecole. Quelle che vengono indicate generalmente come biomolecole, ovvero acidi nucleici, proteine, lipidi e glucidi, interagiscono tra di loro e con molecole più piccole (sia molecole naturali, come per esempio neurotrasmettitori, che anche farmaci) generando conseguenti complesse catene di eventi dai quali dipende la vita delle cellule e degli organismi, che dalle cellule sono composti.

I fenomeni biologici e le loro modificazioni, sia fisiologiche che patologiche, sono tutti regolati da complesse serie di interazioni fra molecole, che possiamo visualizzare in minima parte.

una precisa ricostruzione delle interazioni fra molecole, che colleghi struttura e funzione, può permettere di comprendere come le molecole si riconoscono e così immaginare come si possa intervenire per correggere qualcosa che non funziona.

Per ricostruire la funzione delle molecole e le loro interazioni è necessario decodificare una sorta di linguaggio molecolare, ovvero comprendere le basi chimico-fisiche che permettono alle molecole di interagire fra loro in modo specifico, coordinandosi in una rete complessa, ma rigorosamente ordinata, di reazioni. Imparare a leggere il linguaggio che permette alle molecole di riconoscersi in modo specifico e coordinato, non è soltanto una questione di curiosità scientifica. È la base per esercitare quella che si chiama adesso Medicina Personalizzata, il cui scopo finale è curare ogni paziente in modo mirato, dando medicine specifiche per il problema che ognuno ha. Ne parleremo alla fine, ma occorre subito ricordare che ogni patologia, sempre e comunque, presuppone qualcosa che non va nel modo giusto, e questo qualcosa è sempre un processo molecolare.

Il livello meno complesso del linguaggio delle molecole è quello che conosciamo meglio e da vari decenni, ovvero il **CODICE GENETICO**.

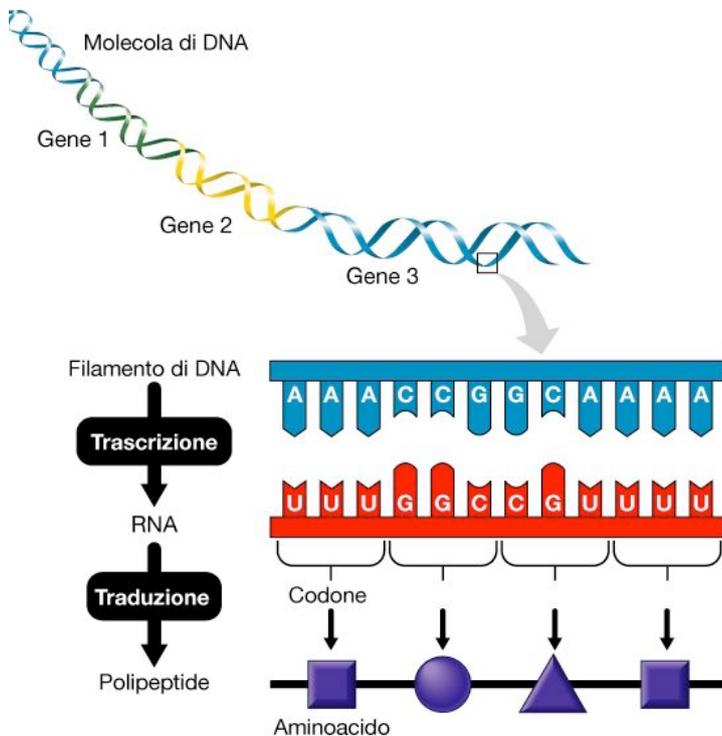
Gli acidi nucleici contengono le informazioni genetiche che trasmettono grazie ad un codice costituito da triplette costruite con 4 diverse basi azotate. Le triplette di basi del DNA vengono trascritte su un filamento di RNA, grazie ad una precisa interazione che accoppia ogni base ad una che le è complementare. Ogni tripletta viene poi tradotta in un aminoacido. Gli aminoacidi compongono una proteina, la cui sequenza è quindi scritta nella sequenza di basi del DNA. Gli acidi nucleici hanno un linguaggio costruito su un alfabeto di 4 lettere. Le proteine usano un linguaggio completamente diverso, costruito su 20 differenti aminoacidi. Agli aminoacidi è affidato un codice molto più complesso che regola la struttura, e di conseguenza la funzione, delle proteine.

Il codice genetico ci permette di predire facilmente la sequenza degli aminoacidi di una proteina ma predire quale sarà poi la struttura che questa proteina assumerà, e quindi la sua funzione, è molto più difficile.

Conosciamo le regole chimiche che permettono le interazioni fra i differenti aminoacidi. Si chiamano **legami deboli**, coinvolgono le diverse catene laterali degli aminoacidi, e sono responsabili sia dell'avvolgimento delle proteine nella loro struttura attiva che della capacità delle proteine di interagire con altre molecole, biomolecole ma anche farmaci. Nonostante si conoscano tutte le leggi che permettono l'avvolgimento di una proteina, la variabilità è così elevata che predire la struttura di una proteina semplicemente dalla sua sequenza è molto difficile. La struttura di una proteina è direttamente responsabile della sua funzione; quindi, sulla

base della sola sequenza, è spesso impossibile predire la funzione di una proteina e predire come e con quali altre molecole una proteina possa interagire.

Le interazioni fra biomolecole possono essere ricostruite in vitro con tecnologie come NMR e raggi x o anche microscopia elettronica ma ognuna di queste richiede tempo, disponibilità di materiale, spesso purificato, e non sempre sono applicabili. Se



potessimo interpretare il linguaggio delle molecole con una accuratezza sufficiente per predirne, sulla base della sequenza, la struttura, ed anche la complementarità e possibilità di interazione con altre biomolecole, potremmo accelerare e semplificare gli interventi terapeutici che potrebbero essere necessari.

Tutte le informazioni che abbiamo sulla chimica del linguaggio molecolare, che regola struttura e funzione delle proteine, sono state associate alla numerosissima collezione di dati su strutture di proteine, derivate da dati sperimentali, per costruire degli algoritmi che permettessero di ottenere informazioni sulla struttura e funzione di proteine, a partire dalla loro sequenza di aminoacidi. Si è usata la **Intelligenza Artificiale** per ricostruire la struttura di proteine e le loro interazioni. Ricordando che ogni fenomeno biologico coinvolge proteine, è evidente la importante ricaduta applicativa di questo approccio, che permette di sviluppare interventi terapeutici per regolare le interazioni alla base di eventuali patologie.

Non a caso, il premio Nobel per la Chimica del 2024 è andato agli inventori di un sistema di intelligenza artificiale che consente di ricostruire la struttura di proteine a partire dalla sequenza di aminoacidi, ed anche di predire efficacemente la loro funzione, incluso la capacità di interagire con altre molecole. Il sistema di Intelligenza Artificiale così costruito permette anche di essere usato 'all'inverso', per

costruire proteine realistiche e funzionanti, a partire da una desiderata funzione. Ovvero costruire ‘allucinazioni’. In un certo senso, il sistema di elaborazione dei numerosissimi dati con i quali lo si è ‘istruito’, somiglia alla capacità del nostro cervello di sognare.

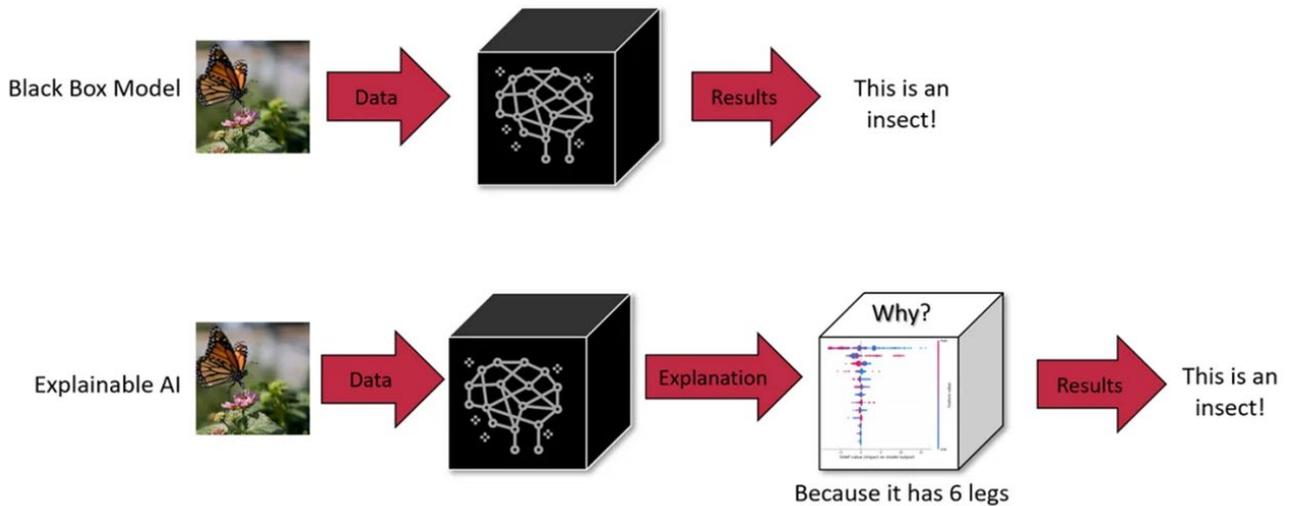
L’intelligenza Artificiale ha cambiato profondamente l’approccio classico alla analisi di un problema scientifico. Il classico metodo scientifico comporta generalmente come punto di partenza la formulazione di una ipotesi, seguita da esperimenti per raccogliere dati che potranno risultare congruenti con la ipotesi o, qualora invece non concordi, porteranno alla formulazione di una seconda ipotesi e ripetizione del ciclo.

La Intelligenza Artificiale procede con un andamento essenzialmente opposto. Si parte da raccolta di un numero elevatissimo di dati ed è poi il sistema che associa ed elabora i dati, portando a conclusioni, che non sono direttamente associate ad una ipotesi specifica.

Questo approccio ha inevitabilmente condotto a discussioni relative all’impatto che conclusioni generate come se le avesse prodotte un oracolo, possono avere sull’avanzamento della conoscenza scientifica.

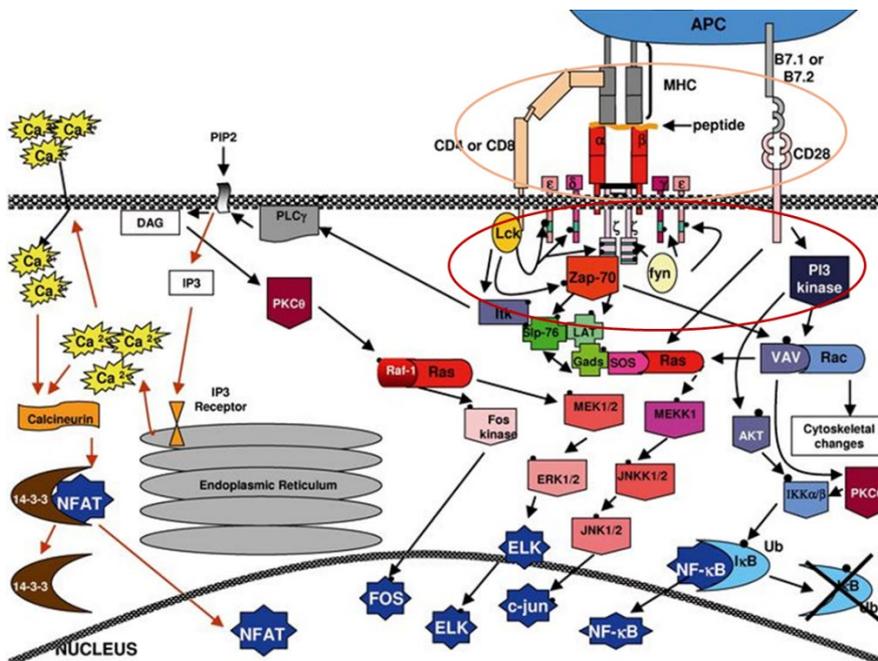
“An oracle that correctly predicts the outcome of every particle physics experiment, the products of every possible chemical reaction or the function of every protein would revolutionize science and technology. However, scientists would not be entirely satisfied because they would want to comprehend how the oracle made these predictions. This is scientific understanding, one of the main aims of science. With the increase in the available computational power and advances in artificial intelligence, a natural question arises: how can advanced computational systems, and specifically artificial intelligence, contribute to new scientific understanding or gain it autonomously?” [1]

Nella maggioranza dei casi, la Intelligenza Artificiale lavora come una **Black Box**, fornendo conclusioni senza dare informazioni su come la analisi dei dati abbiamo permesso di arrivare a queste. In altri casi, è almeno in parte possibile, per i fruitori, avere informazioni su quali siano i passaggi base del sistema che permettono la elaborazione dei dati e le conseguenti conclusioni.



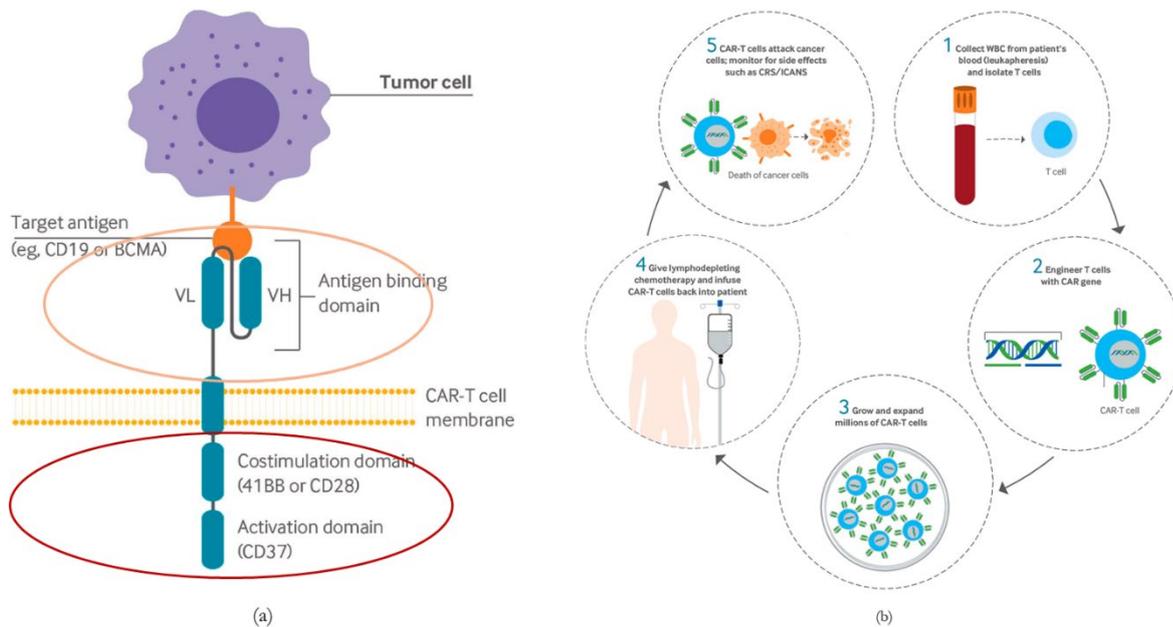
Nonostante le perplessità che la AI può suscitare, anche negli stessi scienziati che la utilizzano, la possibilità di accelerare enormemente la identificazione delle basi molecolari di fenomeni biologici, ha delle ricadute di importanza fondamentale nella medicina. Per convincercene, facciamo alcuni esempi, che riguardano applicazioni cliniche nel campo oncologico.

La prima è una sorta di ‘2001 Odissea nello spazio’ delle tecnologie molecolari, per la sua innovatività, quasi fantascientifica. Si tratta della **CAR-T Cell therapy** (CAR sta per Chimeric Antigen Receptor), dove l’utilizzo delle informazioni molecolari su struttura e funzione di proteine coinvolte nella risposta immunitaria è stata fondamentale per costruire farmaci innovativi.



La CAR T Cell therapy permette di modificare linfociti T di un paziente in modo da indirizzare la loro risposta (anche citotossica, killer) contro cellule tumorali. Questo approccio è stato possibile grazie alla conoscenza dei complessi meccanismi

molecolari che determinano la risposta immunitaria contro patogeni. Con tecnologie molecolari è stato possibile disegnare recettori delle cellule T ‘chimerici’ capaci di iniziare la catena di eventi molecolari, responsabile della attivazione dei linfociti, partendo da un innesco ‘artificiale’, ovvero un frammento di anticorpo diretto verso un antigene espresso selettivamente dalle cellule tumorali, anziché dalla complessa interazione con altre cellule del sistema immunitario, come succede nella risposta contro patogeni.



La CAR T Cell therapy è un raffinato esempio di come la conoscenza del linguaggio delle molecole ci possa permettere di ‘interagire’ con loro. Conoscendo il loro linguaggio, è possibile modificare le interazioni molecolari per scopi terapeutici.

Un’altra terapia antitumorale, meno ‘fantascientifica’ ma anche di più immediata applicazione, che è stata resa possibile dalla comprensione dei meccanismi molecolari della risposta immune, è la **Immunoterapia tramite i Check point inhibitors**.

Lo studio dei complessi dei meccanismi molecolari della risposta immunitaria ha portato alla identificazione dei checkpoints, ovvero interazioni specializzate fra proteine di membrana delle cellule, che permettono di bloccare la eccessiva attivazione e proliferazione dei linfociti reattivi contro patogeni, ma purtroppo anche di quelli (già di per sé poco numerosi) che potrebbero uccidere le cellule tumorali. I checkpoint inhibitors, già correntemente utilizzati nella immunoterapia di vari tumori, sono anticorpi che bloccano i checkpoints, prolungando attivazione e proliferazione dei linfociti.

Ricostruire i dettagli molecolari degli eventi biologici, ed i correlati patologici, è fondamentale per la **Medicina di Precisione**, che è essenzialmente sinonimo di **Medicina Personalizzata**, il cui scopo è quello di fornire ad ogni paziente farmaci mirati, che agiscano sullo specifico evento molecolare che determina la malattia. Dal momento che la osservazione dall'esterno non è sufficiente per determinare quale sia il percorso molecolare patologico da aggredire, la Medicina di Precisione si basa sulla comprensione dei meccanismi molecolari delle malattie, al fine di distinguere gruppi diversi di pazienti, che potranno beneficiare di differenti farmaci, in base al tipo di 'quadro' molecolare che hanno. L'osservazione comune che pazienti con magari la stessa diagnosi clinica e sintomi simili, hanno risposte anche molto diverse ai farmaci, ha sempre una causa molecolare.

La identificazione del corretto trattamento per ogni paziente, può richiedere più cicli di analisi, ovvero raccolta di dati molecolari sempre più mirati. Questo processo di ottimizzazione costituisce la base della **Medicina Traslazionale**, che viene riassunta bene dalla espressione: dal banco di laboratorio al letto del malato e ritorno.

In laboratorio si raccolgono dati, si elaborano ipotesi, si fanno tutte le verifiche sperimentali e alla fine si trova un possibile farmaco per una determinata malattia. Per quanto accurata possa essere stata la raccolta di dati sul paziente per permettere di ricostruire un quadro molecolare, può succedere che, durante il trattamento ancora alcuni pazienti rispondano meno bene al farmaco. Quindi si torna in laboratorio, si analizzano ancora più nel dettaglio le caratteristiche molecolari dei gruppi di pazienti che hanno beneficiato diversamente del farmaco, cercando di definire meglio quali siano le caratteristiche molecolari differenti fra chi risponde bene e chi no.

In Toscana grazie ai finanziamenti del PNRR è stata costituita una rete che comprende tutte le università toscane ed altre istituzioni. La rete è organizzata in 10 diversi Spoke, dei quali 2 sono coordinati dall'Università di Siena e sono proprio due spoke correlati: Medicina di Precisione e Medicina Traslazionale.

Voglio chiudere con un accenno al fenomeno dello scetticismo nei riguardi della scienza. Le molecole, le loro relazioni, le reazioni biologiche, sono diventate qualcosa che fa paura, ed allo stesso tempo, dal momento che non si vedono, vengono a volte assimilate a concetti, idee, dove chiunque può esprimere un'opinione.

Si va verso una dicotomia che dovremo gestire, tra l'approccio della Intelligenza Artificiale dove sono solo i dati ad avere rilievo e dove, addirittura, si possono trarre conclusioni, senza partire da una specifica ipotesi (e magari nemmeno fornendo spiegazioni dirette sul come, partendo dai dati si sia arrivati a quella conclusione) e, dall'altra parte, un approccio opposto dove si parte da ipotesi e conclusioni, senza appoggiarsi su dati scientifici reali e poi, sulla base di opinioni preconcepite, si cerca di giustificare pochi dati, che spesso non sono nemmeno scientificamente dimostrati.