

La bioinformatica nell'era della medicina traslazionale e personalizzata

Le tecnologie di sequenziamento ad alte prestazioni (Next Generation Sequencing, NGS) sono ormai diventate di uso comune nella ricerca biomedica, ma trovano sempre più spazio nella diagnostica, contribuendo fortemente alla realizzazione della medicina traslazionale e gettando le basi per la futura medicina "personalizzata".

Tra le varie applicazioni del NGS occupano un posto di rilievo il sequenziamento dell'intero genoma umano, l'analisi dell'espressione genica in cellule umane esposte a determinati stimoli, l'analisi metagenomica per identificare la componente microbiologica (microbioma) e virologica (viroma) di vari ecosistemi, nonché di diversi distretti corporei. Ovviamente, questo non sarebbe possibile senza il continuo progredire delle tecniche di sequenziamento NGS, il quale ha portato a una riduzione dei costi, sia in termini di tempo che di denaro. Infatti, se nel 2001 il primo sequenziamento del genoma umano ha richiesto una spesa di circa 3 miliardi di dollari e un decennio per l'analisi dei dati, oggi i sequenziatori NGS presenti in commercio e i computer ad alte prestazioni (HPC) portano questi numeri a meno di 1000 Euro e un solo giorno per l'analisi dei dati. Tuttavia, lo sviluppo tecnologico in questo senso sembra essere inarrestabile, visto l'elevato numero di dispositivi sempre più performanti che vengono immessi sul mercato. Negli ultimi 25 anni ha fatto il suo ingresso tra le scienze della vita una nuova disciplina che unisce la biologia molecolare con la matematica, la statistica e l'informatica, ovvero la **bioinformatica**. Il focus di questa disciplina è l'analisi della mole enorme di dati prodotti da sequenziamento massivo, dell'ordine delle decine di gigabyte per singolo campione, e si concentra sullo sviluppo e l'applicazione di software specializzati nell'acquisire, gestire, analizzare e/o visualizzare questi *big data* biologici. Ad oggi il suo ruolo nella medicina è considerato complementare (se non equivalente) alla biostatistica, in cui l'utilizzo di software specializzati, linguaggi di programmazione e computer ad alte prestazioni la fanno da padrone.



FRANCESCO MESSINA



MARIA ROSARIA CAPOBIANCHI
Laboratorio di Virologia, Istituto Nazionale per le Malattie Infettive "L. Spallanzani" IRCCS, Roma

I metodi sviluppati in ambito bioinformatico giocano un ruolo fondamentale nello sviluppo delle biotecnologie ed è per questo motivo che in molti Paesi si assiste ad ingenti piani d'investimento in questa direzione. Già alla fine degli anni Settanta si cominciava a parlare di organizzare ed immagazzinare dati biologici, ma serviva uno sforzo informatico per sfruttarne le potenzialità. La banca dati dell'EMBL (European Molecular Biology Laboratory) nasce nel 1980, mentre nel 1986 viene fondato negli Stati Uniti il National Center for Biotechnology Information (NCBI), depositario della banca dati GenBank, parte integrante del NIH e destinatario di investimenti volti al suo mantenimento. Nel 1996 si assiste alla creazione dell'EBI (European Bioinformatics Institute di Hinxton, Cambridge) che è oggi tra i punti di riferimento più importanti della bioinformatica mondiale. L'EBI è destinatario di ingenti finanziamenti da parte della Commissione europea, dai Governi dei 21 Stati membri dell'EMBL e dai più prestigiosi centri di ricerca europei. In Italia il numero di strutture sanitarie che affrontano sfide relative ai *big data* biologici per scopi diagnostici e di ricerca, sta aumentando rapidamente.

Dal momento che il contributo della bioinformatica alla medicina sta diventando così rilevante, sorge la necessità di definirne gli ambiti e i contenuti, così come il percorso formativo di questa disciplina. Il termine "bioinformatico" è spesso usato per descrivere qualcuno che lavora in vari ambiti della biologia o medicina utilizzando degli strumenti informatici, ma originariamente questo termine descriveva lo studioso della complessità dei sistemi biologici attraverso strumenti informatici. Gli obiettivi generali della bioinformatica possono essere generalmente classificati in tre ambiti: sviluppo di nuovi algoritmi analitici, usando linguaggi di programmazione, modelli matematici e simulazioni *in silico*; costruzione e gestione di database computazionali attraverso la raccolta e l'organizzazione di dati medici e biologici; *data mining*, interpretazione e analisi dei dati per estrapolare il significato biologico dai dati grezzi. Il raggiungimento di tutti questi obiettivi richiede numerose competenze in ambito informatico, ma anche, e soprattutto, una profonda comprensione nei diversi aspetti delle scienze della vita. Come si pone la nuova figura professionale del bioinformatico nei confronti dei ruoli convenzionali previsti dal sistema sanitario? Ad oggi all'interno del Sistema Sanitario Nazionale non è prevista una figura professionale con un ruolo di-

rigenziale riconducibile al bioinformatico. Il mantenimento di questa professionalità all'interno delle strutture ospedaliere è affidato esclusivamente alla capacità dei dirigenti di reperire fondi per la ricerca, oppure ad attitudini dilettantistiche dei dirigenti di estrazione biologica, chimica, fisica o simile. La difficoltà più grande sembra essere l'identificazione dei requisiti minimi per l'integrazione di questa nuova classe di operatori sanitari altamente specializzati nel Sistema Sanitario Nazionale. Infatti, attualmente la maggior parte di coloro che si avvicinano alla bioinformatica provengono da vari ambiti universitari (scienze biologiche, matematica, fisica, informatica) ed acquisiscono le proprie abilità "sul campo" o attraverso una formazione post-laurea (corsi di formazione, master di II livello, ecc.). In Italia, pochi atenei offrono corsi di laurea incentrati su questo settore e non esiste ancora una scuola di specializzazione. In molti casi il dottorato di ricerca, oppure un'esperienza presso rinomati centri di ricerca all'estero, sono un valore aggiunto per la formazione di un bioinformatico competente, specialmente se si tratta di approcci innovativi. La natura multidisciplinare della bioinformatica richiede che i praticanti siano disposti ad acquisire continuamente nuove conoscenze, come linguaggi di programmazione o comprensione di processi biologici, utili a affrontare le problematiche per l'ottenimento del risultato. Le continue richieste per la gestione di casi clinici sempre più complessi sono una sfida per le abilità dei bioinformatici, ma allo stesso tempo rendono difficile riuscire a standardizzare le metodiche utilizzate nei diversi casi. Tuttavia, la profonda comprensione dei dati può contribuire alla risoluzione di casi molto complessi nella pratica clinica, come testimoniato dalla analisi metagenomica in NGS su pazienti con patologie a sospetta eziologia infettiva. In ultima analisi, il bioinformatico è una professionalità che sta acquisendo progressivamente importanza nelle applicazioni, per il momento classificate come avanzate, della diagnostica di laboratorio; visto che questa figura sarà sempre più richiesta in futuro, è opportuno che il percorso formativo del bioinformatico sia definito secondo standard predefiniti, in base al ruolo che si intende far assumere a questa figura emergente nel panorama delle nuove professionalità sanitarie, al fine di garantire che il Sistema Sanitario Nazionale possa cogliere a pieno i benefici e possa essere in posizione trainante nella rivoluzione della medicina personalizzata fondata sui *big data*.