

# Progetto di ricerca

## Analisi statistica e Teoria dei Network applicata a dati high-throughput genomici ed epigenomici in oncoematologia

### Base scientifica

Le ricerche degli ultimi anni in ambito biomedico hanno generato, anche grazie all'introduzione delle tecnologie "high throughput", una nuova visione delle scienze della vita che richiede una **concettualizzazione fisico-matematica** basata su idee come **complessità, interazione, robustezza dinamica ed evoluzione temporale**. Negli ultimi anni questo paradigma, denominato **Biologia dei Sistemi**, ha favorito la nascita di diversi centri di ricerca sia nazionali che internazionali. Di grande interesse è l'impatto sulle scienze mediche che spaziano dalle **Neuroscienze, alla Genomica, all'Immunologia e agli studi sull'Invecchiamento e sulla Nutraceutica**, solo per citarne alcune.

L'indagine dinamica dei processi molecolari a livello di singole cellule sta fornendo una nuova luce su fenomeni biologici ad alto impatto sociale e con fondamentali ricadute per la salute umana come **l'insorgenza di tumori, il controllo del ciclo cellulare, la senescenza replicativa, l'attivazione, da parte di segnali esterni, di geni specifici, l'attivazione dei linfociti T, e il controllo delle risposte cellulari in seguito a perturbazioni**.

Questi studi hanno mostrato che l'applicazione di metodi statistici e matematici avanzati (es. per la ricostruzione di reti di regolazione genica, o per l'identificazione di meccanismi biologici chiave) può permettere l'avanzamento in questi campi, sia in termini di nuove terapie altamente personalizzate, che per l'applicazione di terapie note in ambiti non ancora esplorati (es. altri tipi di tumore).

Parallelamente a queste indagini di tipo modellistico-teorico, si sono sviluppate **metodiche per l'analisi di dati biomedici complessi**, come quelli originati dalle metodiche "high throughput" (**genomica, epigenomica, proteomica e "deep sequencing"**). Questi dati richiedono l'impiego di tecniche statistico-computazionali nuove che tengano conto, ad esempio della non Gaussianità delle distribuzioni sperimentali e la forte interdipendenza degli elementi studiati (interazioni tra geni, proteine e metaboliti) e anche elaborazioni legate all'analisi di sequenze di DNA (DNA motifs localization and 3d configuration).

Il nostro gruppo è coinvolto in progetti europei quali HARMONY (Healthcare alliance for resourceful medicines offensive against neoplasms in hematology), che ha tra i suoi obiettivi:

- Elaborare i dati whole-exome e RNA sequencing acquisiti nel progetto HARMONY
- Sviluppare modelli per descrivere le distribuzioni statistiche ottenute dai dati sperimentali
- Integrare i dati genomici con dati epigenomici (Metilazione DNA, struttura 3d cromatina)
- Integrazione dei dati a disposizione del consorzio con dati pubblici (es. ENCODE)
- Utilizzare la Teoria dei network complessi per scopi quali Drug Repurposing e identificazione di key targets & modules

## Piano di formazione

Il piano di formazione scientifica intende fornire gli strumenti teorici e pratici volti ad approfondire la competenza professionale necessaria per svolgere ricerche sul modelling e sull'analisi di dati biologici. Questa competenza professionale sarà maturata attraverso l'acquisizione progressiva degli elementi fondamentali che caratterizzano principalmente le seguenti attività:

- utilizzo delle tecniche statistico-computazionali per queste analisi (es. analisi di sequenze, first-time return analysis)
- impiego della Teoria dei Network (es. clustering, topological characterization, node/link ranking)
- caratterizzazione modellistica della struttura 3d del DNA associato ai tipi cellulari studiati
- raccolta dei dati e supporto nella gestione ed elaborazione dati all'interno del Consorzio

L'acquisizione degli elementi fondamentali tecnici caratterizzanti l'attività sopra descritta sarà ulteriormente approfondita mediante:

- discussione collegiale dei risultati, dei piani di ricerca e dei protocolli di calcolo;
- approntamento di rapporti, sommari, scritti esplicativi, ecc.;
- ricerche bibliografiche dirette o attraverso banche informatizzate;
- supporto all'organizzazione generale della facility di calcolo;
- acquisizione delle norme fondamentali sulla sicurezza informatica e la protezione dei dati
- istruzione tecnica degli studenti interni del laboratorio

Resta convenuto che:

- la soprastante elencazione delle tecniche ha valore indicativo e potrà essere modificata, pur nell'ambito del filone generale della ricerca, in relazione alle necessità contingenti,
- l'attività di ricerca si svolgerà presso il Dipartimento di Fisica dell'Università di Bologna,
- i dati originati nell'ambito di questa attività sono di proprietà del gruppo di ricerca proponente, e potranno essere divulgati soltanto con il consenso del Consorzio HARMONY.