Bioinformatics

The Bioinformatics unit of the Italian Institute of Genomic Medicine address a wide range of biological issues arising from genomic, transcriptomic and epigenetic experiments. Nowadays, next generation sequencing (NGS) studies are delivering an extraordinary amount of information at an unprecedented speed in the history of biology and medicine. Disentangling the complexity proper of these genomic Big Data to identify key-regulatory biological processes remains challenging. At IIGM, we apply state-of-the-art bioinformatics, statistical and mathematical approaches to extract the proper description, interpretation and understanding of biological processes from genomic Big Data. Our goal is to develop new tools and paradigms for comparing large sets of genomic data and identifying previously hidden patterns associated with a disease, with particular focus on onco-genomics. In doing so, we perform analysis aim at integrating publicly available and *in-house* generated NGS data set to improve the quality of our results. The unit has extensive expertise in mathematics, statistics, scientific programming, database technology, experimental design, genomics and molecular biology. Our computational tools and softwares are designed to be highly parallelizable and to fully exploit the most modern high performance computing (HPC) infrastructures.

We analyse any type of NGS data, such as:

* DNA sequecing: Whole Exome, Whole Genome, Targeted sequencing to identify genomic variants;
* RNA sequencing  to determine transcript abundance and detect alternative spliced variants;
* ChIP sequencing to identify active transcription factor binding dites or to map specific modifications of histone proteins;
* iCLIP sequencing to characterise the regulatory mode of RNA binding proteins;
* Bisulphite sequencing to profile DNA methylation;
* Single Cell Sequencing to understand the transcriptome at single cell level.

Our skills also include:

* data visualisation
* web development
* computer programming expertise in R, Python, C++, Java, Perl, PHP and Bash
* storing of big data.

L’unità di Bioinformatica dell’Istituto Italiano di Medicina Genomica affronta una vasta gamma di problemi biologici derivanti da esperimenti genomici, trascrittomici ed epigenetici. Oggigiorno, gli studi di sequenziamento di nuova generazione (NGS) stanno fornendo una straordinaria quantità di informazioni ad una velocità senza precedenti nella storia della biologia e della medicina. Districare la complessità propria di questi Big Data genomici per identificare i processi biologici regolatori chiave rimane difficile. In IIGM applichiamo i più recenti approcci bioinformatici, statistici e matematici per fornire la descrizione, l’interpretazione e la comprensione appropriata dei processi biologici dai Big Data genomici. Il nostro obiettivo è quello di sviluppare nuovi strumenti e paradigmi per confrontare grandi serie di dati genomici e identificare modelli precedentemente nascosti associati a una malattia, con particolare attenzione alla onco-genomica. In tal senso, eseguiamo analisi finalizzate all’integrazione di set di dati NGS disponibili pubblicamente e generati internamente per migliorare la qualità dei nostri risultati. L’unità ha una vasta esperienza in matematica, statistica, programmazione scientifica, tecnologia di database, progettazione sperimentale, genomica e biologia molecolare. I nostri strumenti e software di calcolo sono progettati per essere altamente parallelizzabili e per sfruttare appieno le più moderne infrastrutture di calcolo ad alte prestazioni (HPC).

Analizziamo qualsiasi tipo di dati NGS, come ad esempio:

* DNA sequencing: Whole Exome, Whole Genome, sequenziamento mirato per identificare le varianti genomiche;
* RNA sequencing per determinare l’abbondanza del trascritto e rilevare varianti spliced alternative;
* ChIP sequencing per identificare i dites attivi del legame del fattore di trascrizione o per mappare specifiche modifiche delle proteine istoniche;
* iCLIP sequencing per caratterizzare la modalità normativa delle proteine leganti l’RNA;
* Bisulphite sequencing per profilare la metilazione del DNA;
* Single Cell Sequencing per comprendere il trascrittoma a livello di singola cellula.

Le nostre abilità includono anche:

* visualizzazione dei dati
* sviluppo web
* programmazione in R, Python, C++, Java, Perl, PHP e Bash
* storage di big data.