

DATA: 11 maggio 2011

TITOLO: *Sequenziamento e analisi bioinformatica del genoma umano*

ABSTRACT:

Dal 2005 sono state introdotte nuove e numerose tecniche di sequenziamento individuate come “next-generation sequencing”.

Durante il seminario dopo una rapida carrellata sull’insieme di queste tecniche si arriverà al dettaglio sui sistemi Illumina Genome Analyzer e HiSeq2000 entrambi presenti al CRS4. Queste macchine “ad alta processività” permettono di effettuare un grande numero di analisi parallele su uno o più campioni di DNA e sono in grado di produrre un numero elevato di sequenze genomiche in pochissimo tempo e a costi ridotti rispetto alle macchine di generazione precedente.

L’enorme mole di dati prodotta in ogni processo, dell’ordine del Terabyte (10¹² byte), viene efficacemente gestita, memorizzata ed elaborata attraverso i sistemi informatici disponibili nel centro di calcolo del CRS4.

Saranno inoltre presentati in dettaglio, i protocolli di analisi bioinformatiche sviluppate al CRS4 in collaborazione con i colleghi informatici e biologi. Si metterà l’accento sui metodi di allineamento di sequenze, sul “management” di un formato generico di dati e anche sulla parte relativa al controllo di qualità, indispensabile per garantire l’accuratezza del dato elaborato.

RELATORE: Frederic Reinier, CRS4



Frederic Reinier è ricercatore nel gruppo di Genomica Avanzata (AG).

Ha conseguito la laurea in Bioinformatica nel 2005 all’Università di Marsiglia, Francia. Ha lavorato con Jean-Michel Claverie e Cedric Notredame presso il laboratorio di Informazioni Strutturali e Genetiche, CNRS (Marsiglia). Ha collaborato allo sviluppo e pubblicazione di PROTOGENE, un server che può trasformare l’allineamento multiplo di una proteina nell’allineamento equivalente del gene codificante del DNA originale.

Lavora al CRS4 dal 2006. I suoi interessi di ricerca includono analisi dei dati di high-throughput sequencing. Si occupa dello sviluppo di programmi per la piattaforma di sequenziamento, principalmente allineamento di sequenze, controlli di qualità e variant-calling.

Collabora con l’Istituto di Neurogenetica e Neurofarmacologia del Consiglio Nazionale delle Ricerche (INN-CNR) con il ruolo di programmatore chiave nei progetti di ricerca sul genoma sardo.

PAROLE CHIAVE: DNA, Sequenziamento, next-generation sequencing, allineamento, calcolo parallelo, calcolo ad alte prestazioni, qualità dei dati.