

**DATA:** 22 giugno 2011**16.30 - 18.00****LUOGO:** Aula Magna Dip. Fisica, Cittadella Universitaria (Monserrato)**TITOLO:**

## *L'inferenza statistica e la lettura dei dati*

**ABSTRACT:**

Negli ultimi anni si ha la possibilità di studiare il DNA con una risoluzione sempre più accurata, ma il costo per alcuni esperimenti è ancora inaccessibile, soprattutto se si intende condurre degli studi su centinaia o migliaia di soggetti.

Grazie ai concetti dell'inferenza statistica e della genetica si è capito che dei progressi possono essere fatti studiando in dettaglio il DNA di un numero relativamente ristretto di individui. Se si dispone ad esempio di un insieme di sequenze di riferimento, ottenute localmente o in altri laboratori, è possibile identificare la sequenza più simile all'individuo in studio, di cui si hanno solo dati parziali, confrontando le basi nucleotidiche alle posizioni comuni. Si può così sequenziare, probabilisticamente, il DNA di decine di migliaia di persone, abbattendo fortemente i costi di laboratorio e aumentando notevolmente la probabilità di identificare le componenti genetiche associate a malattie.

In questo seminario verranno illustrati i metodi statistici e i concetti genetici che stanno alla base di questo approccio, strumento ormai essenziale nell'era del sequenziamento genomico. Verranno discusse utilità e performances in diversi modelli sperimentali, e presentate le applicazioni sui progetti in collaborazione tra il CRS4 e il CNR.

**RELATORE:** Serena Sanna(CNR)

Serena Sanna, laureata in Matematica all'Università di Cagliari, ha iniziato il suo percorso di ricerca nel 2004, con il gruppo di genetisti dell'Istituto CNR di Cagliari, allora guidato dal Prof. Giuseppe Pilia.

Dal 2005 al 2007 ha completato un periodo di alta formazione presso il Centro di Statistica Genetica, all'Università del Michigan, dove ha appreso le tecniche di inferenza statistica e ha condotto, continuando anche al suo ritorno in Sardegna, diversi studi genetici sulla popolazione sarda.

Continua ora il suo interesse per i metodi di inferenza ed in particolare per la loro applicazione ai dati di sequenziamento dell'intero genoma.

**PAROLE CHIAVE:** GWAS, genetica, inferenza statistica, sequenziamento, HapMap, 1000 Genomes, polimorfismi, aplotipi, linkage disequilibrium.