



DIPARTIMENTO DI SCIENZE GINECOLOGICHE, OSTETRICHE E PEDIATRICHE

CATTEDRA DI GENETICA MEDICA  
Presso Pad. 11 - Policlinico S. Orsola-Malpighi  
Via Massarenti 9 - 40138 Bologna  
Tel. 051/2088418– Fax 051/2088416

**AVVISO DI SEMINARIO**

**Mercoledì 11 Novembre 2009**

**Presso U.O. Genetica Medica, Policlinico S. Orsola-Malpighi-Pad.11**

**ore 14.30**

**Dr. Can Alkan**

*(University of Washington and Howard Hughes Medical Institute, Seattle, WA, USA)*

**“Structural Variation Discovery and Characterization of Segmental  
Duplications with Next-Gen Sequencing Technologies”**

**Abstract**

Structural variation, in the broadest sense, is defined as the genomic changes among individuals that are not single nucleotide variants. These include insertions, deletions, duplications, inversions and translocations. Several genome-wide studies on human structural variation have demonstrated that structural variation among individuals is common and ubiquitous. Databases for normal structural variation, such as the Database of Genomic Variants, along with SNPs are needed in order to fully characterize genetic diseases in the future. Structural variation and segmental duplications (duplications of size  $\geq 1$  kb and sequence similarity  $\geq 90\%$ ) are an important source of evolutionary gene innovations. A variety of diseases have been associated (both causative and protective) with structural variation and segmental duplications such as schizophrenia, mental retardation, Crohn's disease, lupus glomerulonephritis, age-related macular degeneration, HIV susceptibility/resistance, color blindness and psoriasis. Structural variation and especially duplicated regions have remained largely intractable due to difficulties in accurately resolving their structure, copy number and sequence content using hybridization based methods. Consequently, a significant fraction of the duplicated genomic content has not been assayed by standard genetic and molecular analyses.

The realization of new ultra-high-throughput sequencing platforms such as Roche/454, Illumina/Solexa and ABI/SOLiD now makes it feasible to detect the full spectrum of genomic variation among many individual genomes, including cancer patients and others suffering from diseases of genomic origin. Recently we have developed a set of computational methods to comprehensively detect and characterize structural variation and segmental duplications using next-gen sequencing technology. We apply our algorithms to characterize structural variation and segmental duplications to genomes sequenced by Illumina and 454 technologies. We initially examine the genomes of three humans and experimentally validate copy-number differences in the organization of these genomes, and extend the application of our methods to study the genomes of  $>160$  individuals sequenced as part of the 1000 Genomes Project.