VanBUG

presents Guest Speaker:

Andrew Clark

Department of Molecular Biology and Genetics, Cornell University

Human genomics with large sample size and full genome sequences

The 1000 Genomes Project has completed its pilot phases including (1) low-coverage whole-genome sequencing of 179 individuals, (2) high-coverage sequencing of two mother–father–child trios; and (3) exon-targeted sequencing of 697 individuals from seven populations. These data present an unprecedented level of detail about the allele frequency spectrum and local haplotype structure of ~15 million SNPs, 1 M indels, and 20,000 structural variants across the human genome. On average, each person is found to carry approximately 250 to 300 loss-of-function variants in annotated genes. Data from the two trios produced an estimate of 1028 germline mutations per genome per generation. These data also present an opportunity to test for past signatures of natural selection and demographic effects. In a separate study we examined sequence variation in just two genes from the ARIC cohort study of 13,715 people. This analysis resulted in nearly 4x the expected count of singleton variants which were subsequently confirmed. These data are consistent with a very recent explosive population growth model that matches the historical and archeological record. The data are also consistent with the observation that each new full human genome that is sequenced typically finds upwards of 100,000 novel SNPs that are not yet in any database.

Introductory speaker (10 mins):

Rodrigo Goya, Marra lab, GSC, BCCA

Thursday, March 10, 2010, 6:00 pm

Gordon and Leslie Diamond Family Theatre,
BC Cancer Research Centre,
675 West 10th Avenue

CIHR/MSFHR BIOINFORMATICS TRAINING PROGRAM FOR HEALTH RESEARCH

MITACS

Genome BritishColumbia

10 a decade of innovation

http://vanbug.org