4th Conference of the African Society of Human Genetics
Addis Ababa, Ethiopia, 3rd-5th June 2006

SAMPLE ABSTRACT

ENHANCING THE UTILITY OF SINGLE NUCLEOTIDE POLYMORPHISM LOCATION TYPES IN THE HUMAN GENOME

R.D. Isokpehi, D.F. Sarpong and F.A. Okojie. Jackson State University, Jackson Mississippi 39217, USA. raphael.isokpehi@jsums.edu

Background
Single Nucleotide Polymorphisms (SNPs) in the human genome can contribute to population level genetic variations that are linked to disease susceptibility. High-throughput SNP genotyping techniques are yielding high-density maps of SNPs that can be used to study complex genetic diseases.

Objectives
The primary objective of this research is to classify and/or prioritize genes in the Human Genome for epidemiological and biological studies using binary integration techniques.

Methods
EnsMart (http://www.ensembl.org/) was used to retrieve Ensembl Gene ID of human gene entries that have the following SNP attributes(entries): Coding (18,143), Intronic (18,423), 5' UTR (6,813), 3' UTR (13,340), 5' Upstream (21,932), 3' Downstream (21,964), Synonymous SNPs (13,842), Non-Synonymous SNPs (15,155), Stop SNPs (981), and SNPs with a ka_ks ratio >0.5 (1,828). A binary integration computational pipeline that encodes the evidence for an attribute as I (present) or 0 (absence) was then used to generate 10-digit binary profiles for each of the over 22,000 gene entries in the ENSEMBL Human database.

Results
Of the possible 1024 gene clusters, 112 had at least one gene. The gene cluster abundant for all the samples attributes contained 66 genes including those not described in the On-line Mendelian Inheritance in Man (OMIM) database.

Conclusions
The binary integration of genomic data is one of the cost- and time-effective methods for reducing massive genomic data into gene clusters of biological or epidemiological relevance. These gene clusters could be useful in genome annotation, haplotype mapping and suggest future areas of genetic research.