## Characterization of Single Nucleotide Polymorphisms (SNPs) in the Korean Population

Seong-Gene Lee, Sunghee Hong, Yongsook Yoon, Insil Yang, Kyuyoung Song

Asan Institute for Life Sciences, Department of Biochemistry, University of Ulsan College of Medicine, Seoul 138-736, Korea

The aim of genetics is to correlate specific molecular variations with phenotype changes. As the human genome sequence was available, it became possible to understand the spectrum of genetic variation in the human gene pool and its relation to diseases, individual responses to environmental factors, and biological processes such as development and aging. As the SNP identification process accelerates, it becomes necessary to characterize a large number of publicly available SNPs in a population.

To construct Korean SNP database, we have determined the allele frequencies of 3000 cSNPs from 1,831 genes and 1,053 ESTs selected from the public database in 24 individuals using a pooled DNA sequencing approach. Of those, thirty six percent of SNPs were monomorphic in the Korean samples. Among polymorphic SNPs, 590 SNPs (30.6%) were uncommon (minor allele <20%), 1,338 SNPs (69.4%) were common (minor allele  $\geq 20\%$ ) in our population. We also compared minor allele frequencies of 406 SNPs derived from 120 genes in Korean with the publicly available data obtained with a mixture of Europeans, Asians, African Americans, and African Pigmies. Of those, fifty seven percent of SNPs (233 markers) have minor allele frequency < 5% in the Korean, whereas 45.8% (186 markers) in the mixture. Of 233 SNP markers, 25% causing non-synonymous amino acid changes, 22% causing synonymous amino acid changes, 12% derived from non-coding regions were common both in the Korean and the mixture samples. Of markers with minor allele frequencies > 5%, about 20% were common regardless of SNP types. Our data suggests that an extensive SNP characterization would be necessary before selecting SNPs informative in diverse populations.