Characterization of Publicly Available SNPs in the Korean Population

Seong-Gene Lee¹, Sunghee Hong², Yongsook Yoon³, Insil Yang², Kyuyoung Song¹²

¹Asan Institute for Life Sciences, ²Department of Biochemistry, University of Ulsan College of Medicine, Poongnap-Dong, Songpa-Gu, Seoul 138-736, Korea

The central aim of all genetics is to correlate specific molecular variation with phenotype changes. As the human genome draft sequence was announced in June, 2000, 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. Geneticists have so far used familial linkage methods for studying simple single-gene diseases. However, complex, high-incidence, multi-gene diseases are thought to require a different approach, for which SNPs can play a key role. Single-nucleotide polymorphisms (SNPs) are the most abundant form of genetic variation and have a great potential for mapping complex genetic traits. Due to its potential for being used as genetic markers, a great deal of effort has been invested in the identification of SNPs, and as of May 2, 2001, approximately 2,841,309 SNPs had been deposited into public databases (http://www.ncbi.nlm.nih.gov/SNP/) compared to 7,000 SNPs in April, 1999. For an association mapping study, SNP allele frequencies in the population would be critical. As a first step toward construction of SNP database of Korean population, we have determined the allele frequencies of over 300 eSNPs from 200 genes selected from the public database. Of those, twenty percent of SNPs were monomorphic in the Korean samples. Among polymorphic SNPs, 69 SNPs (28.6%) were not informative (minor allele <20%), 172 SNPs (71.4%) were informative (minor allele ≥20%) in Korean population. The results suggest that the ethnic and population-based differences should be considered in the selection of SNPs for the study of complex diseases with association mapping methods.