Population-Specific Genetic Risk Stratification Using the Korea Biobank Array v2.0 in Koreans

Inhae Jeong[1], Mi Yeong Hwang[1], Min Jin Seo[1], Young Jin Kim[1], Bong-Jo Kim[1]
[1]Division of Genome Science, Department of Precision Medicine, National Institute of Health, Chungcheongbuk-do 28159, Republic of Korea

To facilitate the precision medicine research in Koreans, Korea Biobank Array (KBA) v2.0 with 1.66 millions markers was designed expanding its predecessor KBA v1.1 by including clinically actionable content (132K clinically actionable and 489K functional). These variants were curated from clinical databases, as well as whole-genome sequencing data from approximately 28,000 Koreans and East Asians. This study evaluated the genetic risk stratification utility of KBA v2.0 in Koreans. KBA v2.0 showed higher usability in East Asians (87%) compared to PangenomiX (84.6%) and GDA (69.4%), reflecting its improved design for this population. Among 14,490 genotyped samples, 1.46 million variants passed quality control. We identified 88,056 deleterious, 12,063 pathogenic, and 12,509 pharmacogenetic variants. Carrier rates for ultra-rare variants (MAF < 0.01%) were high-79% (deleterious), 38% (pathogenic), and 23% (pharmacogenetic)-highlighting the array's utility in moderate-sized cohorts. A genome-wide association study (GWAS) of lipid traits using KBA v2.0 replicated known associations and showed stronger phenotypic effects for functionally annotated variants. Polygenic risk score (PRS) stratification revealed up to a 2.1-fold difference, demonstrating the utility of combining rare allele with PRS: protective alleles more strongly reduced lipids levels in low-PRS individuals, while risk alleles further elevated lipid levels in individuals with already higher lipid levels caused by high-PRS. These findings demonstrate that KBA v2.0 is a high-resolution, Korean-specific platform for rare variant discovery and personalized risk prediction, advancing precision medicine in East Asian populations.