Prevalence of Genetic Variants in the United States: Results from the Third National Health and Nutrition Examination Survey (NHANES III), 1991-1994

Man-huei Chang, Mary L. Lindegren, Mary A. Butler, Stephen J. Chanock, Nicole F. Dowling, Margaret Gallagher, Ramal Moonesinghe, Cynthia A. Moore, Renée M. Ned, Mary Reichler, Christopher L. Sanders, Robert Welch, Ajay Yesupriya, and Muin J. Khoury for the CDC/NCI NHANES III Genomics Working Group

Background

Allele and genotype frequencies of human genetic variants are important for understanding the contribution of genetic variation to human disease susceptibility, progression, and outcomes. Population-based prevalence estimates provide the basis for epidemiologic studies of gene-disease associations, for estimating population attributable fraction, and for informing health policy and clinical and public health practice. However, such prevalence estimates for genotypes of public health importance have not yet been determined for the major racial and ethnic groups in the United States population.

Methods

We used DNA collected from 7,159 participants aged 12 and older enrolled in phase 2 (1991-1994) of the Third National Health and Nutrition Examination Survey (NHANES III), a weighted, minority-enriched population-based sample survey of the United States. Allele and genotype frequency estimates for 90 variants in 50 genes chosen for their potential public health significance were calculated by age, by sex, and for the three major race/ethnic groups (non-Hispanic white, non-Hispanic black, and Mexican-American).

Results

The nationally representative estimates and the 95% confidence intervals of allele frequencies and genotype prevalence were reported for 90 variants in 50 genes for each race/ethnicity, age, and sex subgroup in U.S. population. Allele and genotype frequencies differed significantly by race/ethnic group for ≥96.7% of the genetic variants, but did not differ significantly by sex or by age for most of the studied variants.

Conclusion

These nationally-representative allele and genotype frequency data provide a tremendous resource for future epidemiologic studies in public health in the United States.