Translating Advances from Proportion of Additive Genetic Variance

Roslyn Haar*

Department of Biology, Grenoble University Hospital Center, Grenoble, France

***Corresponding author:** Email: haar_r@gmail.com

Citation: Haar R (2022) Translating advances from Proportion of Additive Genetic Variance. Electronic J Biol, 18(12):1-2

Received date: November 11, 2022, Manuscript No. IPEJBIO-22-15451; **Editor assigned date:** November 13, 2022, PreQC No. IPEJBIO-22-15451 (PQ); **Reviewed date:** November 24, 2022, QC No. IPEJBIO-22-15451; **Revised date:** December 04, 2022, Manuscript No. IPEJBIO-22-15451 (R); **Published date:** December 11, 2022, DOI: 10.36648/1860-3122.18.12.059

Description

The remarkable range of discoveries that Genome-Wide Association Studies (GWASs) have facilitated in population and complex-trait genetics, the biology of diseases, and translation toward new therapeutics. In the introductory sections, we provide a background for this review, summarize its scope and layout, and revisit the scientific rationale for GWASs. We then review general conclusions that can be drawn from GWAS discoveries across a wide range of traits. We subsequently highlight more specific results of discoveries and methods on the path from GWAS to biology and review progress in three exemplar diseases, namely type 2 diabetes, auto-immune diseases, and schizophrenia. We end the review with a number of sections on the limitations of current experimental designs and possible ways to overcome these and a prediction on the future of GWASs for human traits.

Pleiotropy Pervasive

The number of segregating variants in the human population is large but finite. It is not known what proportion of the segregating variants are associated with complex-trait genetic variation, but the fact that each of the many studied traits is associated with variants at hundreds to thousands of loci in the genome strongly suggests that some of the underlying causal variants are the same. Multiple lines of evidence are consistent with widespread pleiotropy for complex traits. First, Mendelian mutations that cause specific syndromes or diseases are frequently associated with multiple phenotypes in an affected individual. Second, pedigree studies have reported genetic correlations between traits, implying that a number of the same variants affect two or more traits in a consistent direction.

Data generated from genome-wide SNP surveys have been exploited for addressing many scientific questions other than SNP-trait associations. We do not have the space to give adequate coverage of discoveries in evolutionary and population genetics, nor can we fully cover the many developments in analytic methods, although we will briefly mention some recent developments. The scope of our review is novel discoveries on the genetics and resulting biology of common adult diseases and their risk factors and the wider implications of those discoveries. GWAS discoveries have and are affecting a wide variety of diseases and traits, many of which have been covered in other in-depth reviews. Our focus is on associations between complex traits and SNPs, but we note that there have been many reported associations between traits and Copy-Number Variants (CNVs) and that there are known mechanisms by which CNVs can be associated with disease. Results from other genome-wide surveys, including exome and Whole-Genome Sequencing (WGS) studies, are not reviewed here.

The GWAS is an experimental design used to detect associations between genetic variants and traits in samples from populations. The primary goal of these studies is to better understand the biology of disease, under the assumption that a better understanding will lead to prevention or better treatment. The path from GWAS to biology is not straightforward because an association between a genetic variant at a genomic locus and a trait is not directly informative with respect to the target gene or the mechanism whereby the variant is associated with phenotypic differences. However, as reviewed herein, new types of data, new molecular technologies, and new analytical methods have provided opportunities to bridge the knowledge gap from sequence to consequence. GWASs have also been successfully implemented for better defining the relative role of genes and the environment in disease risk, assisting in risk prediction, and investigating natural selection and population differences.

Proportion of Additive Genetic Variance

GWASs to date rely on and exploit Linkage Disequilibrium (LD), the correlation structure that exists among DNA variants in the current human genome as a result of historical evolutionary forces, particularly finite population size, mutation, recombination rate, and natural selection. The statistical power to detect associations between DNA variants and a trait depends on the experimental sample size, the distribution of effect sizes of causal genetic variants that are segregating in the population, the frequency of those variants, and the LD between observed genotyped DNA variants and the unknown causal variants.

Therefore, the potential of a GWAS to succeed for a particular trait or disease depends on how many loci affecting the trait segregate in the population, the joint distribution of effect size and allele frequency at those loci, the experimental sample size, the panel of genome-wide variants that are used in the GWAS, and how heterogeneous the trait or disease being studied is. The last relates to both the biology of the trait and the ability to diagnose or measure it with precision. LD between genetic variants is commonly measured as a squared correlation (r2) because this measure is linear in the sample size required for detecting association between an observed genotyped and an unobserved causal variant.

LD r2 can be large only if the allele frequencies at the two loci match and this is the reason why GWASs from common SNP arrays are not powerful enough to detect associations due to rare causal variants (in addition to sample-size considerations; see below). Statistical imputation of unobserved variants can recover some of the information lost because of imperfect LD between observed genotypes and unobserved causal variants. Imputation is enabled by the fact that the genotypes of unobserved genetic variants can be predicted by the haplotypes inferred from multiple observed SNPs and the haplotypes observed from a fully sequenced reference panel.