Genome-wide association scan for height in 6,671 individuals from Finland and Sardinia

Serena Sanna[1,2], A.U. Jackson[1], G. Usala[2], C.J. Willer[1], M. Dei[2], L.L. Bonnycastle[3], S. Lai[2], Y. Li[1], M. Uda[2], R. Nagaraja[4], N. Timpson[5,6], G. Lettre[7], M.R. Erdos[3],, H. Shen[8], A. Shuldiner[8], A. Cao[2], R.M. Bergman[9], D. Schlessinger[2,4], F.S. Collins[3], J. N. Hirschhorn[7], G. D. Smith[5,6], E. Boerwinkle[10], M. Boehnke[1], G.R. Abecasis[1], K.L. Mohlke[11]

1) Dept of Biostatistics, University of Michigan, Ann Arbor, MI; 2) Istituto di Neurogenetica e Neurofarmacologia, CNR, Cagliari, Italy; 3) National Human Genome Research Institute, Bethesda, MD; 4) Gerontology Research Center, NIA, Baltimore, MD; 5) Dept of Social Medicine, University of Bristol, Bristol, UK; 6) 8 MRC CAiTE Centre, University of Bristol, Bristol, UK; 7) Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA; 8) University of Maryland, School of Medicine, Baltimore; MD 9) Keck School of Medicine of USC, Los Angeles, CA; 10) Human Genetics Center and Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX; 11) Dept of Genetics, University of North Carolina, Chapel Hill, NC.

sannas@umich.edu

Introduction

We have imputed ~2 million SNPs to compare GWA scans from two studies that used different genotyping platforms.

Data Description - Sample

•2,371 Finns from the Finland-United States Investigation of NIDDM Genetics (FUSION) study
•4,305 Sardinians from the ProgeNIA project, a longitudinal study of aging-related quantitative traits in the Ogliastra region of Sardinia

Data Description - Genotyping

•FUSION: 308,150 SNPs successfully genotyped using Illumina HumanHap300 BeadChip
•ProgeNIA: 356,359 SNPs successfully genotyped using Affymetrix 500K Array Set

Methods

•We used MACH[1] program to infer ~2 million SNPs in each study. CEU haplotypes were used as a reference pool to probabilistically infer missing genotypes based on haplotype similarity at common markers.
•We used a family based association test[2] to account for family structure while testing the additive effect of each marker.
•We combined single GWA scans using the weighted Z-score method for meta-analysis[3]

Figure 1. Combined GWA scan using 46,836 SNPs present on both the Affymetrix and Illumina platforms

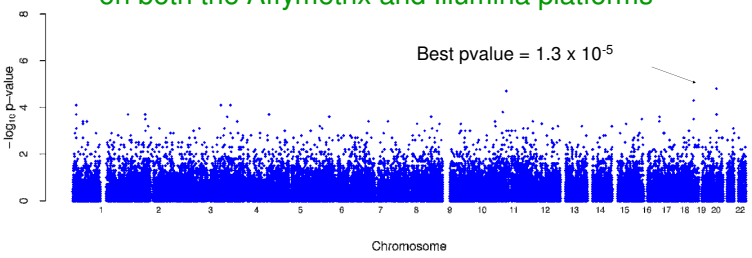

Figure 2. Combined GWA scan using ~2.28 million SNPs (imputed + genotyped)

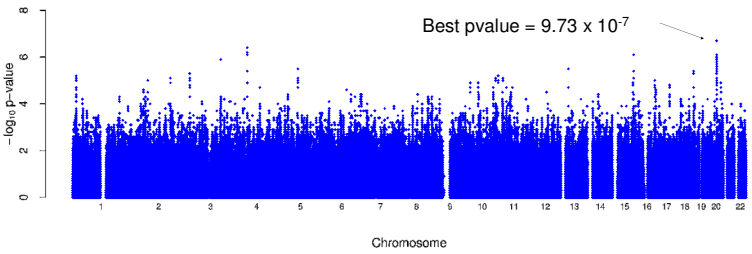

Figure 3. Schematic representation of the methodology

Imputation ⇒ *increased coverage (cost-free) allows comparison across platforms*

\+

Meta-analysis ⇒ *increased sample size (cost-free)*

=

Combined GWAS ⇒ *identification of new loci not covered by the original platforms or not significantly associated in small samples*

Results

The genome-wide scan revealed several loci putatively associated with height. We followed up our top SNP in other populations, succeeding in replication. The SNP falls in the GDF5-BFZB locus

Table 1. Replication data for the strongest associated SNP

Study Group	N	Allele freq	Effect (se) in cm	p-value
FUSION T2D stage 1	1,084	0.416	.343 (.250)	0.071
FUSION NGT stage 1	1,287	0.447	.613 (.235)	0.0024
ProgeNIA	4,305	0.384	.700 (.186)	4.35×10^{-4}
	6,676		*Stage 1 meta-analysis*	*9.73×10^{-7}*
FUSION T2D stage 2	1,075	0.454	.442 (.280)	0.154
FUSION NGT stage 2	1,028	0.453	.449 (.267)	0.022
DGI T2D	1,517	0.447	.449 (.234)	0.044
DGI controls	1,468	0.425	.323 (.249)	0.038
Old Order Amish	2,711	0.383	.424 (.178)	0.028
ARIC European Americans	10,882	0.398	.252 (.085)	0.002
ARIC African Americans	3,860	0.710	.254 (.160)	0.169
Caerphilly	1,097	0.370	.522 (.270)	0.055
BWHHS	3,652	0.362	.560 (.147)	9.71×10^{-5}
	27,295		*Stage 2 meta-analysis*	*2.84×10^{-11}*
	33,966		*Overall meta-analysis*	*4.44×10^{-16}*

T2D = Type 2 diabetes
NGT = Normal glucose tolerant
ARIC = Atherosclerosis Risk in Communities
DGI = Diabetes Genetics Initiative (Finland and Sweden)
Caerphilly = European Men from UK
BWHHS = British Women's Heart and Health Study

Conclusion

With the imputation of millions of SNPs and the meta-analysis we are able to identify a locus responsible for height variation. The locus was poorly associated in the single GWA scans because of the sparse SNP maps of genotyping platforms, and because its effect on height variation is too small to be detected in the original samples. This suggests that combining GWA scans among groups could be an important step in GWA studies.

References

[1]Li Y. et al, *In Silico Genotyping for Genome-Wide Association Studies*. Poster #2071/W
[2]Chen W. & Abecasis G., *Family Based Association test for GWAS*. AJHG(2007)
[3]Whitlock, *The weighted Z-method is superior to Fisher's approach*. J.EV. BIOL. (2005)