

RESEARCH HIGHLIGHT

Identifying population differences in genes that affect body mass index

Scott M Williams

Abstract

Genomic regions of interest can be narrowed by studying populations that have patterns of low linkage disequilibrium. A recent study of body mass index in African Americans demonstrated this point and, through cross-population analyses, revealed additional genomic associations. This comparative analysis showed how rare alleles that associate with traits in specific populations can be detected in cohorts where the same alleles are not rare, and highlights how population diversity can aid genetic analyses.

Genome-wide studies and the search for genes that affect complex disease

Genome-wide association studies (GWAS) have revolutionized the way that geneticists search for loci that affect human diseases and phenotypes [1]. The overall success of using high density chips to interrogate common variants for common disease is a matter of debate [2]; however, it is unarguable that many unexpected signals for disease have been identified and replicated in multiple studies, even if their biological implications still remain a mystery in most cases. GWAS signals typically only identify genomic regions of potential interest, because the single nucleotide polymorphisms (SNPs) on any given chip are generally not functional SNP(s), but proxies, chosen to cover common variation in one or more populations. The sizes of these regions of interest are dependent on many variables, including the strength of the effect of any functional variant(s) and the patterns of linkage disequilibrium (LD) in the population(s) studied. This latter point has led many to argue that fine mapping of functional variation will best be done in populations that have, on average, lower LD, such as those of African descent (Figure 1) [3,4]. This approach was recently used in a large study examining the genetics of body mass index (BMI) in African Americans [5].

Transferability of GWAS results for high BMI from European-descent to African-descent populations

Analyses of BMI across diverse populations - or transferability of association results - is of particular interest in African American populations because they have disproportionately high BMI and obesity, and related diseases such as diabetes. In addition, transferable loci provide a compelling basis to conclude that the sites are real and of universal biological and clinical significance, even in the absence of functional data. Demonstrating association with other variants in the same region may also serve either to narrow a functional site(s) or to provide evidence for allelic heterogeneity. Thus, clearly, searching for BMI-associated variants in African American populations can provide extremely useful information to guide expensive and time-consuming functional studies that may lead to either universal or population-specific prevention and treatment regimens. Gong *et al.* [5] provide an excellent example of how to do this.

Gong *et al.* examined transferability of genetic results using MetaboChip [6,7]. This platform, constructed from GWAS of primarily European-descent populations, comprises approximately 200,000 genetic markers that have effects on a number of metabolic diseases, including obesity and type 2 diabetes, or are promising in these and other studies. The authors first tested 21 genes known to associate with high BMI in people of European descent for a similar association in 30,000 African American individuals. They then performed exploratory meta-analyses (agnostic) on the remaining MetaboChip SNPs to identify additional, and novel, markers associated with BMI in African American individuals.

(Re)testing explicit hypotheses

Unlike most GWAS studies that emphasize discovery and novelty over hypothesis testing, Gong *et al.* tailored their analysis to explicitly look for overlap with previous findings. Their results have provided compelling evidence that eight (*SEC16B*, *TMEM18*, *ETV5*, *GNPDA2*, *TFAP2B*, *BDNF*, *FTO*, *MC4R*) of the 21 loci associated

Correspondence: scott.williams@dartmouth.edu
Department of Genetics, Geisel School of Medicine, Dartmouth College,
Hanover, NH 03755, USA

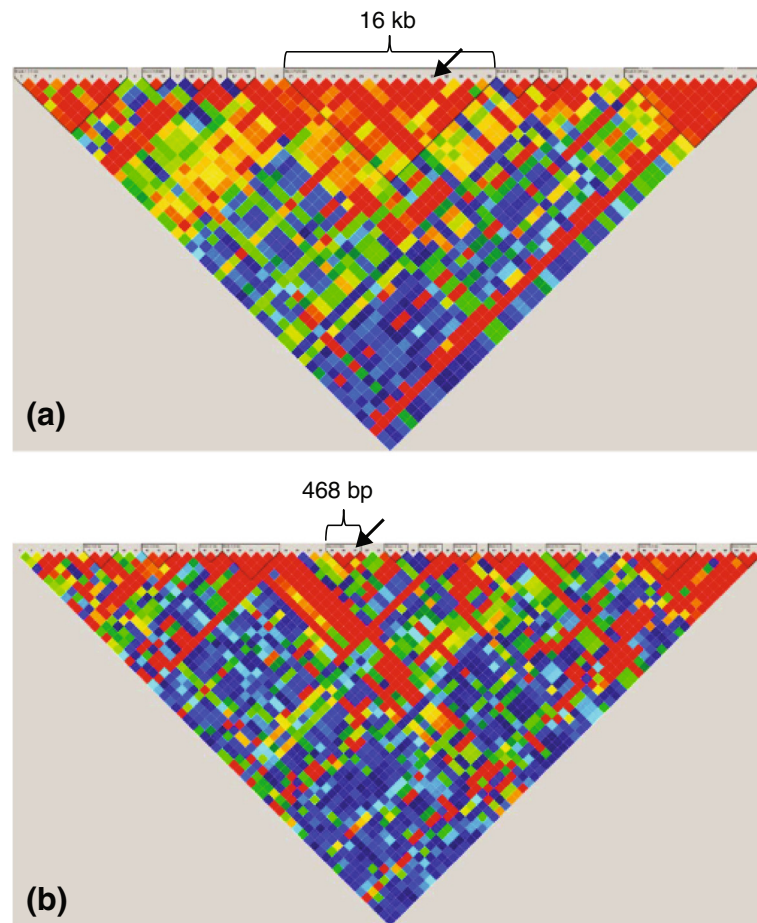


Figure 1 Differences of linkage disequilibrium structure. HapMap based LD around SNP rs629301, near SORT1 (denoted by arrow), in (a) Europeans (CEU) and (b) West Africans (YRI) that associated with low density lipoprotein in both Europeans and African Americans. For this SNP, the region of linkage disequilibrium around it narrowed in the African population to 468 bp from approximately 16 kb in the European population, as indicated by red and yellow diamonds, providing a 30-fold reduction in the associating region. bp, base pair; kb, kilobase; LD, linkage disequilibrium; SNP, single nucleotide polymorphism. (From [4]).

with high BMI in populations of European descent are replicated in African American populations. Most of these loci contained more than one significant SNP; of the eight genes replicated, three identified the same SNP found in European individuals as the most significant in African American individuals (rs543874 in *SEC16B*, rs7647305 in *ETV5* and rs10938397 in *GNPDA2*). These results strongly suggest that the same functional variants, or the same part of each gene, are associated with BMI in both populations. Additionally, as expected, the sizes in kilobases of the BMI-associated regions were smaller in African American populations, except for *TFAP2B*. The reduction of associating regions by physical size in the other genes ranged from approximately 6% (*FTO*) to approximately 94% (*SEC16B*). These results clearly demonstrate that fine-mapping of specific genomic regions is better in African-descent populations due to their lower levels of LD.

Lastly, an additional, independent SNP in *GNPDA2* (rs186117327) was associated with BMI in the African American population, revealing at least one example of allelic heterogeneity in this gene that was not detected in prior analyses. Therefore, analyses of populations with lower LD than European populations cannot only narrow the sizes of associating regions, but can help to differentiate between genetic signals within the same gene.

Agnosticism in the MetaboChip

Gong *et al.* also performed an agnostic study of the remaining MetaboChip SNPs. This identified two new genes significantly associated with BMI, *BRE* on chromosome 2 and *DHX34* on chromosome 19. Both met the threshold of 2.5×10^{-7} , the MetaboChip-wide significance threshold for approximately 200,000 tests. Two of the SNPs in *BRE* reached the normal GWAS threshold. Neither of these genes seemed to be affected by

ancestry adjustment, but were clearly not associated in European population studies. For *BRE*, the failure to find association in European-descent populations was not surprising, as the associated SNPs in this gene all had minor allele frequencies <0.001 in European populations but >0.1 in African American populations. Clearly, even if the same alleles associate in Europeans functionally, they would never be detected statistically. Perhaps this serves as yet another explicit warning about limitations of statistical analyses of rare variants in association studies, especially in GWAS studies where rare variant associations cannot be readily detected. For *DHX34*, the *P*-value was not as small as for *BRE*, not reaching conventional GWAS threshold, and the authors argued for replication in this case. Another common SNP on chromosome 19 (rs3810291) was associated with BMI in a recent study of European-descent populations, but it was 300 kb from the *DHX34* signal and in low LD even in Europeans, indicating that these are independent signals [8].

Neither *BRE* nor *DHX34* appeared in the 1,500 most significantly associating genes in a recent and even larger meta-analysis of BMI in African-descent populations, published earlier this year [9]. This alone does not indicate that the results of Gong *et al.* are false positives, as real signals can certainly be below this arbitrary cutoff of the top 1,500 [10], but does raise some doubts about these findings that larger samples may or may not address. Making all *P*-values <0.05 from GWASs publicly and readily available would be useful to assess the likelihood of association by consensus, if not by single arbitrary thresholds.

Utility of cross-population analyses

The study described above [5] provides an excellent example of how genetic studies in low-LD, African-descent populations can inform the genetics of disease risk. Gong *et al.* discovered virtually all of the things that might have been expected using this approach. They showed that several - far more than expected by chance - gene regions associated with high BMI in both European- and African-descent analyses, providing very strong evidence that these loci are universal genetic risk factors. They used the low LD analyses to narrow most regions, which will make subsequent functional studies a little easier, and found at least one novel region within *GNPDA2* to be associated with BMI risk in African American individuals. They also showed how assessing differences between diverse populations with different risk allele frequencies can reveal new signals. This latter point extends beyond the issue of LD differences to a more generic one: rare allelic associations may be found by carefully evaluating populations with very different patterns of allele frequencies, that is, where the putative

risk alleles are not rare. But, at its most basic, this study shows what is real with respect to BMI risk, and how complex disease genetic analyses can proceed more efficiently when diverse study populations are examined.

Abbreviations

BMI: Body mass index; GWAS: Genome-wide association study; kb: Kilobase; LD: Linkage disequilibrium; SNP: Single nucleotide polymorphism.

Competing interests

The author declares that he has no competing interests.

Published: 29 November 2013

References

1. Visscher PM, Brown MA, McCarthy MI, Yang J: **Five years of GWAS discovery.** *Am J Hum Genet* 2012, **90**:7–24.
2. McClellan J, King MC: **Genetic heterogeneity in human disease.** *Cell* 2010, **141**:210–217.
3. Tishkoff SA, Williams SM: **Genetic analysis of African populations: human evolution and complex disease.** *Nat Rev Genet* 2002, **3**:611–621.
4. Adeyemo A, Bentley AR, Meilleur KG, Doumatey AP, Chen G, Zhou J, Shriner D, Huang H, Herbert A, Gerry NP, Christman MF, Rotimi CN: **Transferability and fine mapping of genome-wide associated loci for lipids in African Americans.** *BMC Med Genet* 2012, **13**:88.
5. Gong J, Schumacher F, Lim U, Hindorf LA, Haessler J, Buyske S, Carlson CS, Rosse S, Buzkova P, Fornage M, Gross M, Pankratz N, Pankow JS, Schreiner PJ, Cooper R, Ehret G, Gu CC, Houston D, Irvin MR, Jackson R, Kuller L, Henderson B, Cheng I, Wilkens L, Leppert M, Lewis CE, Li R, Nguyen KD, Goodloe R, Farber-Eger E, *et al*: **Fine mapping and identification of BMI loci in African Americans.** *Am J Hum Genet* 2013, **93**:661–671.
6. Buyske S, Wu Y, Carty CL, Cheng I, Assimes TL, Dumitrescu L, Hindorf LA, Mitchell S, Ambite JL, Boerwinkle E, Buzkova P, Carlson CS, Cochran B, Duggan D, Eaton CB, Fesinmeyer MD, Franceschini N, Haessler J, Jenny N, Kang HM, Kooperberg C, Lin Y, Le Marchand L, Matise TC, Robinson JG, Rodriguez C, Schumacher FR, Voight BF, Young A, Manolio TA, *et al*: **Evaluation of the metabochip genotyping array in African Americans and implications for fine mapping of GWAS-identified loci: the PAGE study.** *PLoS One* 2012, **7**:e35651.
7. Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, Burtt NP, Fuchsberger C, Li Y, Erdmann J, Frayling TM, Heid IM, Jackson AU, Johnson T, Kilpeläinen TO, Lindgren CM, Morris AP, Prokopenko I, Randall JC, Saxena R, Soranzo N, Speliotes EK, Teslovich TM, Wheeler E, Maguire J, Parkin M, Potter S, Rayner NW, Robertson N, Stirrups K, *et al*: **The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits.** *PLoS Genet* 2012, **8**:e1002793.
8. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Magi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segre AV, Estrada K, Liang L, Nemes J, Park JH, Gustafsson S, Kilpeläinen TO, *et al*: **Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.** *Nat Genet* 2010, **42**:937–948.
9. Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, Lange LA, Ng MC, Adeyemo AA, Allison MA, Bielak LF, Chen G, Graff M, Irvin MR, Rhie SK, Li G, Liu Y, Liu Y, Lu Y, Nalls MA, Sun YV, Wojczynski MK, Yanek LR, Aldrich MC, Ademola A, Amos CI, Bandera EV, Bock CH, Britton A, Broeckel U, Cai Q, *et al*: **A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry.** *Nat Genet* 2013, **45**:690–696.
10. Williams SM, Haines JL: **Correcting away the hidden heritability.** *Ann Hum Genet* 2011, **75**:348–350.

doi:10.1186/gm506

Cite this article as: Williams: Identifying population differences in genes that affect body mass index. *Genome Medicine* 2013 **5**:102.