Published in final edited form as:

Hum Genet. 2011 November ; 130(5): 607–621. doi:10.1007/s00439-011-0974-0.

The efficacy of detecting variants with small effects on the Affymetrix 6.0 platform using pooled DNA

Charleston W. K. Chiang,

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA

Zofia K. Z. Gajdos,

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA

Joshua M. Korn,

Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA. Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA

Johannah L. Butler,

Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA

Rachel Hackett,

Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA

Candace Guiducci,

Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA

Thutrang T. Nguyen,

Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA

Rainford Wilks,

Epidemiology Research Unit, Tropical Medicine Research Institute, University of the West Indies, Kingston, Jamaica

Terrence Forrester,

Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Kingston, Jamaica

Katherine D. Henderson,

Division of Cancer Etiology, Department of Population Sciences, City of Hope National Medical Center, Duarte, CA, USA

Loic Le Marchand,

© Springer-Verlag 2011

Correspondence to: Joshua M. Korn, joelh@broadinstitute.org.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-011-0974-0) contains supplementary material, which is available to authorized users.

Epidemiology Program, Cancer Research Center of Hawaii, University of Hawaii, Honolulu, HI, USA

Brian E. Henderson,

Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

Christopher A. Haiman,

Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

Richard S. Cooper,

Department of Preventive Medicine and Epidemiology, Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA

Helen N. Lyon,

Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA

Xiaofeng Zhu,

Department of Biostatistics and Epidemiology, Case Western Reserve University, Cleveland, OH, USA

Colin A. McKenzie,

Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Kingston, Jamaica

Mark R. Palmer, and

Division of Endocrinology, The Hospital for Sick Children, Toronto, ON, Canada. Department of Pediatrics, University of Toronto, Toronto, ON, Canada

Joel N. Hirschhorn

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA, USA. Children's Hospital Boston, CLS 16065, 300 Longwood Avenue, Boston, MA 02115, USA

Joshua M. Korn: joelh@broadinstitute.org

Abstract

Genome-wide genotyping of a cohort using pools rather than individual samples has long been proposed as a cost-saving alternative for performing genome-wide association (GWA) studies. However, successful disease gene mapping using pooled genotyping has thus far been limited to detecting common variants with large effect sizes, which tend not to exist for many complex common diseases or traits. Therefore, for DNA pooling to be a viable strategy for conducting GWA studies, it is important to determine whether commonly used genome-wide SNP array platforms such as the Affymetrix 6.0 array can reliably detect common variants of small effect sizes using pooled DNA. Taking obesity and age at menarche as examples of human complex traits, we assessed the feasibility of genome-wide genotyping of pooled DNA as a single-stage design for phenotype association. By individually genotyping the top associations identified by pooling, we obtained a 14- to 16-fold enrichment of SNPs nominally associated with the phenotype, but we likely missed the top true associations. In addition, we assessed whether genotyping pooled DNA can serve as an inexpensive screen as the second stage of a multi-stage design with a large number of samples by comparing the most cost-effective 3-stage designs with 80% power to detect common variants with genotypic relative risk of 1.1, with and without pooling. Given the current state of the specific technology we employed and the associated

genotyping costs, we showed through simulation that a design involving pooling would be 1.07 times more expensive than a design without pooling. Thus, while a significant amount of information exists within the data from pooled DNA, our analysis does not support genotyping pooled DNA as a means to efficiently identify common variants contributing small effects to phenotypes of interest. While our conclusions were based on the specific technology and study design we employed, the approach presented here will be useful for evaluating the utility of other or future genome-wide genotyping platforms in pooled DNA studies.

Introduction

With the advent of arrays designed to genotype 500,000 or more SNPs per array, genome-wide association (GWA) studies have identified many SNPs reproducibly associated with various complex traits such as diabetes, height, and obesity (Hindorff et al. 2009; McCarthy et al. 2008). However, most reproducible associations identified through GWA studies represent SNPs with modest effect sizes (odds ratios typically ranging from 1.1 to 1.3 (Hindorff et al. 2009; Iles 2008)), and thus a study design comprising tens of thousands of individuals is required to have sufficient power to detect such associations. Therefore, although genotyping technology has now advanced to the point where it is possible to genotype up to one million SNPs on a single array, the cost of a well-powered GWA study is often prohibitively expensive given the current cost of roughly \$500–700 per array. In situations with limited funding (e.g., when investigating an understudied disease), one possible way to reduce the cost of initial association screening is to genotype pools of DNA from multiple individuals rather than to genotype each individual separately.

Genome-wide genotyping of pooled DNA has previously been used to resolve individuals contributing trace amounts of DNA to a pool (Homer et al. 2008a) and to estimate global ancestry and identify ancestry informative markers (Chiang et al. 2010). For the purpose of conducting GWA studies, numerous analytical tools, approaches, and statistical considerations have been proposed (Craig et al. 2009; Docherty et al. 2007; Homer et al. 2008b; Macgregor et al. 2008; Meaburn et al. 2006; Pearson et al. 2007; Visscher and Le Hellard 2003; Yang et al. 2008; Zhang et al. 2008). However, using pooled genotypes to identify common variants associated with disease, the original goal of genotyping pooled DNA, has had mixed success. Known variants associated with Alzheimer's disease, progressive supranuclear palsy, eye color, pseudoexfoliation syndrome, and age-related macular degeneration have been detected by pooling only several hundred samples in proof-of-principle experiments (Craig et al. 2009; Pearson et al. 2007). Novel variants as well as variant sets that are likely enriched for potential true associations have also been identified using pooled DNA for melanoma susceptibility (Brown et al. 2008), memory performance (Papassotiropoulos et al. 2006), schizophrenia (Shifman et al. 2008b), otosclerosis (Schrauwen et al. 2009), mild mental impairment (Butcher et al. 2005), general cognitive ability (Butcher et al. 2008), mathematical ability (Docherty et al. 2010), and reading ability (Meaburn et al. 2008). Despite these success stories, the common theme among them is that the identified variants all have large effect sizes, with odds ratios ranging from ~1.7 to >15. On the other hand, variants with small effect sizes may not be amenable to identification by pooled genotyping, since the pooling-specific error could have a greater impact (see Pearson et al. 2007 for discussion). Because of these concerns, genotyping pooled DNA has been proposed as a quick and inexpensive screen of understudied diseases for common variants with large effect sizes (Chiang et al. 2010). However, for DNA pooling to serve as a viable cost-saving strategy for conducting GWA studies, the efficacy of genotyping pooled DNA to identify variants of small effect sizes needs to be evaluated.

We performed our evaluation in two ways. First, we evaluated the efficacy of genotyping pooled DNA to identify common variants of small effect in the setting of a single-stage

association study design; second, we assessed whether genotyping pooled DNA can be used as an inexpensive screening stage of a multi-stage design with a large number of samples.

To explore the utility in a single-stage design, we used obesity (as measured by body mass index, BMI) and age at menarche as examples of complex traits for which no common variants with large effect size have been reported (Gajdos et al. 2010; Kang et al. 2010; Ong et al. 2009; Speliotes et al. 2010). We individually genotyped the top associated SNPs identified by pooling to assess the enrichment of validated associations using a pooled design. We note that our focus is on evaluating whether we could identify SNPs with real between-group allele frequency (AF) differences among the actual individuals pooled in our study, as a means to examine the efficacy of the methodology. If the analysis of pooled GWA data is technically robust to small differences in AF, the top associated SNPs from the pooled analysis should show AF differences when genotyped in the same individual samples. Our results should not be interpreted as generalizable findings of associations, as apparent differences in AF between the case and control pools of the size described here are much more likely to reflect sampling variation (and potentially uncorrected error due to pooling as well) than reproducible associations with the phenotype of interest.

To determine if the current state of the specific genotyping technology we employed would generate pooled genotypes useful for a multi-stage GWA design searching for variants with small effects, we compared the most cost-efficient 3-stage designs assuming 50,000 individuals, with or without pooling a subset of the individuals in the second stage. Here we relied on a cohort of African-Americans for whom we have both individual and pooled genotype data to model the efficiency of pooling (Chiang et al. 2010).

We found that we could enrich for nominal associations in a single-stage design by focusing on the top SNPs associated with obesity and age at menarche in a pooled GWA study, although, as expected with an imprecise methodology, the evidence of association was over-estimated in the pooled data and the true top associations were likely missed. In multi-stage designs, the cost of the most optimal 3-stage design involving pooled samples appears to be similar to, but not cheaper than, that of a 3-stage design not involving pooled samples. Together, these results suggest that pooled genotyping, as used here, may not be more cost-efficient for identifying variants with small effect sizes, although future changes in genotyping costs or further improvements in genotyping technology may make pooled genotyping a useful component of future GWA study designs.

Results

Identification and validation of SNPs associated with obesity

We constructed four case pools ($N = 101, 136, 137$, and 137) and two control pools ($N = 79$ and 98) from 688 individuals recruited as part of a study of gene by environment effects from Kingston, Jamaica (GXE), as well as two case pools ($N = 129$ and 135) and two control pools ($N = 105$ and 111) from 480 individuals from Spanish Town, Jamaica (SPT). Pools were constructed according to high (case) and low (control) BMI, stratified by gender (see “Methods”). Each pool was genotyped in triplicate using the Affymetrix 6.0 SNP array. We estimated pooled AF and applied four SNP quality control (QC) filters, as previously described (Chiang et al. 2010). The QC filters were effective in removing the vast majority of SNPs with apparent very large AF differences that are likely to be false positives in an association study (Fig. 1a, b). Moreover, because we were looking for SNPs with very small AF differences, we corrected a systematic bias in AF estimation of SNPs with relatively low AF in the case or control pools by fitting the AF distribution of each pool within a cohort to the overall distribution of AF (see “Methods”). This procedure allowed the distribution of

AF in the case and control pools to be much more comparable and symmetric around the null expectation line of equal AF in cases and controls (Fig. 1c).

The Jamaican pools (GxE and SPT) were separately analyzed to identify variants associated with BMI (see “Methods”). SNPs that passed the QC filters in both cohorts were then meta-analyzed, weighted by the sample size of each cohort. In total, we assayed 252,647 SNPs and observed no obvious signals above the null (Fig. 2a). We divided the top 1,000 SNPs after meta-analysis into three categories, based on the association ranks assigned to the surrounding SNPs in LD with the SNP of interest: “encouraging” associations have other SNPs in LD showing good evidence of association; “discouraging” associations have other SNPs in LD showing poor evidence of association; and “inconclusive” associations have mixed signals or have no other SNPs in LD (see “Methods”). We successfully genotyped a total of 21 SNPs out of the top 36 “encouraging” (ENC) and the top 7 “inconclusive” SNPs (limiting the inconclusive SNPs to those with no SNPs in LD, INC-LD) in the individuals that comprised the GxE and SPT pools, and analyzed them using the same case/control study design. Fourteen of the 21 SNPs showed evidence of association with BMI with $P < 0.05$ in meta-analysis, with an additional two SNPs with $P < 0.05$ in one population that were not genotyped successfully in the other (Table 1). The 14-fold enrichment of SNPs with $P < 0.05$ (16, when 1.15 were expected if SNPs were randomly selected) demonstrates that a GWA study using pooled samples may be an effective method to screen for potential associations in a staged design.

Identification and validation of SNPs associated with age at menarche

For identifying SNPs associated with age at menarche, we constructed five pools each of women with early and late age at menarche from the Multi-Ethnic Cohort (MEC) of Los Angeles and Hawai’i (see “Methods”), for which we have some past genotype data on individual participants useful for validation of the pooled analysis. In total, 745 women with early menarche ($N = 168, 132, 91, 169, \text{ and } 185$) and 729 women with late menarche ($N = 153, 120, 111, 163, \text{ and } 182$) were pooled, stratified by racial/ethnic cohort. The MEC pools were analyzed similarly to the Jamaican pools. Association analyses were conducted within cohort before meta-analysis, and only SNPs passing QC filters in at least three of the five cohorts were meta-analyzed. To account for different AF across diverse panels during meta-analysis, each cohort was weighted by its effective population size, which was calculated based on pooled AF estimates (Gajdos et al. 2008). In total, we assayed 390,175 SNPs and observed no obvious deviation from the null expectation (Fig. 2b). Among the top 2,000 associations, 25 SNPs had been previously individually genotyped in our laboratory (Table 2). Among these 25 SNPs, 15 would have been categorized as ENC or INC-LD (see “Methods”), and 12 out of the 15 SNPs achieved a $P < 0.05$. This represents a 16-fold enrichment with $P < 0.05$ (12, when only 0.75 was expected if SNPs were randomly selected), comparable to that observed for the analysis of BMI above. Of the remaining 10 SNPs in the “inconclusive” (INC, because SNPs in LD showed marginal evidence of association in pooling) and “discouraging” (DIS) categories, half of the SNPs in each category reached a nominal $P < 0.05$ (Table 2). Though limited in number, this suggests that the evidence of association of neighboring SNPs in LD can provide additional useful information for discerning real associations from false positives.

Assessment of the utility of genotyping pooled DNA as a screen in a multi-stage association design

While we were able to enrich for nominal associations by analyzing pooled DNA samples, we tended to over-estimate the evidence of association by at least an order of magnitude (Tables 1, 2), and likely missed a number of the SNPs most differentiated between case and control pools due to both errors associated with pooling and loss of SNP coverage due to

QC. For example, given that ~252,000 SNPs were assayed for BMI, the best P -value should be on the order of $\sim 4 \times 10^{-6}$ under the null hypothesis, and ~9 to 10 SNPs should have had P values more significant than our best associated signal (rs4793952, $P = 4.1 \times 10^{-5}$, Table 1), but were not successfully detected among the top SNPs we individually validated. This implies that the pooling approach may not be adequate enough to serve as a 1-stage design or a first-stage screen. As the most efficient design for GWA studies is typically a multi-stage design (Skol et al. 2006), the power and potential cost reductions of various 2-stage designs involving pooled genotype data have also been investigated (Macgregor et al. 2008; Zuo et al. 2008, 2006). Similarly, we assessed the utility of genotyping pooled DNA to act as a screening stage in a multi-stage design, so to enable fewer putatively associated variants to be genotyped in subsequent stages without substantial loss of power. We conducted our assessment in two ways: first, by empirically determining if information from pooled cohorts would allow better prioritization on GWA signals for follow-up; second, by simulation to compare the most optimal 3-stage study designs with or without pooling.

Specifically, we first considered a scenario in which a GWA study using individual genotyping had already been done with a discovery cohort, and the researchers have additional samples available for replication but with only limited funding for genotyping. In this case, we asked if having pooled information would allow one to prioritize which SNPs to follow up in replication (Fig. 3a). To test this possibility empirically, we used data from a previously conducted GWA study of BMI in African-derived populations (Kang et al. 2010). In this study, all independent SNPs with $P < 1 \times 10^{-5}$ in the discovery stage were screened in additional panels for consistency in the direction of effect (screening stage). In the last stage (replication stage), seven SNPs satisfied the conditions of the first two stages and were genotyped in the same Jamaican populations that had been pooled here (Kang et al. 2010). To determine if the availability of pooling data would assist prioritization of SNPs to be followed up, we also individually genotyped 11 additional SNPs that were associated with BMI with $1 \times 10^{-5} < P < 0.001$ (in the discovery phase, i.e., in the African-derived cohorts) that had also achieved a one-tailed quantile-normalized $P < 0.1$ in the pooled analysis (Fig. 3a). Among these 11 SNPs, 5 reached a one-tailed $P < 0.05$, and one additional SNP achieved a one-tailed $P < 0.07$ (Table 3). In comparison, none of the 7 top SNPs identified in the original GWA study achieved one-tailed P values < 0.05 in GxE and SPT (Table 3), despite having stronger evidence of association in the discovery stage than the 11 SNPs identified with added pooling information and despite having passed the screening stage of that study design. This suggests that while the substantial false-negative rate due to stringent filtering of the pooled data may make pooled genotyping less desirable as a first-stage design, it may be useful for supplementing the list of putative associations for follow-up in multi-staged design, especially by providing additional association evidence (in pooling) for SNPs that only showed modest association in the discovery phase.

To further determine if incorporating information from pooled DNA would produce more cost-efficient designs of multi-staged association studies, we used simulations. We assumed a scenario in which a case-control cohort of 50,000 individuals is available for a 3-stage analysis designed to have 80% power to detect common variants ($AF = 0.3$) with a genotypic relative risk (GRR) of 1.1, a modest effect size consistent with the effect sizes seen among variants associated with common complex diseases (Hindorff et al. 2009; Iles 2008). If no pooling were involved, we assumed that three subsets of the 50,000 individuals would each be genotyped in different stages, first by a genome-wide SNP array and the latter two by individual custom genotyping. The most cost-effective design under this scheme is then compared to a design where all samples used in stage 2 were genotyped genome-wide as pooled DNA, while maintaining the overall 80% power and penalizing the increased rate of false positives that would arise with pooling (see “Methods”). Because of the introduction of pooling, the optimal number of individuals to genotype in each stage is

different under the two schemes (N_{1A} , N_{2A} , and N_{3A} with all-individual genotyping; N_{1B} , N_{2B} , N_{3B} when subset of individuals were pooled (see “Methods”, Fig. 3b)). Note that our simulation only aims to model the cost of the genotyping itself, and thus does not take into account any alterations in the analytical cost due to pooled data or any potential reduction in speed associated with pooled genotyping.

Without pooling, the most efficient design involves genome-wide genotyping of 15,000 individuals in stage 1, followed by genotyping the top 0.1% of the SNPs in 8,000 additional individuals in stage 2. Those that achieved a one-tailed $P < 0.1$ would then be genotyped in the remaining 27,000 individuals in stage 3. This study design has 80% power to detect a common variant ($AF = 0.3$) with $GRR = 1.1$ that passes through the first two stages and achieves a one-tailed $P = 0.0025$ in stage 3, when under the null only 0.2 SNPs would be expected (Table 4). If a pooling design replaced the stage 2 screen and pooled samples were genotyped by a genome-wide platform, the optimal design would involve genotyping 14,000 individuals in stage 1, and 28,000 individuals in pools of 150 in stage 2. SNPs in the top 0.5% of stage 1 that also achieved a quantile-normalized one-tailed $P < 0.35$ in stage 2 would then be genotyped in the remaining 8,000 individuals in stage 3. A true association would then have 80% chance to reach a one-tailed $P < 0.1$ in stage 3. Under the null, 175 such SNPs would exist, and we penalize this increased false positive rate by requiring all 175 SNPs to be individually genotyped in the 28,000 individuals that were assayed in the pools. With a cohort of 28,000 individuals, one would have $>99.8\%$ power to detect a SNP with a one-tailed $P = 0.003$ (the top association among 175 SNPs under the null). Under this optimal design involving pooling, the cost of the entire study would be 1.07 times more expensive than without pooling (Table 4), suggesting that pooled genotyping may not be cost-saving using the platform and methods described here.

The cost-benefit simulation described here is based on current genotyping costs, as well as the efficiency of pooling using the Affymetrix 6.0 platform. We re-simulated the most cost-efficient design with and without pooling with varying conditions of array-to-individual genotyping cost ratios and pooling efficiencies (see “Methods”) to determine if improvements in technology or methodology will make a study design involving pooling more favorable. We found that in general a design with pooling would be more favorable if the cost of the genome-wide SNP array were to decrease, or if the efficiency of pooling were to improve. Conversely, the all-individual genotyping scheme would be favored if individual multiplex custom genotyping cost were to decrease (Fig. 4). Specifically, if the array/individual genotyping cost ratio were to become 1/4 or 1/2 as expensive, or perhaps more importantly if the pooling efficiency were to improve such that the same amount of truly associated SNPs will attain pooling P values that are 1/2 or 1/3 of the current pooling P values, the cost of the entire study could be 0.91 times that of the design without pooling, assuming a $GRR = 1.1$, a variant AF of 0.3, and target power of 80% (Fig. 4).

Discussion

In order for genome-wide genotyping of pooled samples to be a viable cost-saving alternative to individual-level GWA studies, which was the original goal of pooled genotyping, it is necessary to be able to identify associations with small effect sizes, especially given that many common variants irrefutably associated with complex traits have small effect sizes (Hindorff et al. 2009; Iles 2008). Failures of DNA pooling studies to identify variants reproducibly associated with a number of traits, such as neuroticism (Shifman et al. 2008a) and general cognitive ability (Davis et al. 2010), are consistent with the notion that pooled genotyping may be ineffective for identifying variants with small effect sizes, especially since thousands of individuals were pooled in these studies (although other factors such as study design or inadequate power may also be the cause). However, a

systematic assessment of the utility of pooling for identifying such variants had not been conducted previously.

In the present study, we empirically evaluated the use of pooled genotype data obtained using the Affymetrix 6.0 platform to detect associations with obesity and age at menarche. We were able to obtain a 14- to 16-fold enrichment of SNPs nominally associated with our phenotypes (Tables 1, 2), even though the substantial false-negative rate likely prevents our strategy from being used in a single-stage design. Moreover, despite the fact that a great deal of information exists within pooled genotype data, a pooled GWA study design as implemented here does not appear to represent a savings in study cost as compared to a GWA study design employing solely individual genotyping (Table 3).

In order to reliably detect associated variants with small effect sizes, we stringently applied QC filters and also re-parameterized the AF distribution of our pooled data. Concurrent with stringent control of false positive rate is the loss of power due to loss of coverage. Our previous work has shown that the false-negative rate due only to the QC filters applied here is approximately 51% (Chiang et al. 2010). As a result, we were only able to assay ~300,000 to ~500,000 SNPs in each of the pools, and as few as ~250,000 SNPs in meta-analysis for phenotype association. Additionally, while we were able to enrich for nominal associations among the top associated signals by pooling, we tended to over-estimate the effect sizes and probably missed the most differentiated SNPs, likely reflecting pooling-specific errors that remain uncorrected or not modeled in our study design. Taken together, even though it is evident that a great deal of useful information exists in pooled genotype data, it may be difficult, under the conditions employed here, to use pooled genotyping for either a single stage or for the first stage of a multi-stage GWA study design. Instead, it may be useful to consider this information for obtaining additional variants to supplement the list of top putative associations (typically identified in the discovery phase of a GWA study) for follow-up genotyping (Table 3), particularly if the pooled data are readily available.

We also investigated the efficacy of pooled genotyping in a multi-stage design. While a previous report suggested that using pooled samples in a 2-stage design could achieve an approximately 20-fold reduction in study cost compared to all-individual genotyping (Macgregor et al. 2008), the simulation was based on a 30-fold increase in the per genotype cost of individual custom genotyping as compared to genome-wide array genotyping. However, the cost ratio today is closer to 200-fold due to a drop in the price of genome-wide array platforms. This cost reduction makes it more attractive to genotype samples individually in the first stage. We compared the two most cost-efficient 3-stage designs, with and without pooling. We focused on a 3-stage design for two reasons: first, the high false-negative rate as a result of stringent QC implies loss of coverage and power. Thus, we reasoned that a better use of pooled data is to perform little or no QC, and use pooling as a screen in the second stage followed by additional replication. Second, given the need for very large sample sizes to detect reproducible associations with small effects, we reasoned that for both all-individual genotyping and pooled genotyping the optimal cost scenario is likely a 3-stage design with successively fewer markers genotyped rather than a 2-stage design. Indeed, with a sample of 50,000 individuals, a 3-stage design is more cost-efficient than a 2-stage design (Supplemental Table 1). Furthermore, to model pooling efficiency, we relied on the empirical data generated from an African-American cohort pooled with respect to BMI. Because few SNPs with large effect sizes exist in our empirical data, we were prevented from simulating the optimal design using much higher GRR. Given these conditions, and using current estimates of \$0.00065 per genotype for the Affymetrix 6.0 array and \$0.13 per genotype for multiplex genotyping, we determined that a 3-stage design using pooling would not present a reduction in cost, but would in fact represent a 7% cost increase, when compared to a 3-stage design using only individual genotyping (Table 4).

However, it should be noted that the results presented here are limited by both the current state of the specific technology and methodology we employed, as well as by our experimental design. First, our power to detect associated variants may also be limited by the number of replicates genotyped to control for pooling-specific error. Increasing the number of replicates may improve pooling efficiency, although it would also increase the study cost. Moreover, advances in the genotype platform and improvements in sample handling may enhance the overall data quality and AF estimates. For example, increased SNP density and redundancy on the chip should increase SNP coverage and lower the associated false-negative rate, thereby improving pooling efficiency. Indeed, a few of the reproducible associations identified using DNA pooling (Brown et al. 2008; Papassotiropoulos et al. 2006; Schrauwen et al. 2009) may partly be due to the investigators using a genotyping platform in which the reported pooling-specific error is approximately ten times lower than the platform used in this study (Craig et al. 2009; Macgregor et al. 2008). We have also speculated that a 3-stage design involving pooling would be more cost effective if either the cost of array genotyping were to decrease or pooling efficiency were to increase (Fig. 4). Therefore, a re-evaluation altering these different parameters of the study design may produce a different conclusion. The approach described here will be useful for assessing the effects of various parameters on pooling, as well as for evaluating the utility of other existing genome-wide genotyping platforms or of new, potentially more robust, SNP array platforms for conducting GWA studies using pooled DNA.

For the moment, it appears that genotyping pooled DNA remains most useful in scenarios in which either a large genetic effect is present or genome-wide level data are used concurrently to overcome errors in individual marker estimates, such as the case of estimating global admixture proportions (Chiang et al. 2010) or resolving individuals contributing trace amounts of DNA to a pool (Homer et al. 2008a). It should be noted that for many researchers, limited funding prevents a comprehensive GWA study using individual DNA samples from being conducted. In these cases, genotyping pooled DNA in a single-stage design does enrich for nominal associations and remains one of the affordable (albeit less efficient) options. Moreover, given a general lack of a priori expectation of the effect sizes for the variants associated with a previously understudied phenotype, conducting single-stage DNA pooling studies may be informative for the initial characterization of the genetic architecture of the disease as well as for the design of future association studies. For example, one may initially genotype a relatively small cohort in pools to screen for variants of large effects. If no such variants exist, a larger cohort could then be accrued and individually genotyped using a genome-wide array. The top associations from the GWA study discovery cohort could be combined with data from the preliminary pooled DNA study (to include other variants that may not have as strong evidence of association) to prioritize SNPs to be genotyped in a third, replication, cohort.

Methods

Study populations and DNA pool construction

The cohorts used in this study include Jamaican individuals from Kingston (GxE) and Spanishtown (SPT), Jamaica, respectively, as described elsewhere (Chiang et al. 2010; Cooper et al. 1997; Kang et al. 2010), as well as African American (MEC-AA), Native Hawaiian (MEC-H), Japanese American (MEC-J), Latin American (MEC-L), and non-Latina white (MEC-W) women from the Multi-Ethnic Cohort (MEC) of Los Angeles and Hawai'i, as described elsewhere (Chiang et al. 2010; Gajdos et al. 2008; Kolonel et al. 2000). In total, 20 pools were constructed from these cohorts, stratified by phenotype and gender.

In the GxE panel, individuals with BMI <24 were designated as unaffected controls, and individuals with BMI >28 were designated as affected cases. Males and females were pooled separately, but using the same BMI cut-offs. In the SPT panels, individuals with BMI <25 were designated as unaffected controls and individuals with BMI >30 were designated as affected cases. Final sizes for the GxE pools were 101 case males, 79 control males, 136, 137, and 137 case females, and 98 control females. Final sizes for the SPT pools were 129 and 135 case females and 105 and 111 control females.

The MEC samples were pooled by early (<11 years of age) or late (>14 years of age) menarche. The final MEC pools consisted of 321 MEC-AA individuals (153 late menarche and 168 early menarche), 252 MEC-H individuals (120 late and 132 early), 202 MEC-J individuals (111 late and 91 early), 332 MEC-L individuals (163 late and 169 early), and 367 MEC-W individuals (182 late and 185 early).

Pool construction has been described previously (Chiang et al. 2010). Briefly, DNA samples were pooled in an equimolar fashion, and pooled DNA were genotyped with the Affymetrix Genome-Wide Human SNP 6.0 Array according to the manufacturer's instructions (http://www.affymetrix.com/support/technical/byproduct.affx?product=genomewidesnp_6) in triplicate, along with independent individual DNA samples on the same genotyping plate. Pool replicates with excessively low intensity, low call rate, or high heterozygosity compared to the other replicates of the same pool were either re-genotyped or removed: one of the MEC-L late replicates and one of the MEC-J early replicates were each re-genotyped along with a companion replicate on a later plate; one replicate from each of three different pools (GxE control males, GxE case females, and SPT control females) were dropped from the study.

Analysis of pooled genotype data

Pooled AF in the case and control pools of each cohort were estimated as described previously (Chiang et al. 2010). Briefly, the pooled samples were genotyped genome-wide on the same plate with other individual DNA samples. The Birdseed algorithm (Korn et al. 2008) was used to estimate AA, AB, and BB cluster means and covariances of probe intensities for the individual samples as well as to call the genotypes for these samples. Informed by the covariance matrix of the three genotype clusters, for each pool we estimated the pooled AF as the proportion of angular distance observed for the pooled sample relative to that observed for the individual samples on the same plate, averaged over all replicates. SNPs were then subject to QC filtering, and post-QC SNPs were further re-parameterized before analysis (see below).

We applied four QC filters to remove SNPs that tended to genotype poorly or inconsistently in pooled DNA samples when using the Affymetrix 6.0 platform. The four filters include the Fisher's Linear Discriminant (FLD) filter, the radius of pool intensity (r) filter, the minor allele frequency (MAF) filter, and the historical data (hist) filter. Detailed descriptions of each filter can be found elsewhere (Chiang et al. 2010). The FLD filter uses the Fisher's Linear Discriminant as a measure of separation between genotype clusters; a higher value indicates better confidence in the clustering of individuals and yields more accurate estimates of AF. The r -filter measures the radial component of the signal intensity of the pooled samples in polar coordinates, relative to the intensity measured for the individual samples; a low value may indicate poor DNA quality or low quantity. The MAF-filter measures the estimated pooled AF and removes rarer SNPs because the pooled AF estimates tend to be less accurate for such SNPs. Finally, the hist-filter measures the variance in the difference in estimated AF and actual AF among replicate pools based on historical data (Chiang et al. 2010); a large variance may indicate a SNP probe on the array that is not robust through the pooling scheme. Each filter was previously trained by comparing pooled

association results to those based on genome-wide individual genotyping to establish suitable cut-off parameters. The same filter cutoff parameters were employed for both the Jamaican pools and the MEC pools; the filter cut-offs were: FLD filter 20, r -filter 0.8 in at least 80% of the replicates, MAF-filter 0.05, and hist-filter 0.06. The number of SNPs passing QC ranged from 303,627 to 499,969.

To ensure similarity in the AF distributions of the case and control pools during disease association, we also adopted a non-parametric approach to estimate AF for all SNPs passing the initial QC filters. For each cohort, an overall distribution of AF estimates was constructed by combining the pooled AF estimates from each case and control pool, weighted by the sample size of each pool, for all post-QC SNPs. For each pool, post-QC SNPs were then ranked based on pooled AF, and assigned an AF of the corresponding rank from the overall distribution. This measure improved the genomic control inflation factor, lambda, from 1.38 (after QC filters only) to 1.08 in an analysis of African-Americans pooled by high and low BMI, as described previously (Chiang et al. 2010).

For identifying SNPs associated with BMI, we compared the pooled AF (after non-parametric fitting to the overall AF distribution) from the case pool to the control pool, and then deflated the resulting χ^2 statistics to control for pooling-specific error (Chiang et al. 2010; Visscher and Le Hellard 2003). The association statistics were quantile-normalized within cohort before meta-analysis to remove any systematic bias between cohorts that might cause one cohort to dominate the meta-analyzed signal. Only SNPs that passed QC filters in both GxE and SPT were meta-analyzed using the meta-analysis tool Metal (February 2009 release, <http://www.sph.umich.edu/csg/abecasis/Metal/index.html>), weighted by population size. The top 1,000 SNPs from the meta-analyzed results were divided into encouraging, inconclusive, and discouraging categories based on the strength of association evidence in neighboring SNPs in LD ($r^2 > 0.5$ in HapMap YRI panel) with the SNP of interest (also see Chiang et al. 2010). Briefly, data from the meta-analysis of pre-QC-filtered dataset were used as indicators of the strength of association of SNPs in LD with one of the top SNPs: “Encouraging” (ENC) SNPs had at least one SNP in LD with a quantile-normalized P value < 0.05 and had at least half of the SNPs in LD with quantile-normalized P values < 0.1 ; “Discouraging” (DIS) SNPs had none of the SNPs in LD with quantile-normalized P values < 0.1 ; the remaining SNPs were categorized as “Inconclusive” (INC), a category also encompassing SNPs with no other SNPs in LD (INC-LD). After assignment into the appropriate categories, the top encouraging and inconclusive SNPs were selected for validation by individual genotyping.

For age at menarche, the analysis was conducted in the same way, with the following changes: first, each MEC population was weighted by its effective sample size during meta-analysis, which was calculated based on the pooled AF in each population, in order to account for population-specific differences in AF (Gajdos et al. 2008). Second, we meta-analyzed all SNPs that passed QC in at least three of the five MEC cohorts. Third, the LD pattern used for categorical assignment of top SNPs based on strength of association of neighboring SNPs in LD depended on the population that contributed most strongly to the meta-analyzed P value; the HapMap reference panel corresponding to the major ancestry of the population was used.

Cost optimization of 3-stage designs involving 50,000 samples

We compared two 3-stage designs, with and without pooling (Fig. 3b). Without pooling, we assumed samples were individually genome-wide genotyped in the first stage, and then individually custom-genotyped in the second and third stages. With pooling, we assumed that samples were individually genome-wide genotyped in the first stage, genome-wide

genotyped in pools in the second stage, and then individually custom-genotyped in the third stage.

We first determined the most cost-efficient 3-stage design without pooling any samples. We assumed a cohort of 25,000 cases and 25,000 selected controls, a disease prevalence of 0.1, a disease variant frequency of 0.3, and a GRR of 1.1 under the multiplicative model. Using these parameters, we obtained the expected AF among the cases and the controls from the “case-control for discrete traits” module of the genetic power calculator (<http://pngu.mgh.harvard.edu/~purcell/gpc/>). We systematically iterated over different permutations of individuals apportioned to each of stages 1 and 2 (n_1 , n_2), in increments of 1,000, with the remaining individuals assigned to stage 3 (n_3). The case-to-control ratio was 1 in all stages. For each permutation of n_1 , n_2 , and n_3 , the power at each stage was computed for 17 different levels of alpha (α_1 , α_2 , and α_3 ; 1×10^{-7} , 5×10^{-7} , 1×10^{-6} , 5×10^{-6} , 1×10^{-5} , 5×10^{-5} , 1×10^{-4} , 5×10^{-4} , 0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5). Two-tailed power was calculated for stage 1; one-tailed power was calculated for stages 2 and 3. The total power of the 3-stage design is the product of the power at each of the three stages and was calculated for all permutations of alpha levels for each stage.

The pricing for genome-wide genotyping using the Affymetrix 6.0 SNP array and for custom genotyping using the Sequenom MassARRAY platform were obtained from the Children’s Hospital Boston Microarray Core and SNP Genotyping Core Facilities as of February 2010. The full price of the Affymetrix 6.0 array (~1 million SNPs) for outside investigators was \$650 per sample; the custom genotyping cost was \$1,100 per genotyping plate of 384 individuals plus \$20 per SNP primer design (typically 28–35 SNPs are designed and assayed together). These figures translate to a cost of ~\$0.00065 per genotype for the stage 1 simulation and ~\$0.13 to \$0.15 per genotype for custom genotyping in subsequent stages. For simplicity, we assumed a 200× difference between the per genotype cost of genome-wide genotyping and custom genotyping (i.e., \$0.00065 and \$0.13). This estimated pricing includes the labor costs, but not the analytical costs.

The total cost for each all-individual 3-stage design was based on genome-wide genotyping in stage 1, with custom genotyping in the individuals apportioned to subsequent stages only for SNPs that would surpass the alpha level in the previous stage. We considered the optimized design to be the one with the lowest cost that achieved at least 80% power with <1 false positive (the number of SNPs expected to reach the alpha levels for each stage under the null).

The most cost-efficient design involving pooled genotyping in stage 2 was determined similarly, with the following modifications. When determining the power of the second stage, we needed to consider the effect of pooling. Thus, we multiplied the power due to genotyping n_2 individuals at alpha level α_2 by the probability that the association statistic of a SNP reaching α_2 would be ranked above a certain level of quantile-normalized P value in pooling (p_q ; 14 levels were considered: 0.0025, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1). $\Pr(p_q|\alpha_2)$ was modeled empirically using an African-American cohort previously individually genotyped and pooled according to high and low BMI (Chiang et al. 2010; Kang et al. 2010). At each level of p_q and α_2 , $\Pr(p_q|\alpha_2)$ was calculated as the number of SNPs with actual P value $<\alpha_2$ and one-tailed quantile-normalized P value $<p_q$ divided by all SNPs with actual P value $<\alpha_2$. All SNPs with one-tailed quantile-normalized P value $<p_q$ during stage 2 would be followed up in stage 3. When determining the cost of each 3-stage design, individuals in stage 2 were assumed to be pooled in groups of 150 and genotyped in triplicate using the genome-wide genotyping array. Furthermore, in order to maintain the same level of power in the 3-stage design involving pooling, often an increased number of SNPs would reach the specified alpha level under the null. To penalize the increased false

positive rate, these SNPs would be individually genotyped in all individuals apportioned to stage 2, and the cost would be added to the overall cost of the design. The design achieving 80% total power with the lowest cost was then deemed the optimized 3-stage design involving pooling.

In evaluating the cost-efficiency involving pooling under scenarios where genotyping costs have decreased, we assumed the per genotype cost using a genome-wide array to be \$0.000325 and \$0.000163, or the per genotype cost using custom genotyping to be \$0.065 or \$0.0325. To simulate scenarios with 2× and 3× improved pooling efficiency, we assumed that the same level of $\Pr(p_q|a_2)$ could be achieved with p_q that is two times or three times smaller, respectively.

Technical validation by individual genotyping

Twenty-one candidate SNPs associated with BMI (18 of the top 36 encouraging SNPs and three of the top seven inconclusive SNPs) were successfully genotyped in the individuals comprising the GxE and SPT pools. Genotyping was performed as described elsewhere (Gabriel et al. 2002), using the Sequenom MassARRAY platform (Tang et al. 1999) with the iPLEX protocol. The basic protocol involves a multiplex primer extension followed by matrix-assisted laser desorption ionization-time of flight mass spectroscopy detection (Tang et al. 1999). SNPs were considered working if they genotyped successfully in 90% or more of the samples and had no more than one consensus error among 32 triplicates in SPT and 34 triplicates in GxE (for a total of 96 and 102 internal comparisons per SNP, respectively). SNPs were evaluated by these parameters in the SPT and GxE cohorts independently. Association tests for candidate SNPs associated with BMI were conducted using PLINK v1.05 (<http://pngu.mgh.harvard.edu/purcell/plink/>; Purcell et al. 2007).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The principal acknowledgement is to the participants who contributed their time, biological samples, and phenotype data to the different projects. The authors also thank P.I. de Bakker, P. Sklar, and past and present members of the Hirschhorn laboratory for comments, ideas, and discussions; the past and present team members and investigators at the Tropical Medicine Research Institute and the Hawai'i and Los Angeles Multi-Ethnic Cohort for collecting samples and data. This work was supported by a graduate research fellowship from the National Science Foundation to C.W.K.C. and grants from the National Institutes of Health to J.N.H. (R01DK075787), to R.S.C. (R37HL45508 and R01HL53353), to X.Z. (R01HL074166), and to M.R.P. (R01HD048960). The funders had no role in study design, data collection, analysis, interpretation, decision to publish, or preparation of the manuscript.

Abbreviations

AF	Allele frequency
GRR	Genotypic relative risk
GWA	Genome-wide association
LD	Linkage disequilibrium
MEC	Multi-Ethnic Cohort
QC	Quality control

References

- Brown KM, Macgregor S, Montgomery GW, Craig DW, Zhao ZZ, Iyadurai K, Henders AK, Homer N, Campbell MJ, Stark M, Thomas S, Schmid H, Holland EA, Gillanders EM, Duffy DL, Maskiell JA, Jetann J, Ferguson M, Stephan DA, Cust AE, Whiteman D, Green A, Olsson H, Puig S, Ghiorzo P, Hansson J, Demenais F, Goldstein AM, Gruis NA, Elder DE, Bishop JN, Kefford RF, Giles GG, Armstrong BK, Aitken JF, Hopper JL, Martin NG, Trent JM, Mann GJ, Hayward NK. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet.* 2008; 40:838–840. [PubMed: 18488026]
- Butcher LM, Meaburn E, Knight J, Sham PC, Schalkwyk LC, Craig IW, Plomin R. SNPs, microarrays and pooled DNA: identification of four loci associated with mild mental impairment in a sample of 6000 children. *Hum Mol Genet.* 2005; 14:1315–1325. [PubMed: 15800012]
- Butcher LM, Davis OS, Craig IW, Plomin R. Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. *Genes Brain Behav.* 2008; 7:435–446. [PubMed: 18067574]
- Chiang CW, Gajdos ZK, Korn JM, Kuruvilla FG, Butler JL, Hackett R, Guiducci C, Nguyen TT, Wilks R, Forrester T, Haiman CA, Henderson KD, Le Marchand L, Henderson BE, Palmert MR, McKenzie CA, Lyon HN, Cooper RS, Zhu X, Hirschhorn JN. Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. *PLoS Genet.* 2010; 6:e1000866. [PubMed: 20221249]
- Cooper R, Rotimi C, Ataman S, McGee D, Osotimehin B, Kadiri S, Muna W, Kingue S, Fraser H, Forrester T, Bennett F, Wilks R. The prevalence of hypertension in seven populations of west African origin. *Am J Public Health.* 1997; 87:160–168. [PubMed: 9103091]
- Craig JE, Hewitt AW, McMellon AE, Henders AK, Ma L, Wallace L, Sharma S, Burdon KP, Visscher PM, Montgomery GW, MacGregor S. Rapid inexpensive genome-wide association using pooled whole blood. *Genome Res.* 2009; 19:2075–2080. [PubMed: 19801603]
- Davis OS, Butcher LM, Docherty SJ, Meaburn EL, Curtis CJ, Simpson MA, Schalkwyk LC, Plomin R. A three-stage genome-wide association study of general cognitive ability: hunting the small effects. *Behav Genet.* 2010; 40:759–767. [PubMed: 20306291]
- Docherty SJ, Butcher LM, Schalkwyk LC, Plomin R. Applicability of DNA pools on 500K SNP microarrays for cost-effective initial screens in genomewide association studies. *BMC Genomics.* 2007; 8:214. [PubMed: 17610740]
- Docherty SJ, Davis OS, Kovas Y, Meaburn EL, Dale PS, Petrill SA, Schalkwyk LC, Plomin R. A genome-wide association study identifies multiple loci associated with mathematics ability and disability. *Genes Brain Behav.* 2010; 9:234–247. [PubMed: 20039944]
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science.* 2002; 296:2225–2229. [PubMed: 12029063]
- Gajdos ZK, Butler JL, Henderson KD, He C, Supelak PJ, Egyud M, Price A, Reich D, Clayton PE, Le Marchand L, Hunter DJ, Henderson BE, Palmert MR, Hirschhorn JN. Association studies of common variants in 10 hypogonadotropic hypogonadism genes with age at menarche. *J Clin Endocrinol Metab.* 2008; 93:4290–4298. [PubMed: 18728166]
- Gajdos ZK, Henderson KD, Hirschhorn JN, Palmert MR. Genetic determinants of pubertal timing in the general population. *Mol Cell Endocrinol.* 2010
- Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA.* 2009; 106:9362–9367. [PubMed: 19474294]
- Homer N, Szelinger S, Redman M, Duggan D, Tembe W, Muehling J, Pearson JV, Stephan DA, Nelson SF, Craig DW. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet.* 2008a; 4:e1000167. [PubMed: 18769715]
- Homer N, Tembe WD, Szelinger S, Redman M, Stephan DA, Pearson JV, Nelson SF, Craig D. Multimarker analysis and imputation of multiple platform pooling-based genome-wide association studies. *Bioinformatics.* 2008b; 24:1896–1902. [PubMed: 18617537]

- Iles MM. What can genome-wide association studies tell us about the genetics of common disease? *PLoS Genet.* 2008; 4:e33. [PubMed: 18454206]
- Kang SJ, Chiang CW, Palmer CD, Tayo BO, Lettre G, Butler JL, Hackett R, Adeyemo AA, Guiducci C, Berzins I, Nguyen TT, Feng T, Luke A, Shriner D, Ardlie K, Rotimi C, Wilks R, Forrester T, McKenzie CA, Lyon HN, Cooper RS, Zhu X, Hirschhorn JN. Genome-wide association of anthropometric traits in African- and African-derived populations. *Hum Mol Genet.* 2010; 19:2725–2738. [PubMed: 20400458]
- Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, Stram DO, Monroe KR, Earle ME, Nagamine FS. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol.* 2000; 151:346–357. [PubMed: 10695593]
- Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MM, Gabriel SB, Purcell S, Daly MJ, Altshuler D. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet.* 2008; 40:1253–1260. [PubMed: 18776909]
- Macgregor S, Zhao ZZ, Henders A, Nicholas MG, Montgomery GW, Visscher PM. Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays. *Nucleic Acids Res.* 2008; 36:e35. [PubMed: 18276640]
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 2008; 9:356–369. [PubMed: 18398418]
- Meaburn E, Butcher LM, Schalkwyk LC, Plomin R. Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans. *Nucleic Acids Res.* 2006; 34:e27. [PubMed: 16478714]
- Meaburn EL, Harlaar N, Craig IW, Schalkwyk LC, Plomin R. Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Mol Psychiatry.* 2008; 13:729–740. [PubMed: 17684495]
- Ong KK, Elks CE, Li S, Zhao JH, Luan J, Andersen LB, Bingham SA, Brage S, Smith GD, Ekelund U, Gillson CJ, Glaser B, Golding J, Hardy R, Khaw KT, Kuh D, Luben R, Marcus M, McGeehin MA, Ness AR, Northstone K, Ring SM, Rubin C, Sims MA, Song K, Strachan DP, Vollenweider P, Waeber G, Waterworth DM, Wong A, Deloukas P, Barroso I, Mooser V, Loos RJ, Wareham NJ. Genetic variation in LIN28B is associated with the timing of puberty. *Nat Genet.* 2009; 41:729–733. [PubMed: 19448623]
- Papassotiropoulos A, Stephan DA, Huentelman MJ, Hoernkli FJ, Craig DW, Pearson JV, Huynh KD, Brunner F, Corneveaux J, Osborne D, Wollmer MA, Aerni A, Coluccia D, Hanggi J, Mondadori CR, Buchmann A, Reiman EM, Caselli RJ, Henke K, de Quervain DJ. Common Kibra alleles are associated with human memory performance. *Science.* 2006; 314:475–478. [PubMed: 17053149]
- Pearson JV, Huentelman MJ, Halperin RF, Tembe WD, Melquist S, Homer N, Brun M, Szelinger S, Coon KD, Zismann VL, Webster JA, Beach T, Sando SB, Aasly JO, Heun R, Jessen F, Kolsch H, Tsolaki M, Daniilidou M, Reiman EM, Papassotiropoulos A, Hutton ML, Stephan DA, Craig DW. Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies. *Am J Hum Genet.* 2007; 80:126–139. [PubMed: 17160900]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81:559–575. [PubMed: 17701901]
- Schrauwen I, Ealy M, Huentelman MJ, Thys M, Homer N, Vanderstraeten K, Franssen E, Corneveaux JJ, Craig DW, Claustres M, Cremers CW, Dhooge I, Van de Heyning P, Vincent R, Offeciers E, Smith RJ, Van Camp G. A genome-wide analysis identifies genetic variants in the RELN gene associated with otosclerosis. *Am J Hum Genet.* 2009; 84:328–338. [PubMed: 19230858]
- Shifman S, Bhomra A, Smiley S, Wray NR, James MR, Martin NG, Hettema JM, An SS, Neale MC, van den Oord EJ, Kendler KS, Chen X, Boomsma DI, Middeldorp CM, Hottenga JJ, Slagboom PE, Flint J. A whole genome association study of neuroticism using DNA pooling. *Mol Psychiatry.* 2008a; 13:302–312. [PubMed: 17667963]
- Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, Craddock NJ, Kendler KS, Li T, O'Donovan M, O'Neill FA, Owen MJ, Walsh D, Weinberger DR, Sun C, Flint J, Darvasi A.

Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. *PLoS Genet.* 2008b; 4:e28. [PubMed: 18282107]

Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet.* 2006; 38:209–213. [PubMed: 16415888]

Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, 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, Nemesh J, Park JH, Gustafsson S, Kilpelainen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pelliikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebertmann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet.* 2010; 42:937–948. [PubMed: 20935630]

Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Koster H. Chip-based genotyping by mass spectrometry. *Proc Natl Acad Sci USA.* 1999; 96:10016–10020. [PubMed: 10468554]

Visscher PM, Le Hellard S. Simple method to analyze SNP-based association studies using DNA pools. *Genet Epidemiol.* 2003; 24:291–296. [PubMed: 12687646]

Yang HC, Huang MC, Li LH, Lin CH, Yu AL, Diccianni MB, Wu JY, Chen YT, Fann CS. MPDA: microarray pooled DNA analyzer. *BMC Bioinformatics.* 2008; 9:196. [PubMed: 18412951]

Zhang H, Yang HC, Yang Y. PooL: an efficient method for estimating haplotype frequencies from large DNA pools. *Bioinformatics.* 2008; 24:1942–1948.

Zuo Y, Zou G, Zhao H. Two-stage designs in case-control association analysis. *Genetics.* 2006; 173:1747–1760. [PubMed: 16624925]

Zuo Y, Zou G, Wang J, Zhao H, Liang H. Optimal two-stage design for case-control association analysis incorporating genotyping errors. *Ann Hum Genet.* 2008; 72:375–387. [PubMed: 18215207]

Fig. 1.

Estimated allele frequencies in cases versus controls before and after application of QC filters. Estimated allele frequencies of the case and control pools from the SPT panel were plotted against each other **a** before SNP QC filtering, **b** after applying all four SNP QC filters, and **c** after AF distribution fitting of post-QC SNPs. Among the 5,000 SNPs with the largest AF differences between case and control pools, the mean AF difference in the post-QC filtered dataset was significantly reduced (0.54 pre-QC vs. 0.19 post-QC, $P \ll 10^{-15}$, by unpaired two-tailed t test). Fitting of the AF estimates to the overall distribution (see Supplemental Methods) lessened the systemic bias in AF estimation between the case and control pools, and thus the mean difference among the 5,000 SNPs in the post-adjustment dataset was also significantly reduced (0.19 post-QC vs. 0.17 post-fitting, $P \ll 10^{-15}$, by unpaired two-tailed t test). It should be noted that this comparison is based only on the average of the AF estimates, without taking into account the error involved in such estimates, which is corrected for when calculating the association χ^2 statistic (see “Methods”)

Fig. 2.

Meta-analysis of genome-wide association results for BMI and age at menarche. Quantile–quantile plot for **a** BMI in GxE and SPT pools and **b** age at menarche in MEC pools. In both cases, the association statistics in each pool were separately quantile-normalized before meta-analysis. 95% confidence intervals are shown in *gray*. The genomic control inflation factor, lambda, is 0.99 and 1.02 for BMI and age at menarche, respectively

Fig. 3.

Schematics of incorporating pooled genotyping in a multistage design. To assess the utility of genotyping pooled DNA as part of a multi-stage association design, we considered two scenarios. **a** A GWA study by individual genotyping has been done using a discovery cohort, with additional samples for replication available. However, because of the limited funding for individual custom genotyping, only the top hits (depicted here as those that reached $P \leq 10^{-5}$) could be genotyped. With pooling, we evaluated whether the list of SNPs with more moderate association evidence (depicted here as those with $0.001 < P \leq 10^{-5}$) could be prioritized to supplement the top hits as additional SNPs for replication. **b** A 3-stage design where a large cohort of individuals was subdivided into three non-overlapping groups, each genotyped in successive stages. If pooled genotyping were used, we considered pooling stage 2 individuals. N_{1A} , N_{2A} , and N_{3A} represent the number of individuals genotyped in stage 1, 2, and 3, respectively, using all-individual genotyping; N_{1B} , N_{2B} , and N_{3B} represent the number of individuals genotyped in stage 1, 2, and 3, respectively, under the pooling scheme. Note that $\sum N_{iA} = \sum N_{iB} = \text{total } N$, for $i = 1, 2, 3$. In our simulation, we assumed N to be 50,000. Boxed stages denote individual genotyping either by genome-wide SNP arrays or custom multiplex genotyping. *Circled* stages denote pooled genotyping by genome-wide SNP arrays

Fig. 4.

Comparison of the most cost-efficient designs with and without pooling under various scenarios of genotyping cost and pooling efficiency. Under various pricing scenarios for the genome-wide genotyping cost (“Affymetrix Cost”) and custom genotyping cost (“iPLEX Cost”) and different degrees of efficiency due to pooling, we determined the ratio between the most cost-efficient 3-stage design involving pooling to that without pooling. A relative cost < 1 means the scenario with pooling is more cost efficient. The long dashed line used current pricing for genome-wide and custom genotyping, which reflects a roughly 200 $\times$ difference in per genotype cost. To model pooling efficiency, we assumed that same proportion of SNPs at any given level of actual association can be identified by a pooling P value that is half or a third of the current pooling P value (see “Methods” for more details)

Table 1

Technical validation of SNPs showing differences in AF between lean and obese pools

Marker	GxE				SPT				Meta P	Pool P	Category		
	A1	FA	FU	OR	P	FA	FU	OR				P	
rs4793952	C	0.73	0.63	1.57		6.0 × 10 ⁻⁴	0.75	0.69	1.39	0.022	4.1 × 10 ⁻⁵	1.0 × 10 ⁻⁶	ENC
rs7205918	T	0.68	0.62	1.32		0.034	0.68	0.59	1.48	0.0042	5.2 × 10 ⁻⁴	1.7 × 10 ⁻⁵	ENC
rs9975800	A	0.30	0.21	1.61		0.0012	0.30	0.26	1.22	0.17	7.8 × 10 ⁻⁴	1.3 × 10 ⁻⁴	ENC
rs274793	C	0.68	0.62	1.30		0.039	0.72	0.64	1.43	0.011	0.0013	2.2 × 10 ⁻⁴	ENC
rs844850	G	0.44	0.51	0.75		0.019	0.42	0.47	0.81	0.11	0.0045	0.0014	ENC
rs164009	G	0.82	0.77	1.40		0.024	0.85	0.81	1.33	0.098	0.0053	5.0 × 10 ⁻⁴	ENC
rs302848	G	n.a.	n.a.	n.a.		n.a.	0.23	0.30	0.68	0.010	0.010	4.3 × 10 ⁻⁴	ENC
rs7906620	C	0.55	0.64	0.70		0.0053	0.56	0.58	0.93	0.60	0.013	1.5 × 10 ⁻⁴	ENC
rs7717789	A	0.44	0.38	1.27		0.058	0.42	0.39	1.16	0.28	0.031	0.0021	ENC
rs12996442	T	0.91	0.88	1.38		0.11	0.92	0.90	1.36	0.17	0.035	0.0026	ENC
rs6863262	C	0.72	0.65	1.34		0.027	0.70	0.69	1.08	0.59	0.041	9.7 × 10 ⁻⁴	ENC
rs2647447	C	n.a.	n.a.	n.a.		n.a.	0.79	0.74	1.37	0.043	0.043	9.9 × 10 ⁻⁴	ENC
rs13259717	C	0.78	0.74	1.29		0.089	0.78	0.75	1.17	0.29	0.048	0.0022	ENC
rs7545314	C	0.79	0.76	1.18		0.26	0.82	0.77	1.31	0.096	0.054	0.0011	ENC
rs6472664	T	0.76	0.73	1.15		0.33	0.77	0.72	1.29	0.084	0.064	2.1 × 10 ⁻⁴	ENC
rs3829366	A	0.32	0.35	0.89		0.38	0.34	0.38	0.83	0.17	0.12	0.0016	ENC
rs1467328	C	0.63	0.58	1.23		0.11	0.62	0.62	1.01	0.96	0.20	0.0015	ENC
rs2493374	A	0.24	0.24	1.00		0.99	0.27	0.21	1.35	0.051	0.21	7.0 × 10 ⁻⁴	ENC
rs10061785	A	0.80	0.74	1.40		0.020	0.85	0.79	1.49	0.019	9.8 × 10 ⁻⁴	2.9 × 10 ⁻⁵	INC-LD
rs12469822	G	0.86	0.80	1.49		0.013	0.85	0.83	1.20	0.30	0.010	3.8 × 10 ⁻⁵	INC-LD
rs10735098	A	0.55	0.51	1.16		0.217	0.55	0.48	1.32	0.034	0.021	6.2 × 10 ⁻⁵	INC-LD

Eighteen "encouraging" SNPs and three "inconclusive" SNPs were genotyped in the individuals that comprised the pools in GxE and SPT. Each individual was designated as affected (high BMI) or unaffected (low BMI) in the same manner as pool assignments (see "Methods"). Association analysis was performed in each population separately and results were then meta-analyzed. For all SNPs, the direction of effect of the reference allele estimated by pooling and by individual genotyping matched

A1 reference allele on the positive strand, FA allele frequency in affected, FU allele frequency in unaffected, OR odds ratio, P P-value, Meta P meta-analyzed P value using individual genotypes, Pool P P value resulting from meta-analysis of quantile-normalized P using pooled DNA samples, n.a. results not available due to failed genotyping, ENC-SNPs considered "encouraging" because other SNPs in LD also showed strong evidence of association, INC-LD subset of SNPs considered "inconclusive" because there were no SNPs in strong LD (see "Methods")

Table 2
Technical validation of SNPs showing differences in AF between late and early menarche pools

Marker	A1	MEC-AA		MEC-H		MEC-J		MEC-L		MEC-W		Meta P	Pool P	Category
		FA	FU	FA	FU	FA	FU	FA	FU	FA	FU			
rs12414024	A	0.78	0.67	0.55	0.52	0.60	0.53	0.61	0.58	0.61	0.48	3.4×10^{-5}	3.7×10^{-5}	ENC
rs2242045	A	0.29	0.20	0.33	0.27	0.28	0.28	0.30	0.22	0.55	0.45	4.4×10^{-5}	1.1×10^{-5}	ENC
rs6574495	C	0.37	0.32	0.72	0.69	0.82	0.79	0.42	0.38	0.67	0.56	0.0017	8.9×10^{-5}	ENC
rs453186	T	0.10	0.10	0.47	0.53	0.37	0.42	0.25	0.31	0.25	0.33	0.0035	0.0044	ENC
rs216150	G	0.72	0.66	0.63	0.55	0.48	0.51	0.72	0.68	0.68	0.63	0.013	3.1×10^{-4}	ENC
rs11928013	C	0.70	0.63	0.99	0.99	1.00	1.00	0.97	0.97	0.99	0.98	0.046	1.8×10^{-4}	ENC
rs4145975	T	0.57	0.66	0.40	0.45	0.50	0.49	0.45	0.58	0.43	0.47	2.5×10^{-4}	7.7×10^{-5}	INC-LD
rs8258	C	0.66	0.64	0.44	0.41	0.48	0.41	0.46	0.36	0.68	0.64	0.0043	9.9×10^{-6}	INC-LD
rs11623520	A	0.38	0.32	0.48	0.47	0.64	0.67	0.75	0.68	0.73	0.64	0.0047	3.9×10^{-5}	INC-LD
rs4669378	T	0.33	0.38	0.64	0.65	0.79	0.82	0.52	0.58	0.49	0.56	0.0089	7.1×10^{-5}	INC-LD
rs9912513	T	0.80	0.77	0.40	0.52	0.62	0.63	0.54	0.59	0.52	0.59	0.017	0.00079	INC-LD
rs10070239	G	0.38	0.46	0.11	0.15	0.17	0.19	0.27	0.34	0.09	0.05	0.024	4.2×10^{-6}	INC-LD
rs1104039	G	0.18	0.20	0.33	0.40	0.42	0.44	0.39	0.47	0.28	0.24	0.072	4.6×10^{-5}	INC-LD
rs921048	C	0.43	0.35	0.30	0.27	0.12	0.13	0.43	0.40	0.51	0.51	0.15	2.2×10^{-5}	INC-LD
rs11010290	G	0.04	0.03	0.16	0.15	0.06	0.03	0.08	0.06	0.18	0.18	0.27	3.2×10^{-4}	INC-LD
rs9523165	T	0.82	0.80	0.48	0.45	0.42	0.35	0.60	0.56	0.72	0.65	0.0084	2.9×10^{-4}	INC
rs1079727	T	0.89	0.92	0.64	0.62	0.71	0.55	0.67	0.64	0.88	0.81	0.0096	0.0022	INC
rs12807920	T	0.28	0.37	0.49	0.60	0.53	0.48	0.61	0.67	0.51	0.49	0.036	6.5×10^{-5}	INC
rs4295728	T	0.84	0.84	0.45	0.47	0.45	0.39	0.47	0.38	0.66	0.62	0.060	9.8×10^{-4}	INC
rs10278268	G	0.80	0.84	0.74	0.78	0.89	0.82	0.78	0.72	0.72	0.72	0.55	8.3×10^{-4}	INC
rs11015463	T	0.15	0.17	0.08	0.07	0.04	0.01	0.29	0.28	0.20	0.23	0.86	0.0030	INC
rs4783880	A	0.74	0.65	0.80	0.74	0.70	0.66	0.64	0.61	0.86	0.86	0.010	1.6×10^{-5}	DIS
rs7151505	A	0.50	0.51	0.46	0.43	0.31	0.33	0.58	0.63	0.52	0.63	0.048	1.8×10^{-4}	DIS
rs17492120	C	0.98	0.96	0.93	0.96	0.98	0.96	0.93	0.90	0.90	0.87	0.091	0.0020	DIS
rs7788115	T	0.03	0.05	0.07	0.07	0.03	0.04	0.10	0.10	0.13	0.15	0.46	6.8×10^{-4}	DIS

Twenty-five of the top 2,000 associations by pooling had been previously individually genotyped in each of the MEC cohorts. Among the 25, 15 would have been categorized as “encouraging” or “inconclusive” due to no SNPs in strong LD; the remaining 10 would have been categorized as “inconclusive” or “discouraging”. Association analysis was performed in each population separately and results were then meta-analyzed

A1 reference allele on the positive strand, *FA* allele frequency in affected, *FU* allele frequency in unaffected, *Meta P* meta-analyzed *P*-value using individual genotypes, *Pool P* *P*-value resulting from meta-analysis of quantile-normalized *P* using pooled DNA samples, *ENC* SNPs considered “encouraging” because other SNPs in LD also showed strong evidence of association, *INC-LD* subset of SNPs considered “inconclusive” because there were no SNPs in strong LD, *INC* subset of SNPs considered “inconclusive” due to other SNPs in LD showing weak or marginal evidence of association, *DIS* subset of SNPs considered “discouraging” because other SNPs in LD showed poor evidence of association (see “Methods”).

Table 3

Using pooled genotyping to supplement the top association signals from a GWA study when selecting additional variants to follow up

Marker	GWA study association statistics			Replication panel statistics	
	A1	Effect	P value	1-tail pooling P	1-tail meta P
Did not consider pooling information					
rs415277	G	-0.318	4.6×10^{-7}	n.d.	0.75
rs8077681	C	0.188	2.2×10^{-6}	n.d.	0.71
rs7233501	G	0.327	8.5×10^{-6}	n.d.	0.99
rs6495634	A	-0.323	1.6×10^{-6}	n.d.	0.57
rs3003172	G	-0.237	9.3×10^{-7}	n.d.	0.29
rs6564329	G	0.272	8.9×10^{-6}	n.d.	0.90
rs739750	G	-0.196	4.9×10^{-6}	n.d.	0.16
rs415277	G	-0.318	4.6×10^{-7}	n.d.	0.75
Considered pooling information					
rs1862198	A	-0.141	4.0×10^{-4}	0.064	0.0049
rs10778056	T	0.136	3.1×10^{-4}	0.0030	0.016
rs10770107	T	-0.144	1.6×10^{-4}	0.049	0.026
rs12584838	G	-0.446	1.3×10^{-5}	0.046	0.038
rs2525051	G	-0.144	2.5×10^{-4}	0.024	0.048
rs10513376	T	0.144	5.0×10^{-4}	0.049	0.074
rs284320	T	-0.133	4.3×10^{-4}	0.049	0.33
rs7323331	T	-0.124	1.3×10^{-3}	0.0083	0.36
rs10483863	A	0.165	1.2×10^{-4}	0.065	0.40
rs7387855	T	0.146	4.5×10^{-4}	0.031	0.45

An African-derived cohort had previously been surveyed genome-wide for variants associated with BMI as a quantitative trait (Kang et al. 2010). Individuals from GxE and SPT were used for replication in the study. The top independent signals with GWA $P < 1 \times 10^{-5}$ were screened in another cohort of African-Americans for directional consistency, and then genotyped in the Jamaican cohorts. Eight such SNPs were genotyped in the Jamaican cohorts, but none reached nominal significance. Additionally, we supplemented the list of top SNPs with SNPs that did not attain as low P values but that passed QC filtering in pooling and had a one-tailed quantile-normalized pooling $P < 0.1$. Five of ten such SNPs reached nominal significance (one-tailed P value < 0.05) when individually genotyped. Note that here, individual genotyping results in the Jamaican cohorts were analyzed in a case-control fashion using only the individuals that comprised the pools, to allow comparison to the pooling analysis, and thus differ from the published results (Kang et al. 2010).

A / effect allele on the positive strand, *Effect* and *P* value the effect size and *P* value from a previously conducted GWA study in African-derived populations (Kang et al. 2010), *I-tail pooling* *P* the one-tailed quantile-normalized *P* value in analysis of pooled DNA of GxE and SPT, *I-tail meta* *P* the one-tailed *P* value based on individual genotypes for only the individuals that consisted the pools, *n.d.* not determined

Table 4

Comparison of the most cost-efficient 3-stage designs with and without pooling

Overall Power	n_1	a_1	n_2	a_2^a	P_{pool}^a	$\Pr(P a_2)$	n_3	a_3^a	Cost ($\times \$1000$)	Ratio
Without pooling										
0.803	15,000	0.001	8,000	0.1	n.a.	n.a.	27,000	0.0025	11,141	1.00
With pooling										
0.801	14,000	0.005	28,000	5×10^{-5}	0.35	0.9437	8,000	0.1	11,921	1.07

The most cost-efficient 3-stage designs, determined as described in “Methods”, are shown, comparing scenarios with and without pooling. We assumed a variant with GRR = 1.1, a sample size of 25,000 cases and 25,000 controls, and 1 million SNPs were genotyped in the initial stage

n_1 , n_2 , and n_3 the number of samples to be included at stages 1, 2, and 3, a_1 , a_2 , and a_3 the alpha threshold for each of the three stages that together will preserve the overall power targeted, P_{pool} the quantile-normalized P value in pooling (i.e., the percentile for the association statistics of a particular variant), $\Pr(P|a_2)$ the probability that a particular variant reaches one-tail quantile-normalized P value of P_{pool} (or better), given an actual P value a_2 , $n.a.$ not applicable

^aOne-tailed thresholds. We required a_2 in this case to be greater than 5×10^{-5} , as lower values of a_2 may not be well-modeled by our African-American pools (Chiang et al. 2010) due to the small number of observations of such SNPs. The cost of each design is displayed in units of \$1,000. The cost of designs involving pooling assumed pools of 150 individuals and had already penalized the greater number of false positives arising from the pooling design in order to achieve comparable levels of power (see text and “Methods”)