On the Replication of Genetic Associations: Timing Can Be Everything!

Jessica Lasky-Su,[1,2,17,*] Helen N. Lyon,[6,7,8,17] Valur Emilsson,[5,14] Iris M. Heid,[3,4] Cliona Molony,[14] Benjamin A. Raby,[2] Ross Lazarus,[2] Barbara Klanderman,[2] Manuel E. Soto-Quiros,[10] Lydiana Avila,[10] Edwin K. Silverman,[2] Gudmar Thorleifsson,[5] Unnur Thorsteinsdottir,[5] Florian Kronenberg,[9] Caren Vollmert,[3] Thomas Illig,[3] Caroline S. Fox,[11] Daniel Levy,[11] Nan Laird,[12] Xiao Ding,[12] Matt B. McQueen,[13] Johannah Butler,[6,7] Kristin Ardlie,[7] Constantina Papoutsakis,[16] George Dedoussis,[16] Christopher J. O'Donnell,[11] H.-Erich Wichmann,[3,4] Juan C. Celedón,[2] Eric Schadt,[14] Joel Hirschhorn,[6,7,15] Scott T. Weiss,[2] Kari Stefansson,[5] and Christoph Lange[2,12,*]

The failure of researchers to replicate genetic-association findings is most commonly attributed to insufficient statistical power, population stratification, or various forms of between-study heterogeneity or environmental influences.[1] Here, we illustrate another potential cause for nonreplications that has so far not received much attention in the literature. We illustrate that the strength of a genetic effect can vary by age, causing "age-varying associations." If not taken into account during the design and the analysis of a study, age-varying genetic associations can cause nonreplication. By using the 100K SNP scan of the Framingham Heart Study, we identified an age-varying association between a SNP in *ROBO1* and obesity and hypothesized an age-gene interaction. This finding was followed up in eight independent samples comprising 13,584 individuals. The association was replicated in five of the eight studies, showing an age-dependent relationship (one-sided combined p = 3.92×10^{-9}, combined p value from pediatric cohorts = 2.21×10^{-8}, combined p value from adult cohorts = 0.00422). Furthermore, this study illustrates that it is difficult for cross-sectional study designs to detect age-varying associations. If the specifics of age- or time-varying genetic effects are not considered in the selection of both the follow-up samples and in the statistical analysis, important genetic associations may be missed.

Introduction

It is generally accepted that replication of identified associations in other populations is necessary to validate initial findings (that is, the differentiation of true-positive from false-positive results) and possibly generalize them to other populations. However, genetic-association studies have faced the challenge of inconsistent replication, both in the context of genome-wide association (GWA) studies and, more commonly, in the context of candidate-gene studies.[2–7] Recent technological advances enabling large-scale SNP genotyping at relatively low cost, the development of several such platforms with polymorphism content of sufficient density, and the advent of powerful statistical methods for the analysis of such data have resulted in the successful implementation of GWA studies and mapping of complex-trait susceptibility loci in several diseases, including age-related macular degeneration, diabetes, obesity, and inflammatory bowel disease.[8,9] The number of whole-genome scans on large cohorts is increasing rapidly, making replication imperative for the detection of valid associations. A 100K scan in the Framingham Heart Study (FHS) identified an association between the phenotype body mass index (BMI) and a common variant, a SNP with minor allele frequency of 0.30 located near the *INSIG2* gene (*INSIG2*). In the original paper by Herbert et. al[3] as well as in subsequent replication attempts by other groups, the original association between BMI and the SNP near *INSIG2* was confirmed in some but not all cohorts, emphasizing both the complexity and the challenges for replications of association findings for complex traits.[9–14]

Difficulty in replicating genetic-association findings is most commonly attributed to insufficient statistical power, population stratification, or various forms of between-study heterogeneity, including differences in genetic ancestry (i.e., linkage-disequilibrium patterns), ascertainment schema (e.g., controls for different diseases are used to replicate an association for different phenotypes), and environmental influences.[1] Here, we illustrate that genetic

[1]SUNY Upstate Medical University, Syracuse, NY 13210, USA; [2]Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA; [3]GSF National Research Centre for Environment and Health, Institute of Epidemiology, 85764 Neuherberg, Germany; [4]Institute of Medical Informatics, Biometry, and Epidemiology, Ludwig-Maximilians-University, 80539 Munich, Germany; [5]deCode Genetics, IS-101 Reykjavik, Iceland; [6]Divisions of Genetics and Endocrinology, Program in Genomics, Children's Hospital, Boston, MA 02115, USA; [7]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA; [8]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA; [9]Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, 6020 Innsbruck, Austria; [10]Division of Pediatric Pulmonology, Hospital Nacional de Niños, PO Box 1654-1000, San José, Costa Rica; [11]National Heart, Lung, and Blood Institute and its Framingham Heart Study, Framingham, MA 01702, USA; [12]Harvard School of Public Health, Boston, MA 02115, USA; [13]Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80309, USA; [14]Rosetta Inpharmatics, Seattle, WA 98109, USA; [15]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA; [16]Department of Nutrition and Dietetics, Harokopio University, Athens 17671, Greece
[17]These authors contributed equally to this work.
*Correspondence: jessica.a.su@gmail.com (J.L.-S.), clange@hsph.harvard.edu (C.L.)
DOI 10.1016/j.ajhg.2008.01.018. ©2008 by The American Society of Human Genetics. All rights reserved.

Table 1. Screening and Testing of SNPs for Association with Body Mass Index on the Basis of the 86,604 SNP Scan in 1,322 Subjects of the Framingham Heart Study: The Top Ten Powered SNPs

Power Rank	SNP	Allele	HWE	Informative Families	FBAT p Value
1	rs1439358	1	0.183803	22	0.88141
2	rs1875645	1	0.464993	69	0.886052
3	rs9312225	1	0.446051	70	0.848928
4	rs2134456	1	0.466619	72	0.810328
5	rs10520722	2	0.419401	65	0.592788
6	rs778373	2	0.318246	43	0.835302
7	rs293287	2	0.347734	53	0.228776
8	**rs7566605**	2	0.36771	49	0.001879
9	**rs1455832**	1	0.268188	32	0.000624
10	rs10493150	2	0.489362	73	0.662261

The bold indicates the two SNPs that have been identified as significant with the PBAT screening algorithm.

effects for complex traits can vary by age and that such an interaction can prevent replications if the age-varying character of an association is not taken into account in the selection of both the replication samples and the statistical analysis strategy. Although the importance of age-dependent genetic effects for BMI has been suggested,[15] their incorporation into current genetic analyses is not common.

The previously mentioned association between obesity, measured as BMI kg/m^2, and a common genetic variant near *INSIG2* was identified via a scan of 86,604 SNPs among 923 individuals in the FHS offspring cohort.[9] Recently, genotyping of the entire FHS sample with 1322 individuals was completed on the 100K platform, and genetic information for additional 399 subjects became available. We then reanalyzed the data, using the VanSteen Screening and Testing algorithm for family data.[16] Given the availability of longitudinal BMI measurements in the FHS, we selected the family-based association test-principal components (FBAT-PC) approach, which is particularly suited to detect both genetic main effects and time-varying effects without having to make prior assumptions about the presence of either effect.[17] Our reanalysis identified a second common genetic variant (rs1455832) at a different gene locus intronic to the gene roundabout axon guidance receptor, homolog 1 (*ROBO1*) (AHC [MIM 602430]) that achieved genome-wide significance (unadjusted p value p = 0.000624, Table 1). Because these data were longitudinal in nature (six BMI measurements per person), a profile plot that illustrates the effect of genotype on BMI over time was generated (Figure 1). This plot revealed a genetic effect between rs1455832 and BMI that varies by age (Figure 1). For the minor allele homozygote (CC), we observe a gene-age interaction associated with increased BMI that diminishes after age 45 (see Figure 1). We therefore selected as our replication hypothesis a gene-age interaction where the CC genotype is associated with increased BMI until the age of 45, at which point the effect diminishes.

We then attempted to replicate the association in eight additional cohorts totaling 13,584 subjects. These cohorts included studies of unrelated individuals and families, different ethnicities, varying ascertainment conditions, and different age ranges (childhood, adult). The objective of the replication effort was the assessment of whether the age-varying effect that was observed in the FHS was study-specific or could be generalized to the other samples. Further, assuming that the age-varying effect was unknown before the replication analysis, we assessed whether the default statistical analysis of the sample would have detected an age-dependent effect.

Material and Methods

Replication Samples

We used data on 1322 participants from the FHS offspring cohort in this analysis. Details of the study design, recruitment, and follow up are explained in more detail elsewhere.[18–20] In brief, this cohort consisted of the offspring and spouses of offspring of participants in the original FHS cohort, who have been followed prospectively since 1971, and their children. The mean and median ages at the first examination were 31.9 years and 31 years, respectively. Individuals were recruited for this study because they lived in the town of Framingham, MA, without regard to the presence of a particular disease. Obesity was documented in BMI, which is calculated by the division of weight in kilograms by height in meters squared: kg/m^2. Measurements of BMI were taken six times throughout an average time span of 23.5 years. The second examination occurred approximately 8 years after the first examination, and subsequent examinations occurred on average every 4 years thereafter. A total of 116,204 SNPs were initially genotyped, of which 86,604 were used in the association analyses after the appropriate exclusions were made.[9] The samples were genotyped with the Affymetrix GeneChip Human Mapping 100K SNP set. The genotype call rate was 99.1% with no discordance among replicate samples.

Family-based association analyses were performed with the FBAT-PC methodology implemented in PBAT on the six longitudinal measures of BMI.[17] In the initial analysis, additive, dominant, and recessive models were all considered. The FBAT-PC methodology works optimally over multiple time points when the genetic effect varies over time and in the presence of environmental correlation. When constructing the overall phenotype, the FBAT-PC methodology up-weights the time points with heritability evidence in the phenotypes and down-weights the time points with little evidence of heritability in the phenotypes.

Eight independent study samples were used to replicate the initial findings. These study populations were selected to provide information across a range of ethnicities and age ranges. Replication populations include families of children diagnosed with asthma from the Central Valley of Costa Rica (Costa Rica); families of children diagnosed with asthma from a North American clinical trial (CAMP); a child-based sample from Athens, Greece (Gene Diet Attica Investigation on childhood obesity [GENDAI]); a population-based sample from Germany (KORA_S4); a population-based sample from Iceland (Iceland); and three cohorts of lean and obese people from Europe (European Sample), the United States (US sample), and the Merck pharmaceutical company (Merck).
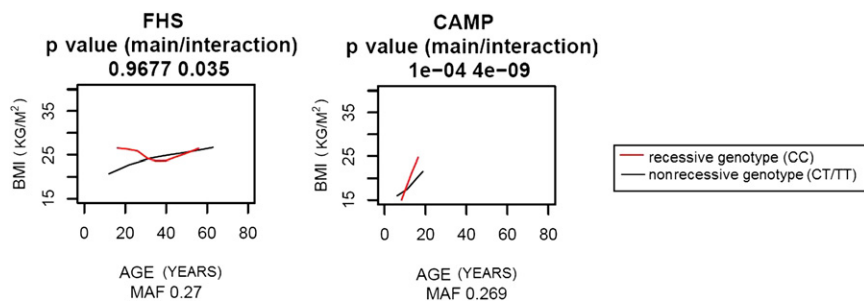
FHS
p value (main/interaction)
0.9677 0.035

MAF 0.27

CAMP
p value (main/interaction)
1e−04 4e−09

MAF 0.269

— recessive genotype (CC)
— nonrecessive genotype (CT/TT)

Figure 1. Profile Plots of Body Mass Index by Age for CC and CT or TT Genotypes in FHS and CAMP
These plots are presented for the two longitudinal studies that we have, FHS and CAMP. The left panel illustrates an initial increase in body mass index (BMI) for the CC genotype that diminishes by age 30 in the FHS. After age 8, the right panel also illustrates an increase in BMI associated with the CC genotype in CAMP. Prior to this age, our sample is limited in size, which is illustrated by the initial cross in genotypes.

Costa Rica

The rs1455832 polymorphism was genotyped in 424 nuclear families of children with asthma in the Central Valley of Costa Rica. Most of the approximately 2.85 million current residents of the Central Valley descend from approximately 4000 individuals counted in the census of 1697. Short screening questionnaires were sent to the parents of 9054 children ages 6 to 14 years enrolled in 95 schools in the Central Valley of Costa Rica from February of 2001 to March of 2005. Children were eligible for inclusion in the study of parent-child trios if they had physician-diagnosed asthma, at least two respiratory symptoms (wheezing, cough, or dyspnea) or a history of asthma attacks in the previous year, and high probability of having at least six great grandparents born in the Central Valley of Costa Rica (as determined by the study genealogist on the basis of the paternal and maternal last names of each of the child's parents).[21] Phenotypic information collected from the subjects included measures of weight and height, from which BMI was calculated. This population was not ascertained on the basis of morphometric phenotypes. Genotyping was performed with the Illumina BeadStation 500G system.[22] Genotyping completion rate was greater than 99.8% with no discordancies among replicate genotypes. Because of Mendelian inconsistencies, nine families were excluded from the analysis. In total, 409 families had sufficient SNP and BMI information to be included in the analysis. Written parental consent was obtained for participating children, for whom written assent was also obtained. The study was approved by the Institutional Review Boards of the Hospital Nacional de Niños (San José, Costa Rica) and Brigham and Women's Hospital (Boston, MA).

CAMP

The Childhood Asthma Management Program (CAMP) is a multicentered North American clinical trial evaluating the long-term effect of inhaled antiinflammatory medications in children with mild to moderate asthma. DNA samples were obtained from participating children and their parents as previously described.[23–25] In brief, children ages 5 through 12 were enrolled between December 1993 and September 1995. Otherwise healthy children were eligible for enrollment if they had a diagnosis of asthma as defined by methacholine hyperreactivity ($FEV_1PC_{20} < 12.5$ mg/ml) and at least one of the following criteria for 6 months in the year prior to enrollment: (1) asthma symptoms at least two times per week, (2) at least two usages per week of an inhaled bronchodilator, and (3) use of daily asthma medication. Measures of height and weight were collected at baseline and throughout the study. Because chronic inhaled corticosteroid use (one of the three treatment arms in the CAMP trial) is known to influence height (and in turn BMI),[25] we restricted the statistical analysis to the baseline measure of BMI when we could not appropriately control for the treatment arm assignment in another way. When the treatment arm could be adjusted for properly, we used all repeated measurements. A total of 408 individuals of European descent with BMI and SNP information were used in the analysis. Genotyping was performed with the Sequenom system with a genotyping completion rate of 98%.[26] Informed assent and consent were obtained from the study participants and their parents to collect DNA for genetic studies. The Institutional Review Board of the Brigham and Women's Hospital, as well as those of the other CAMP study centers, approved the study.

GENDAI

DNA samples were obtained from 1020 school-aged children in Athens, Greece, as previously described.[27] The study group was composed of children between 9 and 14 years of age from 40 of 593 schools in selected areas of the Athens Attica region. Enrollment began in November 2005 and continued through 2006. The children were each examined by a physician and surveyed through parent questionnaires. Exclusions included lack of interest, refusal, chronic or acute illness, absence from school, relocation, or incomplete data. Data collected include anthropometrics, a fasting blood sample, 24 hr dietary and physical activity recall, food frequency questionnaire, and questionnaires on meal patterns, lifestyle, personal, and family characteristics. Genotyping was completed by an allele-specific primer extension of amplified products with detection by matrix-assisted laser desorption ionization time-of-flight mass spectroscopy with a Sequenom platform. Genotyping call rate was 99.6% on 796 samples with adequate DNA concentration, showing no consensus errors. The IRB of Children's Hospital Boston and of Harokopio University and the Greek Ministry of Education approved the study. Subjects with poor genotyping quality were excluded. In total, 796 subjects were included in the statistical analysis.

KORA_S4

This is the fourth population-based survey (S4) of the platform "Kooperative Gesundheitsforschung in der Region Augsburg" (KORA_S4). Details about the study, the BMI distribution and its dependence upon gender and age can be found elsewhere.[28] In brief, this survey was conducted during the years 1999–2001. The sample was drawn to represent the population of the city of Augsburg and surrounding counties in a stratified fashion by 10 year age groups and by gender. Subjects were recruited via registry, were of German nationality, 25–75 years of age, 50% females, and mostly white Europeans (99%). Anthropometric measurements including weight and height were assessed according to the WHO MONICA protocol by trained staff.[28] The association was analyzed in all subjects with genotype and phenotype available,

yielding data on 3861 subjects. Genotyping was performed with MALDITOF MS (Sequenom, Mass ARRAY System, San Diego, CA).[29] The genotype call rate was 95.84% with no discordancies among replicate samples. All study participants gave informed written consent according to the ethics committee of the Bavarian Medical Association and every attempt was made to ensure anonymity of the participants.

Iceland

DNA samples were obtained from a population-based group of 4865 Icelanders. Any possible relatedness between Icelanders was accounted for as described previously.[10] The study group was composed of individuals who participated in studies of the genetic etiology of cardiovascular and metabolic diseases, and the majority of these subjects were recruited as unaffected relatives of subjects or as controls and did not have any history of metabolic or cardiovascular diseases. All participants in the study signed informed consent. All personal identifiers associated with tissue samples, clinical information, and genealogy were encrypted by the Data Protection Authority (DPA), with a third-party encryption system in which DPA maintains the code.[30] The genotyping procedure has been previously described.[31] Genotype call rate was 93.72%. Ethical approval for the present study was granted by the National Bioethics Committee (NBC reference number 01-033) and the Icelandic DPA.

The p values and confidence intervals were adjusted for relatedness of the Icelandic individuals with simulations as previously described.[32] Genotypes for the causal SNP were simulated through the Icelandic Genealogy Database, based on more than 600,000 Icelanders, and the association test was performed. This procedure was repeated 50,000 times to generate an empirical distribution, from which the standard deviation under the null hypothesis of no association was generated and used to adjust both the p value and the confidence interval.

European and US Cohorts

The rs1455832 polymorphism was genotyped in obese and lean Caucasian subjects from Poland (EUROPEAN) and the United States (US) with data provided by Genomics Collaborative. Obese subjects were selected from among individuals in the 90th–97th percentile of the BMI distribution, and lean subjects were selected from the 5th–12th percentile. Such distributions were determined with information on gender, country of origin, and age by decade as previously described.[2] The obese and lean subjects were matched by gender, decade of age, and country of origin of grandparents. Genotyping completion rates were 99.0% for the European sample and 99.6% for the US sample. There were a total of 701 cases and 331 controls in the European sample and 1218 cases and 624 controls in the US sample. Ages for these samples range from 31 to 70 years. Approval for this work was obtained from the IRB at Children's Hospital Boston.

Merck

The rs1455832 polymorphism was genotyped in a case-control study in which the cases had been recruited as part of a multicenter, double-blind, randomized, placebo-controlled study. Obese patients with BMI 30-43 kg/m^2 and between the ages of 21 and 65 who met other entry criteria were eligible to participate. Only white patients were considered for the genetic study. Prior to randomization, there was a 2 week single-blind placebo run-in period. The BMI measures used in the genetic analysis were collected at the end of the run-in period and before any drug was administered as part of the trial. Genetic and phenotypic information was collected on a total of 275 cases. The controls were acquired from GHC from a databank of individuals selected and consented for scientific study. We selected and genotyped 547 lean, healthy, white controls equally split between males and females, with BMI less than 25 and age greater than 50. In addition, the controls were selected because they were free of all forms of diabetes, cardiovascular disease, and obesity, and they had no first-degree relatives with these diseases. All samples and patient data were handled in accordance with the policies and procedures approved by a Merck IRB.

Statistical Methods

The PBAT screening approach, as described previously, was implemented in the 100K analysis of the FHS data to minimize the multiple comparisons problem.[9,16] In summary, the PBAT screening and testing algorithm follows three steps: (1) Power is calculated for all SNPs with the conditional mean model, (2) the SNPs are rank ordered according to the power calculations, (3) the ten SNPs with the greatest power are selected, and (4) statistical analyses are performed on the selected subset of SNPs. We employed the PBAT screen while using FBAT-PC, which constructs one maximally heritable BMI component for each SNP from the six measures of BMI. Family-based association analysis is then applied to the univariate BMI phenotype.[17] Golden Helix PBAT was used to perform these analyses.

Replication Analyses

We performed two analyses for each replication cohort: (1) a "uniform-analysis" approach that is the same for all eight replication cohorts and tests for both a genetic main effect and a gene-age interaction and (2) a "design-specific analysis" of the genetic main effect alone. When we refer to the design-specific analysis, we mean the default analysis that would typically be performed for a given study design. For the uniform analysis, a standard linear regression adjusting for the covariates age and gender was run, with genotype and genotype-by-age interactions as the primary predictors of interest and BMI as the response variable. The linear model used for these analyses was BMI = β_0 + $\beta_1$genotype + $\beta_2$age + $\beta_3$sex + $\beta_4$age*genotype + ε, where genotype = 1 if CC and genotype = 0 otherwise. Obesity affection status was used as an additional covariate for the case-control studies. In the CAMP study, a multivariate longitudinal regression analysis was performed in which the 14 BMI measures were the dependent variable. All of the independent variables listed above were included, with the addition of an indicator variable for asthma treatment arm, because CAMP was part of a clinical trial. In all models, the genotype was modeled in a recessive fashion to replicate the initial findings. F tests were used to evaluate the significance of both the main genetic effect and the age-by-genotype interaction term.

Based on the results of this analysis, a study was considered either a replication or a nonreplication. If the parameter estimate for the gene-age interaction variable was significant (α < 0.05), associated with the recessive genotype (CC), and the parameter estimate was negative (indicating a decrease in BMI with older ages for the recessive genotype), the study was called a replication. Otherwise, it was referred to as a "nonreplication." However, this created a dilemma for the pediatric cohorts (CAMP, Costa Rica, GENDAI) that were included in this replication attempt. The hypothesis of a gene-age interaction for subjects younger than 45 years was generated on the basis of the FHS, whose youngest subjects were 16–20 years old, whereas the pediatric cohorts covered an age range of 4–18 years. Because the age ranges of the pediatric cohorts and the FHS overlapped for only 2 years and a negative estimate

for the gene-age interaction would have meant an increase in BMI with younger ages for the recessive genotype, a peak for the association with BMI must occur in the age range below 20 years. One could envision three hypotheses: (1) The peak age of association is at birth or a very early age, (2) the peak association age coincides with the ages of the pediatric cohorts, or (3) the peak association age is around 20, when the pediatric cohorts and the FHS overlap. Testing and distinguishing of all three hypotheses for the pediatric cohorts is technically challenging and requires larger sample sizes than those we have at hand. These three hypotheses have one thing in common that makes them consistent with the FHS findings: They predict higher BMI values for the recessive genotypes in the 15–20 year age range. We therefore considered a pediatric cohort to be a replication of the FHS finding if its interaction term was significant ($\alpha < 0.05$) and the interaction had a positive parameter estimate associated with the recessive genotype (CC).

A Fisher's combined p value was calculated to determine the overall evidence for an age-by-genotype interaction for all models. Fisher developed a method for combining p values from k independent tests by summing over the log of the p value and multiplying the sum by −2. This statistic follows a chi-square distribution with 2k degrees of freedom. In this calculation, one-sided p values were used to account for scenarios in which the trend tended in the opposite direction.

Standard design-based analyses were performed for each of the replicate cohorts as follows. In the two family-based studies, Costa Rica and CAMP, we performed family-based association analysis of standardized BMI with age and sex as covariates, assuming a recessive genetic model, with Golden Helix PBAT. The standardization used in these analyses was recommended by the Center for Disease Control (CDC; see CDC-Defined BMI for Adolescents and Children in the Web Resources) and accounts for the effect of age and gender on BMI throughout childhood and adolescence. We restricted our analysis to the baseline measure of BMI in CAMP because one of the three treatment arms of the trial would have to be removed because of the known association of steroid use with BMI, and this would greatly reduce the sample size and therefore the power for this study. For the three case-control studies (European, US, and Merck), a logistic regression analysis of case-control status (obese and nonobese subjects) on rs1455832 was used. We performed the logistic regression analysis assuming a recessive model with a main genetic effect and adjusting for age and sex. Because the European and US samples were matched on age and gender, the addition of the variables as covariates contributed little to the analysis. In the population-based designs (KORA, Iceland, and GENDAI), we performed linear regression analysis of BMI on rs1455832 assuming a recessive genetic model and adjusting for age and sex. The Iceland analyses were performed with log-transformed BMI. No pooled analyses of the replication samples were performed because we did not have access to all of the individual genotype data.

Results

Preliminary Association of rs1455832 in the 100K FHS Scan

As previously reported,[9] a total of 86,604 SNPs were screened for an association with BMI in the FHS. These SNPs were subsequently genotyped in an additional 399 individuals, totaling 1322 subjects. We then reanalyzed the data with the same methodology from the original paper; the results for the top ten powered SNPs are presented in Table 1. Of these, two SNPs reached genome-wide significance (unadjusted p value for rs7566605 = 0.002, unadjusted p value for rs1455832 = 0.0006). Details of the significant association between rs7566605 and obesity were described in the previous report.[9] SNP rs1455832 is also common (minor allele frequency [MAF] = 0.271) and did not deviate from Hardy-Weinberg Equilibrium (HWE) (p = 0.27). SNP rs1455832 is located in intron 1 of the *ROBO1* gene, which contains 30 exons. Of the additional 26 SNPs that were genotyped in the *ROBO1* gene, 16 SNPs had allele frequencies that resulted in a large enough sample to evaluate genetic associations (the number of informative families greater than 20). The two SNPs adjacent to rs1455832 also had significant associations with BMI (p value for rs1455824 = 0.0092, p value for rs2311350 = 0.0181); however, this is not surprising because linkage disequilibrium exists between rs1455832 and both of these SNPs ($r^2 = 0.81$). Details of the age-dependent relationship between BMI and the rs1455832 genotypes in the FHS cohort across a 50 year age range are presented in the first graph of Figure 1. In this analysis, we observed a strong age-varying genetic association in which the CC genotype was associated with increased BMI compared with the CT or TT genotypes until age 45 years, when this effect diminishes. Not surprisingly, the significance of the main genetic effect diminishes after an interaction term is added to the regression model, because this interaction term captures the age-dependent effect (Table 2).

Replication Studies in Eight Independent Samples

To validate this initial age-dependent association (i.e., age-by-genotype interaction with BMI) finding between rs1455832 and BMI, as well as assess its generalizability to other populations, we genotyped this variant and tested for evidence of association in eight independent populations. Descriptive summaries of the eight replication cohorts are provided in Table 3. Despite different ancestral histories, the allele frequency was similar across populations, ranging from 0.211 in Costa Rica to 0.289 in Iceland. Genotype distributions were in Hardy-Weinberg equilibrium in all eight populations.

For the uniform analysis across all replication studies, we performed a population-based linear regression analysis with an age-by-genotype interaction term, with a recessive model. The analysis was restricted to individuals younger than 45.

A summary of the main genetic effect and gene-age interaction for the age-varying effect on BMI is provided in Table 2. In all replication studies, with the exception of CAMP, the main genetic effect was not significant when an age-by-genotype interaction term was included in the model. In CAMP, the CC genotype was associated with an increased BMI compared with the CT or TT genotype. As illustrated in the second graph in Figure 1, this effect is not evident in the early age ranges but is apparent

Table 2. Findings in Replication Samples

Study	p Value Main Effect	Main Genetic Effect Estimate	p Value Interaction	Interaction Effect Estimate	Direction of Interaction Effect Estimate Necessary for Replication	Genotype Associated with Increased BMI at Earlier Ages	Replication
FHS	0.97	3.01	0.018	−0.073		CC	Initial Finding
Pediatric Studies							
CAMP[a]	<0.001	−7.83	2×10^{-9}	0.838	Positive	CC	Replication
Costa Rica	0.278	−5.29	0.044	0.429	Positive	CC	Replication
GENDAI	0.311	9.36	0.85	−0.865	Positive	CT or TT	Nonreplication
Adult Studies							
US	0.464	3.06	0.05	−0.025	Negative	CC	Replication
European	0.870	−1.62	0.02	−0.053	Negative	CC	Replication
Merck	0.473	4.35	0.003	−0.067	Negative	CC	Replication
KORA_S4	0.481	−2.97	0.97	0.075	Negative	CT or TT	Nonreplication
Iceland	0.265	−0.96	0.91	0.020	Negative	CT or TT	Nonreplication

Standard linear regression analysis for all studies including a main genetic effect term and a genotype-by-age interaction term.
[a] Included treatment arm as a covariate.

after age 10. This graph illustrates the relationship between BMI and the rs1455832 genotypes in the CAMP cohort across a 10 year age range, where 14 repeated measures of BMI were used in a multivariate regression model for each individual.

Replications of the FHS finding must meet three criteria: (1) The interaction term must be significant ($\alpha = 0.05$), (2) the overall associated genotype must be consistent with the FHS (i.e., during the time frame in which we hypothesize an effect, the CC genotype is associated with an increased BMI), and (3) the interaction parameter estimate must be positive for pediatric cohorts or negative for adult cohorts. Five of the eight cohorts met these criteria: US, European, Merck, CAMP, and Costa Rica. Significant age-by-genotype interactions were observed in three adult cohorts whose age ranges match the FHS: US, European, and Merck. In these studies, the parameter estimate for the interaction variable is slightly negative, which is in the same direction and approximately the same magnitude of the effect observed in the FHS. This is to be expected because their age ranges are similar to those of the initial study. Of the pediatric cohorts, CAMP and Costa Rica replicated the FHS findings. As expected for very young ages, under the alternative hypothesis of an age-dependent genetic effect, the interaction parameter estimate for both of these studies was positive, which is in the opposite direction from the adult cohorts. For all five replications, the CC genotype was associated with an increased BMI ($p < 0.05$) during the time frame of interest. In sum, these studies reflect a stronger increase in BMI for the recessive genotype at younger ages for the pediatric cohorts and the diminishing effect of the recessive genotype in the adult cohorts. As discussed above, these five studies are considered valid replications of the initial finding under the alternative hypothesis of an age-dependent genetic effect.

Despite the lack of evidence for interaction in the Iceland, KORA_S4, and GENDAI cohorts (Iceland p = 0.910, KORA_S4 p = 0.968, GENDAI p = 0.851), the one-sided Fisher's combined p value for the age-dependent relationship in all replicate samples was highly significant (p = 3.92×10^{-9}). For the pediatric and adult cohorts separately, the combined p values were 2.21×10^{-8} and

Table 3. Descriptive Statistics on Replication Cohorts

Study	Design	Sample Size	Sample Size ≤ 50 Years	HWE p Value[a]	MAF[b]	Average Age (Range)	Average BMI (Range)	Percent Male
Costa Rica	Family based	409 Trios	409	0.99	0.211	9.1 (6.0 – 14.2)	17.9 (11.3–32.8)	62.6
CAMP	Family based	408 Trios	408	0.95	0.269	8.7 (5.2 – 17.5)	17.8 (12.7–28.6)	62.3
GENDAI	Population based	796	796	0.89	0.240	11.1 (9.9 – 13.6)	20.2 (12.9–35.9)	52.8
KORA_S4	Population based	3861	2053	0.51	0.269	49.2 (25 – 74)	27.2 (15.8–55.1)	49.0
Iceland	Population based	4865	2161	0.92	0.289	53.3 (14 – 93)	27.3 (13.7–60.2)	47.7
European Sample	Case control	1031 (700 cases)	286	0.90	0.235	55.7 (31 – 70)	Cases: 21.4 (19.9–22.7), Controls: 31.7 (27.8–35.3)	47.0
US Sample	Case control	1842 (1218 cases)	401	0.30	0.273	57.0 (31 – 70)	Cases: 36.9 (32.3–43.3), Controls: 21.5 (19.3–23.1)	45.6
Merck	Case control	422 (181 cases)	201	0.61	0.271	52.3 (40 – 60)	Cases: 34.3 (29.8–40.9), Controls: 22.3 (12.9–24.9)	39.6

[a] Reported for parents of affected individuals in family-based studies, for control subjects in case-control studies, and for all subjects in population-based studies.
[b] Minor allele frequency.

Table 4. Findings in Replication Samples for the Main Genetic Effect: Analysis by Study Design

Study Design and Study	Test	p Value for Main Genetic Effect[a]	Associated Genotype[b]
Family-Based Studies			
Costa Rica	PBAT	0.060	CC
CAMP	PBAT	0.017	CC
Case-Control Studies			
European	Logistic regression	0.920	N/A
US	Logistic regression	0.311	N/A
Merck	Logistic regression	0.665	N/A
Population-Based Studies			
GENDAI	F test	0.679	N/A
KORA_S4	F test	0.848	N/A
Iceland	F test	0.596	N/A

The family-based studies were analyzed with a standard family-based association test. The case-control studies were analyzed with a standard logistic regression with obese or lean status as the outcome variables and genotype, age, and gender as covariates. The population-based studies were performed over the entire sample and stratified on age. A linear regression with genotype as the main predictor variable and BMI as the outcome variable was performed.

[a] p value with a recessive model.

[b] Genotype associated with increased BMI is only presented when the finding reaches nominal significance at an α level of 0.10.

0.00422, respectively. In these calculations, information from GENDAI, KORA_S4, and Iceland cohorts do not contribute positively to this calculation because the direction of the association was opposite. For the computation of the combined p value, we accounted for the effect in the opposite direction by using one-sided p values. Because Fisher's method of combining p values is sensitive to highly significant results (which we have in the CAMP sample), we recalculated the combined p value eight times, each time removing one of the replication studies. The overall significance of the result did not depend on the findings from any individual replication sample because the combined p value remained significant at an α level of 0.01 when any single replication sample was removed from the analysis. It is apparent that the results from the CAMP study are highly significant and thus highly influential on the overall findings. However, if the CAMP study were completely removed from the analysis, all results and conclusions would remain the same, with the only difference being the magnitude of the combined p value (p = 0.0037). To assure that the data from the CAMP study are not spurious, we thoroughly checked the genotyping and analysis methods used. We found the quality of these data to be excellent and therefore have kept this sample in the paper.

In addition to the uniform analysis approach we used for all studies, standard study-design-specific analyses were conducted, as well. These analyses were focused only on genetic main effects. The analyses were performed for the family-based, case-control, and population-based cohorts separately. These findings are summarized in Table 4. With Golden Helix PBAT for the family-based designs, the association between standardized BMI and rs1455832 was of borderline statistical significance (p = 0.06) in 63 informative families in Costa Rica and tended in the risk direction. This association was found to be significant in the CAMP cohort with 96 informative families (p = 0.017). Therefore, the main genetic effect was detected in both of these child and/or adolescent samples. Logistic regression analyses did not result in a significant main genetic effect for any of the case control studies. In the KORA_S4 cohort, there was no association between the genotype and BMI among all subjects (p = 0.848). In the Iceland cohort, the genetic effect was also not significant (p = 0.596). No significant association was observed between genotype and BMI with the GENDAI cohort (p = 0.679).

Discussion

We initially observed an age-varying (i.e., age-by-genotype interaction) association between BMI and rs1455832 in the FHS cohort, where the CC genotype was associated with increased BMI early in life and diminished over time. The initial finding was arrived at with the FBAT-PC approach, which is particularly suited to detect both genetic main effects and time-dependent effects. We sought to replicate the age-varying association in eight independent replication samples. The age-by-rs1455832 interaction with the same directional effect was observed in five of the eight replicate samples with statistically significant results. These replication samples each had varying demographic properties, ascertainment conditions, and study designs. Notably, the genetic association was not detectable in all populations (three of which had no significant age-by- rs1455832 interactions); however, the combined p value from all of the replication samples indicates a significant overall association. It is also important to note that, in all of the replication samples except CAMP, the genetic association effect would not have been detected had we only been testing for a main genetic effect and not an age-by-rs1455832 interaction.

One important aspect in identifying associations of interest is selecting replication samples in the age range of the effect that was initially observed. Therefore, if we are simply interested in evaluating the observed main genetic effect, samples in which all individuals are less than 30 years of age are ideal. If the cohort includes individuals younger and older than 30 years of age and an interaction term is not incorporated into the analysis, the genetic effect may be missed. In cohorts whose members are all younger than 30 years, we would expect to observe the main genetic effect. The two family-based replication samples, CAMP and Costa Rica, identified the genetic association with at least marginal significance ($\alpha < 0.10$), whether or not the time-dependent effect was considered. This is probably because both of these studies recruited children

and adolescents and are therefore well-powered to detect the association. The profile plot for the CAMP study suggests that the genetic effect is less clear at very young ages (less than 10 years). This might point toward an effect starting at or after puberty, and it might be speculated that hormone status plays a role. In contrast to the two family-based studies, the GENDAI cohort does not show the main genetic effect or the interaction. This could be attributable to the small age range in GENDAI; most subjects were either age 11 or 12, making it very difficult to detect a gene-age interaction.

We also observed that the findings are sensitive to how strongly the sample was ascertained for obesity. A big advantage of population-based studies (FHS, KORA_S4, Iceland, and GENDAI) is that their phenotypic range is not restricted through the ascertainment condition, making them ideal samples in which to test for gene-age effects. Although the CAMP and Costa Rica samples were ascertained for asthma, which has a weak association with obesity, the asthma cases are mild enough and the asthma-obesity relationship tenuous enough to conclude that neither of these samples was ascertained for obesity. Therefore FHS, KORA_S4, Iceland, GENDAI, CAMP, and Costa Rica can essentially be thought of as unbiased cohorts for BMI. When looking for an age-dependent relationship between a disease and SNP, the analysis of these cohorts is straightforward. In contrast, the analytic strategy to detect interaction with age is less clear for the Merck, European, and US samples, all of which were ascertained for obesity. When the data are ascertained by the case-control status, the phenotypic variation is reduced so substantially that the interaction effect is obscured. Therefore, case-control studies are better suited to detect main effects, but their design is limited to detect interactions.

These analyses also illustrate that longitudinal phenotypic measures provide more power to detect modest age-related effects in genetic studies. Age-dependent effects can easily be missed if only one time point is collected, because the data collection may not occur when the genetic effect is strongest. Longitudinal measures are also a powerful way to increase the power of genetic studies, which is important when the genetic effects are modest. In our analyses, the two strongest effects were observed with the FHS and CAMP cohorts, both of which had many repeated-measures of BMI at different ages. Population-based cohorts like the FHS, with longitudinal data during the proper time period, seem to be optimal for identifying age-varying genetic effects. Although it is true that such a model can be estimated with cohort or case-control data at a single time point, the statistical power of such a model to detect significant genetic associations, particularly gene-time interactions, is much higher when longitudinal data are available.

A potential cause for the observed age-related genetic effects could be period effects; however, given the relatively consistent findings across several samples collected at different time points (some of them more than 40 years apart)

and in different geographical locations (US, Europe, Iceland), it is unlikely that the age-by-genotype interaction is an artifact of a cohort effect.

ROBO1 encodes for a receptor that is a member of the neural cell adhesion molecule family and is highly conserved. High tissue expression is found in brain, cochlea, mouth, adipose, and embryonic tissue (UniGene Hs. 13640). In humans, the gene maps on chromosome 3p12, and inactivating mutations were found in lung, breast, and kidney cancer. Genetic variations within the *ROBO1* gene were reported to be associated with dyslexia.[33] We can only speculate whether *ROBO1* plays a role in obesity. *ROBO1* is expressed in adipose tissue, though its function in this tissue is completely unknown.

As important findings arise from the increasing number of genetic-association studies and subsequent replication attempts become more common, it is vital that replication attempts carefully examine specific attributes of the initial study. Among necessary considerations are possible age-dependent effects from the initial association findings. In a common disease such as obesity, population-based samples are better suited than highly ascertained studies to replicate the initial findings. In conclusion, an age-genotype interaction has caused apparent inconsistencies in the attempts to replicate the finding of an association between rs1455832 and obesity in the FHS. This highlights the pitfalls of replication in cohorts with ascertainment conditions and age ranges different than those of the original study. Population-based samples with longitudinal data have greater statistical power to detect genetic effects that vary in strength by age. Given that current genotyping technology has made genotyping affordable in large-scale population-based studies, such cohorts seem to be ideal for both the discovery of new genetic effects and their replication.

Web Resources

The URLs for data presented herein are as follows:

CDC-Defined BMI for Adolescents and Children, http://www.cdc.gov/nccdphp/dnpa/bmi/childrens_BMI/about_childrens_BMI.htm#interpreted%20the%20same%20way

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

References

1. Ioannidis, J.P., Ntzani, E.E., Trikalinos, T.A., and Contopoulos-Ioannidis, D.G. (2001). Replication validity of genetic association studies. Nat. Genet. 29, 306–309.

2. Lyon, H.N., Florez, J.C., Bersaglieri, T., Saxena, R., Winckler, W., Almgren, P., Lindblad, U., Tuomi, T., Gaudet, D., Zhu, X., et al. (2006). Common variants in the ENPP1 gene are not reproducibly associated with diabetes or obesity. Diabetes 55, 3180–3184.

3. Heo, M., Leibel, R.L., Boyer, B.B., Chung, W.K., Koulu, M., Karvonen, M.K., Pesonen, U., Rissanen, A., Laakso, M., Uusitupa, M.I., et al. (2001). Pooling analysis of genetic data: The association of leptin receptor (LEPR) polymorphisms with variables related to human adiposity. Genetics 159, 1163–1178.

4. Heo, M., Leibel, R.L., Fontaine, K.R., Boyer, B.B., Chung, W.K., Koulu, M., Karvonen, M.K., Pesonen, U., Rissanen, A., Laakso, M., et al. (2002). A meta-analytic investigation of linkage and association of common leptin receptor (LEPR) polymorphisms with body mass index and waist circumference. Int. J. Obes. Relat. Metab. Disord. 26, 640–646.

5. Jellema, A., Zeegers, M.P., Feskens, E.J., Dagnelie, P.C., and Mensink, R.P. (2003). Gly972Arg variant in the insulin receptor substrate-1 gene and association with Type 2 diabetes: A meta-analysis of 27 studies. Diabetologia 46, 990–995.

6. Geller, F., Reichwald, K., Dempfle, A., Illig, T., Vollmert, C., Herpertz, S., Siffert, W., Platzer, M., Hess, C., Gudermann, T., et al. (2004). Melanocortin-4 receptor gene variant I103 is negatively associated with obesity. Am. J. Hum. Genet. 74, 572–581.

7. Swarbrick, M.M., Waldenmaier, B., Pennacchio, L.A., Lind, D.L., Cavazos, M.M., Geller, F., Merriman, R., Ustaszewska, A., Malloy, M., Scherag, A., et al. (2005). Lack of support for the association between GAD2 polymorphisms and severe human obesity. PLoS Biol. 3, e315.

8. Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A., et al. (2006). A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314, 1461–1463.

9. Herbert, A., Gerry, N.P., McQueen, M.B., Heid, I.M., Pfeufer, A., Illig, T., Wichmann, H.E., Meitinger, T., Hunter, D., Hu, F.B., et al. (2006). A common genetic variant is associated with adult and childhood obesity. Science 312, 279–283.

10. Lyon, H.N., Emilsson, V., Hinney, A., Heid, I.M., Lasky-Su, J., Zhu, X., Thorleifsson, G., Gunnarsdottir, S., Walters, G.B., Thorsteinsdottir, U., et al. (2007). The association of a SNP upstream of INSIG2 with body mass index is reproduced in several but not all cohorts. PLoS Genet 3, e61.

11. Rosskopf, D., Bornhorst, A., Rimmbach, C., Schwahn, C., Kayser, A., Kruger, A., Tessmann, G., Geissler, I., Kroemer, H.K., and Volzke, H. (2007). Comment on "A common genetic variant is associated with adult and childhood obesity". Science 315, 187, author reply 187.

12. Loos, R.J., Barroso, I., O'Rahilly, S., and Wareham, N.J. (2007). Comment on "A common genetic variant is associated with adult and childhood obesity". Science 315, 187, author reply 187.

13. Dina, C., Meyre, D., Samson, C., Tichet, J., Marre, M., Jouret, B., Charles, M.A., Balkau, B., and Froguel, P. (2007). Comment on "A common genetic variant is associated with adult and childhood obesity". Science 315, 187, author reply 187.

14. Hall, D.H., Rahman, T., Avery, P.J., and Keavney, B. (2006). INSIG-2 promoter polymorphism and obesity related phenotypes: Association study in 1428 members of 248 families. BMC Med. Genet. 7, 83.

15. Gorlova, O.Y., Amos, C.I., Wang, N.W., Shete, S., Turner, S.T., and Boerwinkle, E. (2003). Genetic linkage and imprinting effects on body mass index in children and young adults. Eur. J. Hum. Genet. 11, 425–432.

16. Van Steen, K., McQueen, M.B., Herbert, A., Raby, B., Lyon, H., Demeo, D.L., Murphy, A., Su, J., Datta, S., Rosenow, C., et al. (2005). Genomic screening and replication using the same data set in family-based association testing. Nat. Genet. 37, 683–691.

17. Lange, C., Van Steen, K., Andrew, T., Lyon, H., DeMeo, D.L., Raby, B., Murphy, A., Silverman, E.K., MacGregor, A., Weiss, S.T., and Laird, N.M. (2004). A family-based association test for repeatedly measured quantitative traits adjusting for unknown environmental and/or polygenic effects. Stat. Appl. Genet. Mol. Biol. 3, Article17.

18. Kannel, W.B. (2000). Review of recent Framingham study hypertension research. Curr. Hypertens. Rep. *2*, 239–240.

19. Feinleib, M., Kannel, W.B., Garrison, R.J., McNamara, P.M., and Castelli, W.P. (1975). The Framingham Offspring Study. Design and preliminary data. Prev. Med. *4*, 518–525.

20. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J., and Castelli, W.P. (1979). An investigation of coronary heart disease in families. The Framingham offspring study. Am. J. Epidemiol. *110*, 281–290.

21. Escamilla, M.A., Spesny, M., Reus, V.I., Gallegos, A., Meza, L., Molina, J., Sandkuijl, L.A., Fournier, E., Leon, P.E., Smith, L.B., and Freimer, N.B. (1996). Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. Am. J. Med. Genet. *67*, 244–253.

22. Oliphant, A., Barker, D.L., Stuelpnagel, J.R., and Chee, M.S. (2002). BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. Biotechniques Suppl. *56–8*, 60–61.

23. Lyon, H., Lange, C., Lake, S., Silverman, E.K., Randolph, A.G., Kwiatkowski, D., Raby, B.A., Lazarus, R., Weiland, K.M., Laird, N., and Weiss, S.T. (2004). IL10 gene polymorphisms are associated with asthma phenotypes in children. Genet. Epidemiol. *26*, 155–165.

24. The Childhood Asthma Management Program Research Group (1999). The Childhood Asthma Management Program (CAMP): Design, rationale, and methods. Control. Clin. Trials *20*, 91–120.

25. The Childhood Asthma Management Program Research Group (2000). Long-term effects of budesonide or nedocromil in children with asthma. N. Engl. J. Med. *343*, 1054–1063.

26. Tang, K., Fu, D., Kotter, S., Cotter, R.J., Cantor, C.R., and Koster, H. (1995). Matrix-assisted laser desorption/ionization mass spectrometry of immobilized duplex DNA probes. Nucleic Acids Res. *23*, 3126–3131.

27. Papoutsakis, C., Vidra, N.V., Hatzopoulou, I., Tzirkalli, M., Farmaki, A.E., Evagelidaki, E., Kapravelou, G., Kontele, I.G., Skenderi, K.P., Yannakoulia, M., and Dedoussis, G.V. (2007). The Gene-Diet Attica investigation on childhood obesity (GENDAI): Overview of the study design. Clin. Chem. Lab. Med. *45*, 309–315.

28. Heid, I.M., Vollmert, C., Hinney, A., Doring, A., Geller, F., Lowel, H., Wichmann, H.E., Illig, T., Hebebrand, J., and Kronenberg, F. (2005). Association of the 103I MC4R allele with decreased body mass in 7937 participants of two population based surveys. J. Med. Genet. *42*, e21.

29. Buetow, K.H., Edmonson, M., MacDonald, R., Clifford, R., Yip, P., Kelley, J., Little, D.P., Strasberg, R., Koester, H., Cantor, C.R., and Braun, A. (2001). High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Proc. Natl. Acad. Sci. USA *98*, 581–584.

30. Gulcher, J.R., and Stefansson, K. (2000). The Icelandic Healthcare Database and informed consent. N. Engl. J. Med. *342*, 1827–1830.

31. Kutyavin, I.V., Milesi, D., Belousov, Y., Podyminogin, M., Vorobiev, A., Gorn, V., Lukhtanov, E.A., Vermeulen, N.M., and Mahoney, W. (2006). A novel endonuclease IV post-PCR genotyping system. Nucleic Acids Res. *34*, e128.

32. Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A., Gudnadottir, V.G., et al. (2005). A common inversion under selection in Europeans. Nat. Genet. *37*, 129–137.

33. Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kaariainen, H., and Kere, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet *1*, e50.