

Genomic predictors of the maximal O₂ uptake response to standardized exercise training programs

Claude Bouchard,¹ Mark A. Sarzynski,¹ Treva K. Rice,² William E. Kraus,³ Timothy S. Church,⁴ Yun Ju Sung,² D. C. Rao,² and Tuomo Rankinen¹

¹Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, Louisiana; ²Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri; ³Department of Cardiovascular Medicine, Duke University School of Medicine, Durham, North Carolina; and ⁴Laboratory of Preventive Medicine Research, Pennington Biomedical Research Center, Baton Rouge, Louisiana

Submitted 20 August 2010; accepted in final form 17 December 2010

Bouchard C, Sarzynski MA, Rice TK, Kraus WE, Church TS, Sung YJ, Rao DC, Rankinen T. Genomic predictors of the maximal O₂ uptake response to standardized exercise training programs. *J Appl Physiol* 110: 1160–1170, 2011. First published December 23, 2010; doi:10.1152/jappphysiol.00973.2010.—Low cardiorespiratory fitness is a powerful predictor of morbidity and cardiovascular mortality. In 473 sedentary adults, all whites, from 99 families of the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study, the heritability of gains in maximal O₂ uptake ($\dot{V}O_{2max}$) after exposure to a standardized 20-wk exercise program was estimated at 47%. A genome-wide association study based on 324,611 single-nucleotide polymorphisms (SNPs) was undertaken to identify SNPs associated with improvements in $\dot{V}O_{2max}$. Based on single-SNP analysis, 39 SNPs were associated with the gains with $P < 1.5 \times 10^{-4}$. Stepwise multiple regression analysis of the 39 SNPs identified a panel of 21 SNPs that accounted for 49% of the variance in $\dot{V}O_{2max}$ trainability. Subjects who carried ≤ 9 favorable alleles at these 21 SNPs improved their $\dot{V}O_{2max}$ by 221 ml/min, whereas those who carried ≥ 19 of these alleles gained, on average, 604 ml/min. The strongest association was with rs6552828, located in the acyl-CoA synthase long-chain member 1 (*ACSL1*) gene, which accounted by itself for $\sim 6\%$ of the training response of $\dot{V}O_{2max}$. The genes nearest to the SNPs that were the strongest predictors were PR domain-containing 1 with ZNF domain (*PRDMI*); glutamate receptor, ionotropic, *N*-methyl-D-aspartate 3A (*GRIN3A*); K⁺ channel, voltage gated, subfamily H, member 8 (*KCNH8*); and zinc finger protein of the cerebellum 4 (*ZIC4*). The association with the SNP nearest to *ZIC4* was replicated in 40- to 65-yr-old, sedentary, overweight, and dyslipidemic subjects trained in Studies of a Targeted Risk Reduction Intervention Through Defined Exercise (STRRIDE; $n = 183$). Two SNPs were replicated in sedentary obese white women exercise trained in the Dose Response to Exercise (DREW) study ($n = 112$): rs1956197 near dishevelled associated activator of morphogenesis 1 (*DAAMI*) and rs17117533 in the vicinity of neccin (*NDN*). The association of SNPs rs884736 in the calmodulin-binding transcription activator 1 (*CAMTA1*) locus and rs17581162 ~ 68 kb upstream from regulator of G protein signaling 18 (*RGS18*) with the gains in $\dot{V}O_{2max}$ in HERITAGE whites were replicated in HERITAGE blacks ($n = 247$). These genomic predictors of the response of $\dot{V}O_{2max}$ to regular exercise provide new targets for the study of the biology of fitness and its adaptation to regular exercise. Large-scale replication studies are warranted.

endurance training; trainability; human variation; high and low responders

LOW CARDIORESPIRATORY FITNESS, lack of exercise, and spending a considerable amount of time in sedentary activities, such as

Address for reprint requests and other correspondence: C. Bouchard, Human Genomics Laboratory, Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808 (e-mail: claud.bouchard@pbr.edu).

sitting for prolonged periods, are all associated with less favorable cardiovascular and diabetes risk profiles, increased morbidities, and higher mortality from all causes and from cardiovascular diseases (3, 6, 18, 30, 35, 40, 42). To alleviate the health burden associated with sedentary behavior and poor fitness, public health recommendations are that adults be physically active at moderate intensity most days of the week (6, 35, 42). The assumption is that as people become more active, they increase cardiorespiratory fitness, thereby reducing the risk of disease and premature death. This approach has been validated in several studies (2, 21). Cardiorespiratory fitness levels are also more strongly related to health outcomes than are general physical activity levels (17, 23).

Maximal O₂ uptake ($\dot{V}O_{2max}$), the maximal amount of O₂ per unit of time that can be delivered to peripheral organs, including skeletal muscle, where it is used to sustain muscular contraction at peak exercise, is considered the gold standard measure of cardiorespiratory fitness. $\dot{V}O_{2max}$ is characterized by wide interindividual differences even among sedentary adults. For instance, in the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study, the heritability of $\dot{V}O_{2max}$ in the untrained state adjusted for age, sex, body mass, and body composition was estimated at $\sim 50\%$ (7), a level comparable with what has been observed in other family studies and in sets of identical and fraternal twins (11). However, little is known about the genes and DNA sequence variants that account for this genetic effect (12). Regular exercise is the most effective way to augment cardiorespiratory fitness, as evidenced by an increase in $\dot{V}O_{2max}$ or an improved tolerance to a given absolute level of submaximal exercise.

There are considerable individual differences in $\dot{V}O_{2max}$ responses to exercise training. In the HERITAGE Family Study, 473 adult whites from 99 nuclear families completed a fully standardized 20-wk endurance training program. The training program induced an average $\dot{V}O_{2max}$ increase of ~ 400 ml O₂/min with a SD of ~ 200 ml/min and a range from -114 to $+1,097$ ml/min in HERITAGE whites (8). The heritability estimate of the $\dot{V}O_{2max}$ training response was 47%. The genes and sequence variants responsible for this substantial genetic effect remained unidentified.

Here, we report the results of an investigation to identify the genetic variants associated with gains in $\dot{V}O_{2max}$ using the resources of the subsample of whites in the HERITAGE Family Study. A total of 324,611 single-nucleotide polymorphisms (SNPs) were genotyped for this purpose and subjected to extensive quality control. The most significant SNPs were tested for replication in the subsample of blacks from the

HERITAGE Study, women in the Dose Response to Exercise (DREW) Study, and men and women in the Studies of a Targeted Risk Reduction Intervention Through Defined Exercise (STRRIDE), who were all exposed to different but standardized and supervised exercise training programs. This is the first genome-wide association study (GWAS) of changes in response to exercise training. Our assumption was that it would be easier to identify SNPs and genes using an intervention design because of the low error variance for the phenotypes of interest and the low probability of confounders playing a major role compared with a typical GWAS based on cross-sectional, observational data.

METHODS

Studies Used in the GWAS

The three studies upon which the present report is based are described succinctly here. Additional information on each can be found in other publications (5, 14, 22).

HERITAGE Family Study. The sample, study design, and exercise training protocol of the HERITAGE Family Study have been described elsewhere (5).

Briefly, 834 subjects from 218 families were recruited to participate in an endurance exercise training study and have measurements of baseline $\dot{V}O_{2\max}$ taken. Among them, 483 adults from 99 families of Caucasian descent were defined as completers (5). Likewise, 259 blacks from 105 families or sibships completed the training and testing requirements. Parents were 65 yr of age or less, whereas offspring ranged in age from 17 to 41 yr old. Participants were sedentary at baseline, normotensive or mildly hypertensive (<160/100 mmHg), and did not take medications for hypertension, diabetes, or dyslipidemia (5). The study protocol was approved by the Institutional Review Boards (IRBs) at each of the five participating centers of the HERITAGE Family Study consortium. Written informed consent was obtained from each participant.

Each subject in the HERITAGE Family Study exercised 3 times/wk for 20 wk on cycle ergometers. The intensity of the training was customized for each individual based on heart rate (HR) and $\dot{V}O_2$ measurements taken at the baseline test. Details of the exercise training protocol can be found elsewhere (5). Briefly, subjects trained at the HR associated with 55% of baseline $\dot{V}O_{2\max}$ for 30 min/session for the first 2 wk. The duration and intensity were gradually increased every 2 wk until 50 min and 75% of the HR associated with baseline $\dot{V}O_{2\max}$ were reached. This level was maintained for the final 6 wk of training. All training was performed on Universal Aerobicycles (Cedar Rapids, IA), and power output was controlled by direct HR monitoring using the Universal Gym Mednet (Cedar Rapids, IA) computerized system. The protocol was standardized across all clinical centers and was supervised to ensure that the equipment was working properly and that the participants were compliant with the protocol.

Two maximal exercise tests to measure $\dot{V}O_{2\max}$ were performed on 2 separate days at baseline and again on 2 separate days after training on a SensorMedics 800S (Yorba Linda, CA) cycle ergometer and a SensorMedics 2900 metabolic measurement cart (38). The tests were conducted at about the same time of day, with at least 48 h between the two tests. In the first test, subjects exercised at a power output of 50 W for 3 min followed by increases of 25 W each 2 min until volitional exhaustion. For older, smaller, or less fit individuals, the test was started at 40 W, with increases of 10–20 W each 2 min thereafter. In the second test, subjects exercised for ~10 min at an absolute (50 W) and a relative power output equivalent to 60% $\dot{V}O_{2\max}$. They then exercised for 3 min at a relative power output that was 80% of their $\dot{V}O_{2\max}$, after which resistance was increased to the highest power output attained in the first maximal test. If the subjects were able to

pedal after 2 min, power output was increased each 2 min thereafter until they reached volitional fatigue. The average $\dot{V}O_{2\max}$ from these two sets was taken as the $\dot{V}O_{2\max}$ for that subject and used in analyses if both values were within 5% of each other. If they differed by >5%, the higher $\dot{V}O_{2\max}$ was used.

DREW exercise training program. A complete description of the design and methods (31) and details of the study participants (14) in the DREW Study have been previously published. In brief, the study was a randomized, dose-response exercise trial with sedentary, postmenopausal overweight or obese women assigned to either a nonexercise control group or to exercising groups that expended either 4, 8, or 12 kcal·kg⁻¹·wk⁻¹. The study was originally reviewed annually by The Cooper Institute and subsequently approved by the IRB of the Pennington Biomedical Research Center for continued analysis. Before participation, all participants signed a written informed consent document outlining the procedures involved in the DREW Study.

Exercising women participated in three or four (for the higher volume arm) training sessions each week for 6 mo with training intensity at the HR associated with 50% of each woman's $\dot{V}O_{2\max}$. During the first week, each group expended 4 kcal/kg. Those assigned to that level continued to expend 4 kcal·kg⁻¹·wk⁻¹ for 6 mo. All other groups increased their energy expenditure by 1 kcal·kg⁻¹·wk⁻¹ until they reached the amount prescribed for their group. All exercise sessions were performed under supervision in an exercise laboratory with complete and strict monitoring of the amount of exercise completed in each session. Women in the exercise groups alternated training sessions on semirecumbent cycle ergometers and treadmills. For the purpose of this study, we pooled the data of the women in the 8 ($n = 50$) and 12 ($n = 62$) kcal·kg⁻¹·wk⁻¹ groups only and excluded the 4 kcal·kg⁻¹·wk⁻¹ group.

Baseline and posttraining $\dot{V}O_{2\max}$ values were an average of two maximal exercise tests completed on separate days using a cycle ergometer. The maximal tests consisted of subjects exercising at 30 W for 2 min and 50 W for 4 min followed by increases of 20 W every 2 min until exhaustion, as described elsewhere (31).

STRRIDE I and II exercise training programs. A complete description of the STRRIDE study design and eligibility has been published elsewhere (22). The research protocol was reviewed and approved by the relevant IRBs of Duke University and East Carolina University. All subjects provided written informed consent. In brief, the STRRIDE I cohort was 40 to 65 yr old, sedentary, overweight or class I obese (body mass index: 25–35 kg/m²), dyslipidemic (either LDL-cholesterol of 130–190 mg/dl or HDL-cholesterol of <40 mg/dl for men or <45 mg/dl for women), and postmenopausal. Subjects in STRRIDE I were randomly assigned to one of three training groups as follows: 1) high-amount/vigorous-intensity exercise, 2,000 kcal/wk (170 min/wk), or the caloric equivalent of jogging ~20 miles/wk for a 90-kg person at 65–80% $\dot{V}O_{2\max}$; 2) low-amount/vigorous-intensity exercise, 1,200 kcal/wk (~120 min/wk), or the caloric equivalent of jogging 12 miles/wk at 65–80% $\dot{V}O_{2\max}$; and 3) low-amount/moderate-intensity exercise, 1,200 kcal/wk (170 min/wk), or the caloric equivalent of walking 12 miles/wk at 40–55% $\dot{V}O_{2\max}$. For the high-amount/vigorous-intensity group, the specific prescription was to expend 23 kcal·kg body wt⁻¹·wk⁻¹. For the two low-amount groups, energy expenditure was 14 kcal·kg body wt⁻¹·wk⁻¹. There was an initial ramp period of 2–3 mo followed by 6 mo at the appropriate exercise prescription, so that the total duration was 8–9 mo of training. Although the amount of exercise was expressed in terms of walking or jogging a certain distance to simplify the description of the exercise groups, the main exercise modalities were treadmills and elliptical trainers, with some use of cycle ergometers.

The STRRIDE II cohort was very similar to the STRRIDE I cohort with the exceptions being that the age range was 18 to 70 yr old and women did not have to be postmenopausal. Subjects in STRRIDE II were assigned to one of four exercise training groups as follows: 1) aerobic training (AT) only, 1,300 kcal/wk, with exercise performed 3 times/wk at an intensity of 65–80% $\dot{V}O_{2\max}$; 2) resistance training

(RT) only, with 3 sets of 12–15 repetitions performed 3 times/wk; 3) combined AT and RT (AT/RT), the combination of the protocols for AT and RT described above performed 3–5 times/wk; and 4) high AT, 2,200 kcal/wk, performed 3 times/wk at an intensity of 65–80% $\dot{V}O_{2\max}$. All exercise sessions were verified by direct supervision or by HR monitors that provided recorded data (Polar Electro, Woodbury, NY).

All subjects underwent $\dot{V}O_{2\max}$ tests with a 12-lead ECG and expired gas analysis on a treadmill (16). These tests were performed twice at baseline and after the exercise program was completed. Expired gases were analyzed continuously (model 2900 U, Sensor-Medics, Yorba Linda, CA; or TrueMax 2400 ParvoMedics, Sandy, UT). The protocol consisted of 2-min stages, which increased the workload by ~ 1 metabolic equivalent unit/stage. The same protocol and same metabolic cart were used before and after training in each subject. The last 40 s were averaged to determine $\dot{V}O_{2\max}$. The actual work rate, correlating to the prescribed exercise intensity, was determined during a submaximal exercise test performed on a separate day during the first 2–3 wk of exercise training.

GWAS SNP genotyping

Genomic DNA was prepared from immortalized lymphoblastoid cell lines by a commercial DNA extraction kit (Gentra Systems, Minneapolis, MN). GWAS SNPs were genotyped using Illumina HumanCNV370-Quad v3.0 BeadChips on an Illumina BeadStation 500GX platform. Genotype calls were done with Illumina GenomeStudio software, and all samples were called in the same batch to eliminate batch-to-batch variations. All GenomeStudio genotype calls with a GenTrain score of < 0.885 were checked and confirmed manually. Monomorphic SNPs and SNPs with only one heterozygote as well as SNPs with $> 30\%$ missing data were filtered out with GenomeStudio. Quality control of the GWAS SNP data confirmed all family relationships and found no evidence of DNA sample mix-ups. Only 78 SNPs (0.023%) had $> 10\%$ missing data. Minor allele frequency was $< 1\%$ for 1,301 SNPs (0.39%). Hardy-Weinberg equilibrium (HWE) test P values were $< 10^{-5}$ and 10^{-6} for 55 (0.017%) and 12 (0.0037%) SNPs, respectively. Twelve samples were genotyped in duplicate with 100% reproducibility across all the SNPs.

The 15 SNPs most strongly associated from the final regression model were selected for replication experiments. Two of these SNPs received very low Illumina GoldenGate assay design scores (rs10921078: 0.30 and rs824205: 0.21, score range: 0–1.1; values of > 0.6 indicate a satisfactory a priori likelihood for a successful assay), and we used the International HapMap Caucasian database to find replacements for them: SNP rs824205 was replaced with rs17117533 (pairwise $r^2 = 1.0$, design score = 0.99) and SNP rs10921078 was replaced with rs17581162 (pairwise $r^2 = 0.84$, design score = 0.96). DNA for the replication experiments was extracted from immortalized lymphoblastoid cell lines (HERITAGE), buffy coats (DREW and STRRIDE), and skeletal muscle biopsies (STRRIDE). SNPs were genotyped using an Illumina GoldenGate assay and Veracode technology on the BeadXpress platform. Genotype calls were done using Illumina GenomeStudio software. The average genotype call rate was 0.998 for SNPs and 99.5% for DNA samples (HERITAGE: 100%, DREW: 99.7%, and STRRIDE: 98.5%). None of the SNPs showed Mendelian errors in the HERITAGE families, and all SNPs were in HWE [tested using the exact HWE test implemented in the PEDSTATS software package (44)]. In addition, five Centre d'Etude du Polymorphisme Humain (CEPH) DNA samples included in the HapMap Phase II CEU panel (NA10851, NA10854, NA10857, NA10860, and NA10861) were genotyped in triplicate. Concordance between the replicates as well as with the SNP genotypes from the HapMap database was 100%. A 100% concordance between GoldenGate (replication study) and Infinium (GWAS) assays was confirmed by genotyping 20 HERITAGE whites with both methods.

Statistical Analyses

The $\dot{V}O_{2\max}$ training response was adjusted for age and baseline $\dot{V}O_{2\max}$ using a stepwise regression procedure. Residuals were standardized to a zero mean and unit variance within sex-by-generation subgroups as previously described (9).

Associations between the GWAS SNPs and $\dot{V}O_{2\max}$ training responses were analyzed using the MERLIN software package. The total association model of MERLIN uses a variance-component framework to combine the phenotypic mean model and estimates of additive genetic, residual genetic, and residual environmental variances from a variance-covariance matrix into a single likelihood model. The evidence of association is evaluated by maximizing the likelihoods under two conditions: the null hypothesis (L_0) restricts the additive genetic effect of the marker locus to zero ($\beta_a = 0$), whereas the alternative hypothesis does not impose any restrictions on β_a . The quantity of twice the difference of the log likelihoods between the alternative (L_1) and null hypotheses $\{2[\ln(L_1) - \ln(L_0)]\}$ is distributed as χ^2 with 1 difference in the number of parameters estimated.

Multivariable regression procedures were used to evaluate the overall contribution of the most significant GWAS SNPs on the $\dot{V}O_{2\max}$ training response. All GWAS SNPs with $P \leq 1.5 \times 10^{-4}$ were included. First, a regression model with backward elimination was used to filter out redundant SNPs [mainly due to a strong pairwise linkage disequilibrium (LD)]. The threshold for keeping the SNPs in the model was $P = 0.05$. SNPs that were retained in the final backward elimination model were then analyzed with a multivariate regression model using forward selection.

Associations between the $\dot{V}O_{2\max}$ training response and replication SNPs were analyzed using the MERLIN association model in HERITAGE blacks and with general linear models (Proc GLM in SAS version 9.1) in DREW and STRRIDE subjects. Data adjustment in HERITAGE blacks was done the same way as in HERITAGE whites. For DREW and STRRIDE, age, sex (STRRIDE only), baseline $\dot{V}O_{2\max}$, and intervention group were used as covariates in the GLM model. In these replication samples, we retained SNPs significant at the 0.05 level.

RESULTS

The distribution of $\dot{V}O_{2\max}$ changes with the HERITAGE exercise training program is shown in Fig. 1 for the whites. About 7% of subjects registered a gain of 100 ml/min or less, whereas 8% of subjects improved by 700 ml/min or more. On

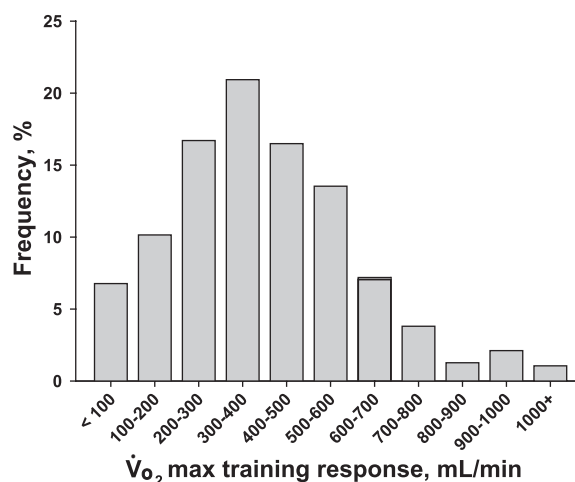


Fig. 1. Distribution of the maximal O₂ uptake ($\dot{V}O_{2\max}$) training responses in whites in the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study.

average, subjects increased $\dot{V}O_{2max}$ by ~ 400 ml O_2 /min with a SD of >200 ml/min.

An overview of single-SNP GWAS results across all 22 autosomes is shown as a Manhattan plot in Fig. 2. Altogether, 39 SNPs with minor allele frequency of 8% or more were associated with a $\dot{V}O_{2max}$ training response at the $P < 1.5 \times 10^{-4}$ significance level (Table 1); 5 SNPs showed associations at $P < 1 \times 10^{-5}$, whereas 20 SNPs [including a cluster of 6 SNPs with pairwise $r^2 > 0.95$ in the HLA complex group 22 (*HCG22*) locus] were associated at a significance level of $1.4 \times 10^{-5} < P < 9.3 \times 10^{-4}$. Full names of the gene loci closest to the SNPs shown in Table 1 and used elsewhere in this report are provided in the Supplemental Material (Supplemental Table S1).¹

The strongest evidence of association ($P = 1.31 \times 10^{-6}$) was detected with SNP rs6552828 (chromosome 4), which is located in the first intron of the acyl-coA synthase long-chain member 1 gene (*ACSL1*; also known as long-chain acyl-CoA synthetase), 715 bp and 718 bp upstream of exon 2 and the start codon, respectively. Homozygotes of the rs6552828 minor allele (A/A) had 125 ml/min (28%) and 63 ml/min (17%) lower $\dot{V}O_{2max}$ responses than the common allele homozygotes (G/G) and heterozygotes (A/G), respectively. In the single-SNP analyses, rs6552828 explained 6.1% of the variance in the response of $\dot{V}O_{2max}$.

All 39 SNPs that showed P values of $<1.5 \times 10^{-4}$ in single-SNP analyses were selected for multivariate analyses with the change in $\dot{V}O_{2max}$ as the dependent variable. First, all 39 SNPs were tested using a regression model with backward elimination to filter out redundant markers (due to pairwise LD); 21 SNPs were retained ($P < 0.05$) in the final model. These 21 SNPs were then entered into a multivariate regression model with forward selection (Table 2). The full-forward selection model with 21 SNPs explained 48.6% of the variance in the response of $\dot{V}O_{2max}$ to the exercise program: 6 SNPs each contributed at least 3%, 10 SNPs contributed between 1% and 3%, and 4 SNPs contributed $<1\%$ to the

total variance. The six SNPs most strongly associated with the gains in $\dot{V}O_{2max}$ were rs10499043, which was located 287 kb from the nearest gene, PR domain-containing 1 with ZNF domain (*PRDMI*); rs1535628, which mapped 516 kb from the gene for glutamate receptor, ionotropic, *N*-methyl-D-aspartate (NMDA) 3A (*GRIN3A*); rs4973706, which was 268 kb from the gene for K^+ channel, voltage gated, subfamily H, member 8 (*KCNH8*); rs12115454, which was 33 kb from chromosome 9 open reading frame 27 (*C9orf27*); rs6552828, which was located in the first intron of *ACSL1*; and rs11715829, which was 146 kb from zinc finger protein of the cerebellum 4 (*ZIC4*).

A summary “predictor score” was constructed using the 21 SNPs shown in Table 2. Each SNP was recoded to reflect the number of high $\dot{V}O_{2max}$ training response alleles: 0 = homozygote for the low-response allele, 1 = heterozygote, and 2 = homozygote for the high-response allele. The sum of the 21 recoded SNPs was used as the predictor score (theoretical range: 0–42). The observed predictor SNP summary score values ranged from a minimum of 7 to a maximum of 31 (Table 3). The mean increase in $\dot{V}O_{2max}$ among the subjects with a predictor SNP score of ≤ 9 (36 subjects) was 221 ml O_2 /min, whereas those with a score of ≥ 19 (52 subjects) had a 2.7 times greater (604 ml/min) $\dot{V}O_{2max}$ training response (Fig. 3).

Replication of associations with the 15 most strongly associated SNPs among these 21 SNPs was attempted using the smaller cohorts of HERITAGE blacks and DREW and STRIDE subjects. Detailed results for these 15 SNPs across all replication cohorts are shown in Supplemental Table S2. We used a P value of ≤ 0.05 to verify if associations could be replicated in these other cohorts. Subjects in DREW and STRIDE were ~ 20 yr older than the HERITAGE whites and blacks (Table 4). Baseline $\dot{V}O_{2max}$ was significantly lower in DREW subjects, with a mean of ~ 16 ml·kg⁻¹·min⁻¹. Subjects in DREW and STRIDE increased their $\dot{V}O_{2max}$ by 8.7% and 10%, respectively, in contrast to the mean increases of 16.9% and 18.9% in HERITAGE whites and blacks (Table 4).

We first considered the 6 SNPs among the 21 most strongly associated SNPs (regression model) with the exercise training

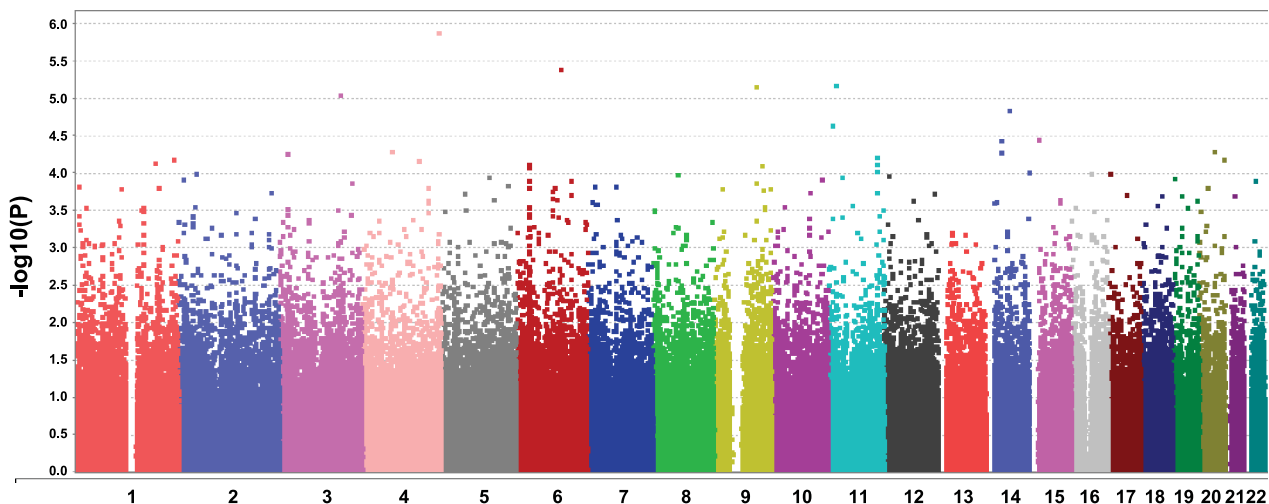


Fig. 2. A Manhattan summary plot of the $\dot{V}O_{2max}$ training response genome-wide association study (GWAS) results across 22 autosomes. Chromosomes are on the x-axis, and P values are shown on the y-axis as $-\log_{10}$ of P .

¹ Supplemental Material for this article is available at the *Journal of Applied Physiology* website.

Table 1. List of GWAS SNPs associated with $\dot{V}O_{2\max}$ training responses with $P < 1.5 \times 10^{-4}$ in HERITAGE whites

SNP	Chromosome	MAP	H^2	P Value	MAF	Gene
rs884736	1	6,937,692	3.89	1.49×10^{-4}	0.41	<i>CAMTA1</i>
rs10921078	1	190,325,645	3.96	7.17×10^{-5}	0.15	<i>RGS18</i> (68 kb)
rs7531957	1	235,856,279	3.84	6.42×10^{-5}	0.08	<i>RYR2</i>
rs2003298	2	11,003,601	3.40	1.21×10^{-4}	0.42	<i>KCNF1</i> (32 kb)
rs4952535	2	41,985,027	3.87	1.01×10^{-4}	0.41	<i>LOC400950</i> (11 kb)
rs4973706	3	18,896,776	4.09	5.31×10^{-5}	0.24	<i>KCNH8</i> (268 kb)
rs11715829	3	148,439,856	4.94	8.68×10^{-6}	0.08	<i>ZIC4</i> (146 kb)
rs2030398	3	174,488,667	3.84	1.32×10^{-4}	0.20	<i>NLGN1</i> (110 kb)
rs2272040	4	71,041,636	3.84	5.05×10^{-5}	0.13	<i>CSN1S2B</i>
rs2053896	4	137,374,246	4.03	6.62×10^{-5}	0.10	<i>LOC100289626</i> (134 kb)
rs6552828	4	185,962,410	6.14	1.31×10^{-6}	0.37	<i>ACSL1</i>
rs2176830	5	159,380,714	3.66	1.42×10^{-4}	0.13	<i>TTC1</i>
rs2523849	6	31,133,030	4.03	7.53×10^{-5}	0.17	<i>HCG22</i>
rs2523848	6	31,133,083	4.02	7.53×10^{-5}	0.17	<i>HCG22</i>
rs2428514	6	31,135,495	4.03	8.22×10^{-5}	0.15	<i>HCG22</i>
rs2517518	6	31,136,324	4.02	7.53×10^{-5}	0.17	<i>HCG22</i>
rs2523840	6	31,138,404	4.03	7.53×10^{-5}	0.17	<i>HCG22</i>
rs2517506	6	31,139,659	4.02	7.53×10^{-5}	0.17	<i>HCG22</i>
rs10499043	6	106,353,830	5.66	3.93×10^{-6}	0.13	<i>PRDM1</i> (287 kb)
rs10452621	6	132,127,094	3.87	1.23×10^{-4}	0.12	<i>ENPP3</i> (17 kb)
rs3814991	7	18,601,428	3.52	1.46×10^{-4}	0.11	<i>HDAC9</i>
rs12538806	7	70,200,777	3.57	1.47×10^{-4}	0.30	<i>WBSCR17</i> (35 kb)
rs13235325	7	70,202,943	3.57	1.47×10^{-4}	0.30	<i>WBSCR17</i> (33 kb)
rs1535628	9	104,056,570	5.26	6.81×10^{-6}	0.09	<i>GRIN3A</i> (516 kb)
rs959066	9	104,081,084	3.41	1.35×10^{-4}	0.27	<i>GRIN3A</i> (540 kb)
rs12115454	9	117,759,871	3.77	7.74×10^{-5}	0.11	<i>C9orf27</i> (33 kb)
rs2198009	11	10,360,153	4.35	2.28×10^{-5}	0.50	<i>LOC100130460</i>
rs10500872	11	20,202,299	4.60	6.49×10^{-6}	0.15	<i>DBX1</i> (64 kb)
rs353625	11	35,125,122	3.88	1.12×10^{-4}	0.32	<i>CD44</i>
rs4938561	11	118,223,695	3.61	9.29×10^{-5}	0.23	<i>CXCR5</i> (36 kb)
rs7933007	11	118,235,879	3.71	7.35×10^{-5}	0.23	<i>CXCR5</i> (24 kb)
rs12580476	12	13,435,330	3.70	1.08×10^{-4}	0.14	<i>C12orf36</i> (14 kb)
rs12896790	14	37,343,673	4.13	3.59×10^{-5}	0.09	<i>TTC6</i>
rs8018889	14	37,353,342	4.02	5.25×10^{-5}	0.09	<i>TTC6</i>
rs1956197	14	58,547,167	4.90	1.43×10^{-5}	0.16	<i>DAAMI</i> (174 kb)
rs824205	15	21,559,164	4.05	3.45×10^{-5}	0.15	<i>NDN</i> (75 kb)
rs7256163	19	6,909,134	3.83	1.13×10^{-4}	0.31	<i>EMR4</i>
rs6090314	20	61,327,997	3.82	6.48×10^{-5}	0.16	<i>BIRC7</i> and <i>YTHDF1</i> (10 kb)
rs738353	22	26,460,072	3.51	1.23×10^{-4}	0.35	<i>MNI</i> (14 kb)

Maximal O_2 uptake ($\dot{V}O_{2\max}$) training responses were adjusted for age, sex, and baseline $\dot{V}O_{2\max}$ (see METHODS). GWAS, genome-wide association studies; SNP, single-nucleotide polymorphism; HERITAGE, Health, Risk Factors, Exercise Training, and Genetics Family Study; MAP, indicates the map position of the polymorphism; MAF, minor allele frequency.

induced increase in $\dot{V}O_{2\max}$ in HERITAGE whites. Among them, rs10499043 (287 kb upstream from *PRDM1*) was not associated with the changes in $\dot{V}O_{2\max}$ in HERITAGE blacks or DREW or STRRIDE subjects. No associations were found also with rs1535628 (516 kb from *GRIN3A*), rs4973706 (268 kb from *KCNH8*), rs12115454 (33 kb from *C9orf27*), and rs6552828 (in *ACSL1*) in any of the replication cohorts. However, rs11715829 (146 kb from *ZIC4*) was associated with the gains in $\dot{V}O_{2\max}$ in STRRIDE subjects, with the major allele homozygotes gaining $\sim 30\%$ less than the minor allele carriers, just as in HERITAGE whites. The latter results are shown in Fig. 4.

Even though there was little replication with the six most strongly associated SNPs as evidenced in HERITAGE whites, there were several positive findings for the other SNPs of the genomic predictors of trainability. In HERITAGE blacks, rs884736 [in calmodulin-binding transcription activator 1 (*CAMTA1*)] was significantly associated with $\dot{V}O_{2\max}$ increases ($P = 0.032$), with a pattern identical to that of HERITAGE whites (Fig. 5, top). Similarly, for two SNPs in complete LD in the vicinity of the regulator of G protein signaling 18 gene (*RGS18*), the

common allele homozygotes gained less than the minor allele carriers both in HERITAGE blacks and whites (Fig. 5, bottom).

More replications were observed in DREW subjects. As shown in Fig. 6, rs1956197 [174 kb from the dishevelled associated activator of morphogenesis 1 gene (*DAAMI*), $P = 0.0193$] was significantly associated with the gains in $\dot{V}O_{2\max}$ with the same response patterns as in HERITAGE whites. Moreover, SNPs rs17117533 and rs824205, located ~ 75 kb upstream of the neccin gene (*NDN*), found to be significant in HERITAGE whites were also found to be associated ($P = 0.05$) with the gains in $\dot{V}O_{2\max}$ in DREW subjects (Fig. 7).

DISCUSSION

On average, the gains in $\dot{V}O_{2\max}$ in groups of subjects exposed to several months of endurance exercise were of the order of 15–25%. Most studies have not found an age or a sex effect in the relative $\dot{V}O_{2\max}$ response to training, i.e., in the gains of $\dot{V}O_{2\max}$ adjusted for the baseline value or in the percentage of improvement (8, 26, 39). However, there are

Table 2. Results of the multivariate regression model with forward selection for the $\dot{V}O_{2max}$ response to training in HERITAGE whites

SNP	Chromosome	MAP	Gene	MAF	Regression Model			
					Partial R ²	Model R ²	F Value	Probability > F
rs10499043	6	106,353,830	PRDM1 (287 kb)	0.13	0.070	0.070	34.73	<.0001
rs1535628	9	104,056,570	GRIN3A (516 kb)	0.09	0.052	0.122	27.67	<.0001
rs4973706	3	18,896,776	KCNH8 (268 kb)	0.24	0.045	0.167	25.11	<.0001
rs12115454	9	117,759,871	C9orf27 (33 kb)	0.11	0.041	0.208	23.71	<.0001
rs6552828	4	185,962,410	ACSL1	0.37	0.035	0.243	21.15	<.0001
rs11715829	3	148,439,856	ZIC4 (146 kb)	0.08	0.032	0.275	20.6	<.0001
rs884736	1	6,937,692	CAMTA1	0.41	0.029	0.304	18.82	<.0001
rs10921078	1	190,325,645	RGS18 (70 kb)	0.15	0.025	0.329	17.23	<.0001
rs6090314	20	61,327,997	BIRC7 (10 kb)	0.16	0.022	0.351	15.48	<.0001
rs10500872	11	20,202,299	DBX1 (65 kb)	0.15	0.018	0.369	13.29	0.0003
rs1956197	14	58,547,167	DAAMI (174 kb)	0.16	0.017	0.386	12.88	0.0004
rs824205	15	21,559,164	NDN (75 kb)	0.15	0.014	0.401	10.53	0.0013
rs7933007	11	118,235,879	CXCR5 (24 kb)	0.23	0.013	0.414	10.16	0.0015
rs12896790	14	37,343,673	TTC6	0.09	0.013	0.427	10.3	0.0014
rs4952535	2	41,985,027	LOC400950 (11 kb)	0.41	0.012	0.439	9.29	0.0024
rs2053896	4	137,374,246	LOC100289626 (134 kb)	0.10	0.012	0.451	9.44	0.0022
rs2198009	11	10,360,153	LOC100130460	0.50	0.008	0.459	6.57	0.0107
rs2030398	3	174,488,667	NLGN1 (110 kb)	0.20	0.008	0.467	6.61	0.0105
rs738353	22	26,460,072	MNI (14 kb)	0.35	0.009	0.476	7.83	0.0054
rs353625	11	35,125,122	CD44	0.32	0.005	0.481	4.42	0.0362
rs10452621	6	132,127,094	ENPP3 (17 kb)	0.12	0.005	0.486	4.04	0.0451

$\dot{V}O_{2max}$ training responses were adjusted for age, sex, and baseline $\dot{V}O_{2max}$ (see METHODS). Gene indicates the gene located nearest to the SNP. The distance to the gene in kb (1,000 bp) is shown in parentheses. If no distance is shown, the SNP is located within the gene locus.

considerable individual differences in the magnitude of benefits derived from a physically active lifestyle (4, 20, 26, 45). Thus, the heterogeneity in the response in $\dot{V}O_{2max}$ ranges from zero or very low gains up to >50%. In a series of intervention studies performed with identical twins, we (4, 36) found that the between-identical twin pairs variance in response to regular exercise was from two to nine times larger than the within-pair variance for cardiorespiratory fitness, hemodynamic, and metabolic phenotypes. These observations were amplified and confirmed by the HERITAGE Family Study (8, 10, 37).

Table 3. Distribution of the $\dot{V}O_{2max}$ response (in mL/min) GWAS predictor SNP summary scores in HERITAGE whites

Score	No. of Subjects	Frequency	Mean	SE
7	4	0.8	258.8	24.1
8	13	2.62	188.7	38.5
9	19	3.82	204.7	34.0
10	25	5.43	204.5	34.7
11	28	6.44	269.4	31.2
12	54	11.27	302.2	22.9
13	57	12.07	349.5	20.6
14	48	10.46	386.2	22.3
15	60	12.88	393.9	24.8
16	51	11.47	480.8	26.8
17	28	6.04	594.5	35.2
18	27	5.84	515.3	31.6
19	25	5.23	503.1	30.6
20	12	2.41	505.5	31.0
21	6	1.21	827.6	117.0
22	1	0.2	557	
23	1	0.2	747	
24	2	0.4	727.5	189.5
28	2	0.4	800.4	56.6
29	1	0.2	894.7	
31	2	0.4	1062	35.5

The summary score was calculated using the 21 SNPs shown in Table 2.

Why do we have such human variation in human adaptive capacity? From a prior report (8), we know that the heritability of the response variance adjusted for age, sex, and baseline $\dot{V}O_{2max}$ is in the range of 45–50%. The present study showed that almost 50% of the variance in age, sex, and baseline level-adjusted $\dot{V}O_{2max}$ response to an exercise program can be predicted with a panel of 21 SNPs in 473 whites from the HERITAGE Family Study, indicating a high degree of concordance between the heritability level and the variance accounted for by the panel of 21 SNPs. Interestingly, each of six of these SNPs accounted for at least 3% of the variance in the $\dot{V}O_{2max}$ response. This observation is in sharp contrast with the numerous reports on the effect size of SNPs associated with

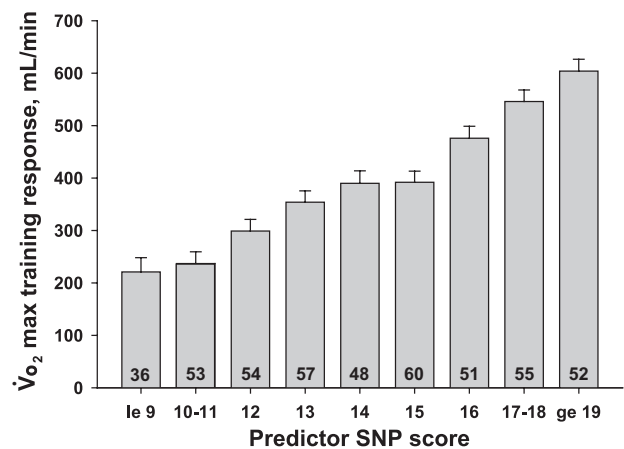


Fig. 3. Age, sex, and baseline $\dot{V}O_{2max}$ -adjusted $\dot{V}O_{2max}$ training responses across nine GWAS predictor single-nucleotide polymorphism (SNP) score categories in HERITAGE whites. Numbers of subjects within each SNP score category are indicated inside each histogram bar. Le, “less or equal to;” ge, “greater or equal to.”

Table 4. Descriptive data, including baseline $\dot{V}O_{2\max}$ and its response to training, for HERITAGE whites and blacks as well as DREW and STRRIDE cohorts

	HERITAGE		DREW	STRRIDE
	Whites	Blacks		
No. of subjects	470	247	112	183
Age, yr	35.8 (14.5)	33.6 (11.5)	57.9 (6.5)	51.1 (8.9)
Baseline body mass index, kg/m ²	25.8 (4.9)	27.8 (5.8)	31.5 (3.9)	29.9 (2.9)
Baseline $\dot{V}O_{2\max}$, ml/min	2,458 (740)	2,086 (629)	1,312 (240)	2,508 (736)
Baseline $\dot{V}O_{2\max}$, ml·kg ⁻¹ ·min ⁻¹	33.2 (8.9)	27.3 (7.3)	15.8 (2.5)	28.2 (6.0)
$\dot{V}O_{2\max}$ response, ml/min	395 (215)	362 (171)	108 (132)	253 (270)
$\dot{V}O_{2\max}$ response, %	16.9 (9.0)	18.9 (10.3)	8.7 (10.5)	10.0 (10.0)

Values are given as means (SD). The $\dot{V}O_{2\max}$ response = posttraining $\dot{V}O_{2\max}$ minus baseline $\dot{V}O_{2\max}$ (a positive value represents an improvement in $\dot{V}O_{2\max}$). DREW, Dose Response to Exercise Study; STRRIDE, Studies of a Targeted Risk Reduction Intervention Through Defined Exercise.

common chronic diseases or normal human quantitative traits as reported in the GWAS literature over recent years, in which the effect sizes of SNPs rarely exceed 1% (25, 28, 43, 46). We speculate that the difference between the HERITAGE Family Study and the other GWAS reports arises primarily from the fact that $\dot{V}O_{2\max}$ was measured quite precisely (twice before the exercise program and twice posttraining) in the HERITAGE Family Study, that $\dot{V}O_{2\max}$ is a clear, well-defined physiological trait, and that the changes were generated by a rigorous intervention with all completers having achieved a very high degree of compliance. In contrast, the genome-wide explorations reported to date (e.g., obesity and type 2 diabetes) have been based largely on observational data, with subjects measured or reporting their status at one time point and in cohorts in which young or middle-aged control subjects may carry the risk alleles but have not yet expressed the disease. However, there is also a probability that some of the significant SNPs are false positives given the modest sample size of HERITAGE whites.

SNP rs6552828 encoded in *ACSL1* accounted for 3.5% of the $\dot{V}O_{2\max}$ trainability variance in the multivariable regression model (Table 2) but 7% in the individual SNP analysis (Table 1). Besides the quantitative evidence of a strong association, *ACSL1* is a serious candidate gene because of its potential relevance to the adaptation to regular exercise. *ACSL1* contributes most of the acyl-CoA synthetase activity in adipose

tissue and is highly expressed in skeletal muscle and in the heart. *ACSL1* encodes an enzyme that is involved in the partitioning of fatty acids and the control of lipid influx but, even more importantly, the control of lipid efflux (24). It plays a key role in normal insulin metabolism and fatty acid-induced insulin resistance. A SNP in *ACSL1* has been associated with fasting glucose, insulin resistance, and the presence of metabolic syndrome, particularly in the presence of a high-fat diet (34). In mice, heart-specific overexpression of *ACSL1* increases triglyceride accumulation in cardiac myocytes, which leads to cardiac hypertrophy, left ventricular dysfunction, and premature death (13).

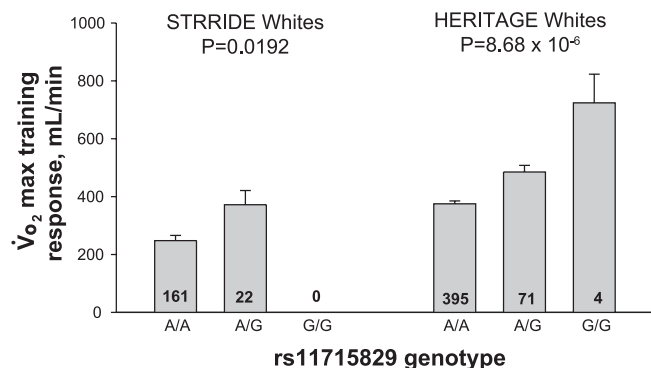


Fig. 4. Association of SNP rs11715829 with age, sex, and baseline $\dot{V}O_{2\max}$ -adjusted $\dot{V}O_{2\max}$ training responses in Studies of a Targeted Risk Reduction Intervention in Defined Exercise (STRRIDE) whites (left) and HERITAGE whites (right). P values for the main effect of genotype in each cohort are shown at the top of each graph. Numbers of subjects with each genotype are indicated inside each histogram bar (no STRRIDE whites had the G/G genotype).

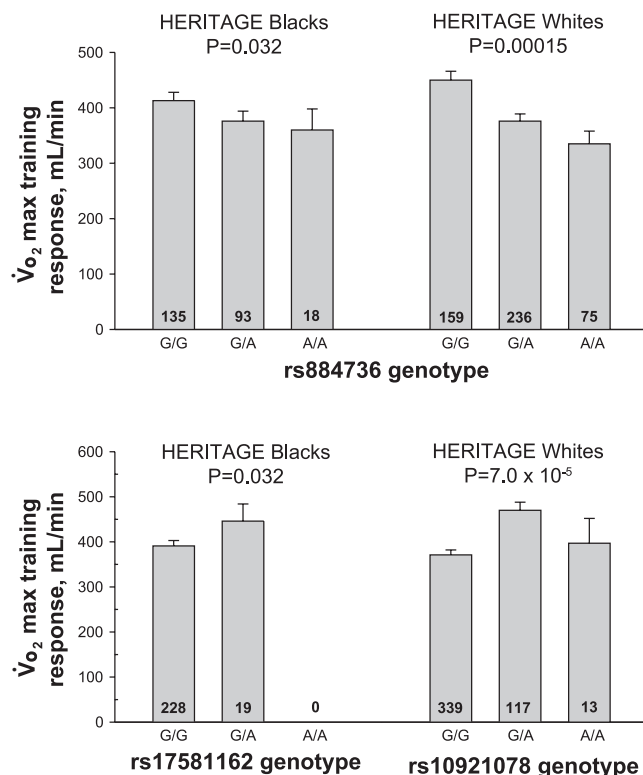


Fig. 5. Association of calmodulin-binding transcription activator 1 (*CAMTA1*) SNP rs884736 (top) and regulator of G protein signaling 18 (*RGS18*) SNPs rs17581162 and rs10921078 (bottom) with age, sex, and baseline $\dot{V}O_{2\max}$ -adjusted $\dot{V}O_{2\max}$ training responses in HERITAGE blacks (left) and whites (right). P values for the main effect of genotype in each cohort are shown at the top of each graph. Numbers of subjects with each genotype are indicated inside each histogram bar. GWAS SNP rs10921078 in HERITAGE whites was replaced with rs17581162 in the replication experiments (see METHODS).

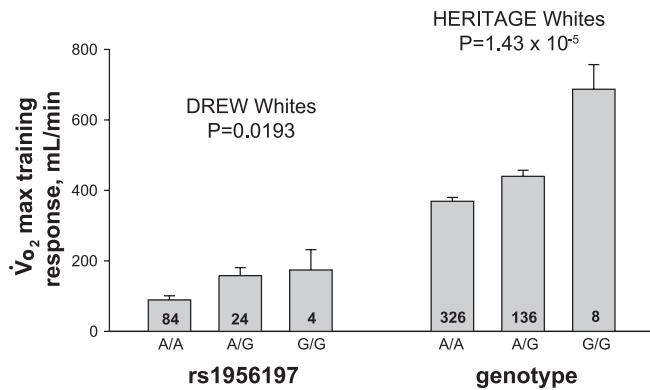


Fig. 6. Association of SNP rs1956197 with age, sex, and baseline $\dot{V}O_{2max}$ -adjusted $\dot{V}O_{2max}$ training responses in Dose Response to Exercise (DREW) whites (left) and HERITAGE whites (right). P values for the main effect of genotype in each cohort are shown at the top of each graph. Numbers of subjects with each genotype are indicated inside each histogram bar.

In the multivariable regression analysis of association with the $\dot{V}O_{2max}$ response, a marker 287 kb from *PRDM1* was the strongest predictor, accounting for 7% of the variance (Table 2). *PRDM1* (also known as *BLIMP1*) encodes a protein with a positive regulatory domain 1 element and a zinc finger domain. The protein acts as a repressor of β -interferon gene expression. *PRDM1* is widely expressed and has been implicated in skeletal muscle fiber type differentiation (1), germ cell lineage formation (33), and T cell homeostasis and activation (29), to name but a few potential roles. It has also been suggested that *PRDM1* may be the target of epigenetic downregulation (32). It is not clear yet how sequence variation in *PRDM1* or in another gene in the vicinity could influence the $\dot{V}O_{2max}$ training response, but given the estimated effect size observed in HERITAGE whites, this genomic region deserves further, intensive investigation.

Among other strong members of the SNP predictor score, rs1535628, mapping 516 kb upstream from *GRIN3A*, accounted for ~5% of the response variance in the multivariate regression model (Table 2). *GRIN3A* encodes ionotropic glutamate (NMDA) receptor subunit 3A, which belongs to the superfamily of glutamate-regulated ion channels. *GRIN3A* is widely expressed in neural cells, and a targeted deletion of the gene in mice suggests that it is involved in the development of synaptic elements (15). In periodontal ligament cells, it has been shown that tensile mechanical stress induces glutamate signaling, including *GRIN3A*, to activate cytodifferentiation and mineralization of these cells (19). Sequence variants in *GRIN3A* have also been associated with nicotine dependence (27).

KCNH8 is primarily expressed in the human nervous system (47). A SNP ~268 kb away from the gene locus accounted for 4.5% of the $\dot{V}O_{2max}$ response to exercise training (Table 2). Likewise, a SNP in the vicinity (33 kb) of an open reading frame sequence (frame 27) on chromosome 9 (*C9orf27*) accounted for 4.1% of the training response variance.

A SNP 146 kb away from *ZIC4* explained 3.2% of the $\dot{V}O_{2max}$ response to training in HERITAGE whites (Table 2). *ZIC4* is a transcription factor expressed in the central nervous system, and deletion of *ZIC4* is apparently sufficient to cause Dandy-Walker syndrome, which is characterized by delayed motor development, hypotonia, and often mental retarda-

tion. It is not clear how these markers and genes would functionally relate to the trainability of $\dot{V}O_{2max}$, but, if confirmed in independent studies, they could provide insight into the biology of cardiorespiratory fitness and its adaptation to regular exercise.

Interestingly, the significant HERITAGE association in whites with the SNP in the vicinity of *ZIC4* was replicated in the STRRIDE cohort (Supplemental Table S2). The associations of SNPs rs884736 in the *CAMTA1* locus and rs17581162 (serving as a surrogate for rs10921078) ~68 kb upstream from *RGS18* with the gains in $\dot{V}O_{2max}$ in HERITAGE whites were replicated in HERITAGE blacks. Finally, two SNPs were replicated in DREW subjects: rs1956197 near *DAAMI* and rs17117533 (surrogate for rs824205) in the vicinity of *NDN*. Each accounted for ~1–2% of the $\dot{V}O_{2max}$ response variance in HERITAGE whites. Thus, only 5 of the 15 markers entering in the genomic predictor score calculation that were tested for potential replication could be partially replicated (Supplemental Table S2 and Figs. 4–7).

It is not uncommon to find that associations in whites cannot be replicated in individuals of African descent. In general, genotyping a much larger panel of SNPs covering the genes or chromosomal regions encompassed in the genomic predictor scores may potentially have yielded many more significant markers. In the present study, we typed in the 3 other cohorts only the top 15 SNPs that were identified in HERITAGE whites. While the HERITAGE blacks were selected based on the same set of inclusion and exclusion criteria and were exposed to the same exercise training program and testing protocols as the HERITAGE whites, the situation was quite different for the participants of the DREW and STRRIDE studies (see Table 4). The latter were composed of subjects who were quite different from those of the HERITAGE Family Study at baseline, and the STRRIDE and DREW exercise programs on average did not yield as high an increase in $\dot{V}O_{2max}$ as was observed in the HERITAGE samples. This is probably due to the differences in the training programs. For example, exercise intensity was considerably lower (HR ~50% of $\dot{V}O_{2max}$ HR) in DREW subjects than in HERITAGE participants (HR 55% progressing to 75% of $\dot{V}O_{2max}$ HR).

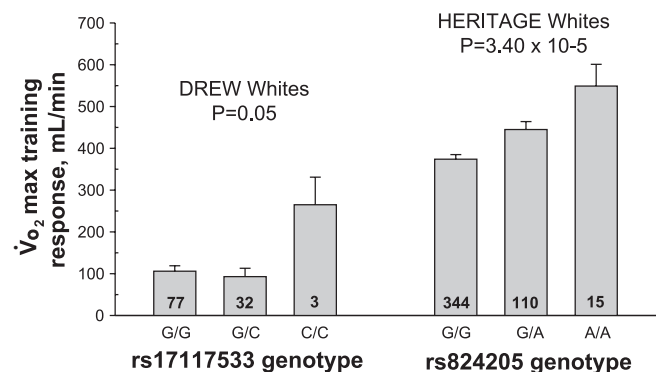


Fig. 7. Association of SNPs rs17117533 and rs824205 [75 kb upstream of neccin (*NDN*)] with age, sex, and baseline $\dot{V}O_{2max}$ -adjusted $\dot{V}O_{2max}$ training responses in DREW whites (left) and HERITAGE whites (right). P values for the main effect of genotype in each cohort are shown at the top of each graph. Numbers of subjects with each genotype are indicated inside each histogram bar. GWAS SNP rs824205 in HERITAGE whites was replaced with rs17117533 in the replication experiments (see METHODS).

Finally, none of the SNPs/transcripts identified in a previous study by Timmons et al. (41) are part of the 21 genomic predictors identified in this report, but 4 of their SNPs are significant at the 0.008 level or better (see below). There were considerable differences in the approaches between the two studies. Timmons et al. used a transcriptomic exploration of the mRNAs from skeletal muscle biopsies, whose level of expression at baseline was predictive of $\dot{V}O_{2\max}$ trainability. Starting with 25 of these candidate transcripts, Timmons et al. added 10 candidate genes previously identified in the HERITAGE cohort. Eleven SNPs from these 35 targets accounted for 23% of the variance in the gains in $\dot{V}O_{2\max}$ in HERITAGE whites (41). When these 11 SNPs were compared with the panel of 21 SNPs resulting from the optimized genomic predictor score of the present study, there were no common genes. The fact that the Timmons et al. molecular predictor relies heavily on baseline expressed transcripts from skeletal muscle is undoubtedly responsible for this divergence between the two genomic predictor schemes. However, on close inspection, markers in 4 of the 11 genes of the Timmons et al. molecular predictors presented with some evidence of associations, albeit not at the level that would have justified their selection for the multivariable regression analyses. One supervillin (*SVIL*) SNP was nominally associated with the gains in $\dot{V}O_{2\max}$ (rs7358069: $P = 0.00028$). Other associations of interest were those in neuropilin 2 (*NRP2*; rs3755233: $P = 0.003$), titin (*TTN*; rs10497520: $P = 0.008$), and carboxypeptidase, vitellogenic-like (*CPVL*; rs1052200: $P = 0.0053$) (plus a few others at $P = 0.05$ or better).

Despite the limitations identified above, the findings of the first GWAS for the trainability of $\dot{V}O_{2\max}$ are quite exciting. Even though the sample size of the HERITAGE whites was limited compared with the large GWAS performed for common diseases on observational cohorts, the findings were impressive. The present study suggests that there are several oligogenes exhibiting allelic variation that contribute to the ability to improve cardiorespiratory fitness with regular exercise. This is in contrast to the findings of most GWA reports based on observational data. We interpret this to mean that the experimental intervention resulted in a clean phenotype with less confounding influences, which improved our ability to find associations with large effects despite the modest sample size. Moreover, the phenotype used herein reflects genotype-exercise training interaction effects that may include a larger genetic variance commonly seen in observational study designs. The genomic predictor score appears to be rather efficient in identifying a priori the low and high responders to regular exercise. Each of the genomic markers and candidate genes retained in the genomic predictor score should be further investigated through appropriate series of experiments and intervention studies to establish their exact functions and the conditions under which they operate in vivo. In particular, to provide a solid foundation for the new biology of adaptation to exercise that is likely to arise from this line of research and for the development of the exercise component of personalized preventive and therapeutic medicine, it will be essential to develop a large, well-characterized cohort of sedentary individuals and to expose them to a fully standardized and carefully monitored exercise training program.

ACKNOWLEDGMENTS

The authors express thanks to Dr. Steven Blair (DREW) and Dr. Arthur Leon, Dr. James Skinner, and Dr. Jack Wilmore (HERITAGE) for contributions to the data collection. The authors also thank Jessica Watkins and Kathryn Cooper for the expert contributions to GWAS and replication genotyping and DNA bank maintenance.

GRANTS

The HERITAGE Family Study was funded by National Heart, Lung, and Blood Institute (NHLBI) Grants HL-45670, HL-47323, HL-47317, HL-47327, and HL-47321 (to C. Bouchard, T. Rankinen, D. C. Rao, A. Leon, J. Skinner, and J. Wilmore). C. Bouchard is partially funded by the John W. Barton Sr. Chair in Genetics and Nutrition. Support for DREW came from NHLBI Grant HL-66262; Life Fitness provided exercise equipment. Support from STRIDE came from NHLBI Grant HL-57354.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

1. Beermann ML, Ardelt M, Girgenrath M, Miller JB. Prdm1 (Blimp-1) and the expression of fast and slow myosin heavy chain isoforms during avian myogenesis in vitro. *PLoS One* 5: e9951, 2010.
2. Blair SN, Kohl HW 3rd, Barlow CE, Paffenbarger RS Jr, Gibbons LW, Macera CA. Changes in physical fitness and all-cause mortality. A prospective study of healthy and unhealthy men. *JAMA* 273: 1093–1098, 1995.
3. Blair SN, Kampert JB, Kohl 3rd HW, Barlow CE, Macera CA, Paffenbarger RS Jr, Gibbons LW. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *JAMA* 276: 205–210, 1996.
4. Bouchard C, Dionne FT, Simoneau JA, Boulay MR. Genetics of aerobic and anaerobic performances. *Exerc Sport Sci Rev* 20: 27–58, 1992.
5. Bouchard C, Leon AS, Rao DC, Skinner JS, Wilmore JH, Gagnon J. The HERITAGE family study. Aims, design, and measurement protocol. *Med Sci Sports Exerc* 27: 721–729, 1995.
6. Bouchard C, Blair SN, Haskell WL. (editors). Physical Activity and Health. Champaign, IL: Human Kinetics, 2007.
7. Bouchard C, Daw EW, Rice T, Perusse L, Gagnon J, Province MA, Leon AS, Rao DC, Skinner JS, Wilmore JH. Familial resemblance for $\dot{V}O_{2\max}$ in the sedentary state: the HERITAGE family study. *Med Sci Sports Exerc* 30: 252–258, 1998.
8. Bouchard C, An P, Rice T, Skinner JS, Wilmore JH, Gagnon J, Perusse L, Leon AS, Rao DC. Familial aggregation of $\dot{V}O_{2\max}$ response to exercise training: results from the HERITAGE Family Study. *J Appl Physiol* 87: 1003–1008, 1999.
9. Bouchard C, Rankinen T, Chagnon YC, Rice T, Perusse L, Gagnon J, Borecki I, An P, Leon AS, Skinner JS, Wilmore JH, Province M, Rao DC. Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE Family Study. *J Appl Physiol* 88: 551–559, 2000.
10. Bouchard C, Rankinen T. Individual differences in response to regular physical activity. *Med Sci Sports Exerc* 33, Suppl 6: S446–S451, 2001.
11. Bouchard C, Rankinen T. Genetic determinants of physical performance. In: *Olympic Textbook Science in Sport*, edited by Maughan RJ. Hoboken, NJ: Wiley-Blackwell, 2009.
12. Bray M, Hagberg J, Perusse L, Rankinen T, Roth S, Wolfarth B, Bouchard C. The human gene map for performance and health-related fitness phenotypes: the 2006–2007 update. *Med Sci Sports Exerc* 41: 34–72, 2009.
13. Chiu CH, Kovacs A, Ford DA, Hsu FF, Garcia R, Herrero P, Saffitz JE, Schaffer JE. A novel mouse model of lipotoxic cardiomyopathy. *J Clin Invest* 107: 813–822, 2001.
14. Church TS, Earnest CP, Skinner JS, Blair SN. Effects of different doses of physical activity on cardiorespiratory fitness among sedentary, overweight or obese postmenopausal women with elevated blood pressure: a randomized controlled trial. *JAMA* 297: 2081–2091, 2007.
15. Das S, Sasaki YF, Rothe T, Premkumar LS, Takasu M, Crandall JE, Dikkes P, Conner DA, Rayudu PV, Cheung W, Chen HS, Lipton SA, Nakanishi N. Increased NMDA current and spine density in mice lacking the NMDA receptor subunit NR3A. *Nature* 393: 377–381, 1998.
16. Duscha BD, Slentz CA, Johnson JL, Houmard JA, Bensimhon DR, Knetzger KJ, Kraus WE. Effects of exercise training amount and intensity on peak oxygen consumption in middle-age men and women at risk for cardiovascular disease. *Chest* 128: 2788–2793, 2005.

17. Erikssen G, Liestol K, Bjornholt J, Thaulow E, Sandvik L, Erikseen J. Changes in physical fitness and changes in mortality. *Lancet* 352: 759–762, 1998.
18. Fletcher GF, Balady G, Blair SN, Blumenthal J, Caspersen C, Chaitman B, Epstein S, Sivarajan Froelicher ES, Froelicher VF, Pina IL, Pollock ML. Statement on exercise: benefits and recommendations for physical activity programs for all Americans. A statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. *Circulation* 94: 857–862, 1996.
19. Fujihara C, Yamada S, Ozaki N, Takeshita N, Kawaki H, Takano-Yamamoto T, Murakami S. Role of mechanical stress-induced glutamate signaling-associated molecules in cytodifferentiation of periodontal ligament cells. *J Biol Chem* 285: 28286–28297, 2010.
20. Kohrt WM, Malley MT, Coggan AR, Spina RJ, Ogawa T, Ehsani AA, Bourey RE, Martin WH, Holloszy JO. Effects of gender, age and fitness level on response of $\dot{V}O_{2\max}$ to training in 60–71 yr olds. *J Appl Physiol* 71: 2004–2011, 1991.
21. Kokkinos P, Myers J, Faselis C, Panagiotakos DB, Pittaras A, Manolis A, Kokkinos JP, Karasik P, Greenberg M, Papademetriou V, Fletcher R. Exercise capacity and mortality in older men: a 20-year follow-up study. *Circulation* 122: 790–797, 2010.
22. Kraus WE, Torgan CE, Duscha BD, Norris J, Brown SA, Cobb FR, Bales CW, Annex BH, Samsa GP, Houmard JA, Slentz CA. Studies of a targeted risk reduction intervention through defined exercise (STRRIDE). *Med Sci Sports Exerc* 33: 1774–1784, 2001.
23. Lee DC, Sui X, Ortega FB, Kim YS, Church TS, Winett RA, Ekelund U, Katzmarzyk PT, Blair SN. Comparisons of leisure-time physical activity and cardiorespiratory fitness as predictors of all-cause mortality in men and women. *Br J Sports Med*; doi:10.1136/bjism.2009.066209.
24. Lobo S, Wiczer BM, Bernlohr DA. Functional analysis of long-chain acyl-CoA synthetase 1 in 3T3-L1 adipocytes. *J Biol Chem* 284: 18347–56, 2009.
25. Loos RJ. Recent progress in the genetics of common obesity. *Br J Clin Pharmacol* 68: 811–829, 2009.
26. Lortie G, Simoneau JA, Hamel P, Boulay MR, Landry F, Bouchard C. Responses of maximal aerobic power and capacity to aerobic training. *Int J Sports Med* 5: 232–236, 1984.
27. Ma JZ, Payne TJ, Li MD. Significant association of glutamate receptor, ionotropic N-methyl-D-aspartate 3A (GRIN3A), with nicotine dependence in European- and African-American smokers. *Hum Genet* 127: 503–512, 2010.
28. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 461: 747–753, 2009.
29. Martins GA, Cimmino L, Shapiro-Shelef M, Szabolcs M, Herron A, Magnusdottir E, Calame K. Transcriptional repressor Blimp-1 regulates T cell homeostasis and function. *Nat Immunol* 7: 457–465, 2006.
30. Marwick TH, Hordern MD, Miller T, Chyun DA, Bertoni AG, Blumenthal RS, Philippides G, Rocchini A; Council on Clinical Cardiology, American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee; Council on Cardiovascular Disease in the Young; Council on Cardiovascular Nursing; Council on Nutrition, Physical Activity, and Metabolism; Interdisciplinary Council on Quality of Care and Outcomes Research. Exercise training for type 2 diabetes mellitus: impact on cardiovascular risk: a scientific statement from the American Heart Association. *Circulation* 119: 3244–3262, 2009.
31. Morss GM, Jordan AN, Skinner JS, Dunn AL, Church TS, Earnest CP, Kampert JB, Jurca R, Blair SN. Dose response to exercise in women aged 45–75 yr (DREW): design and rationale. *Med Sci Sports Exerc* 36: 336–344, 2004.
32. Nie K, Zhang T, Allawi H, Gomez M, Liu Y, Chadburn A, Wang YL, Knowles DM, Tam W. Epigenetic down-regulation of the tumor suppressor gene PRDM1/Blimp-1 in diffuse large B cell lymphomas. A potential role of the microRNA Let-7. *Am J Pathol* 177: 1470–1479, 2010.
33. Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, Sano M, Barton SC, Obukhanych T, Nussenzweig M, Tarakhovskiy A, Saitou M, Surani MA. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* 436: 207–213, 2005.
34. Phillips CM, Goumidi L, Bertrais S, Field MR, Cupples LA, Ordoas JM, Defoort C, Lovegrove JA, Drevon CA, Gibney MJ, Blaak EE, Kiec-Wilk B, Karlstrom B, Lopez-Miranda J, McManus R, Hercberg S, Lairon D, Planells R, Roche HM. Gene-nutrient interactions with dietary fat modulate the association between genetic variation of the ACSL1 gene and metabolic syndrome. *J Lipid Res* 51: 1793–1800, 2010.
35. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report, 2008. Washington, DC: U. S. Department of Health and Human Services, 2008.
36. Prud'homme D, Bouchard C, Leblanc C, Landry F, Fontaine E. Sensitivity of maximal aerobic power to training is genotype dependent. *Med Sci Sports Exerc* 16: 489–493, 1984.
37. Rankinen T, Bouchard C. Genes, genetic heterogeneity, and exercise phenotypes. In: *Molecular and Cellular Exercise Physiology*, edited by Mooren FC, Volker K. Champaign, IL: Human Kinetics, 2005.
38. Skinner JS, Wilmore KM, Krasnoff JB, Jaskolska A, Jaskolska A, Gagnon J, Province MA, Leon AS, Rao DC, Wilmore JH, Bouchard C. Adaptation to a standardized training program and changes in fitness in a large, heterogeneous population: the HERITAGE Family Study. *Med Sci Sports Exerc* 32: 157–161, 2000.
39. Skinner JS, Jaskolski A, Jaskolska A, Krasnoff J, Gagnon J, Leon AS, Rao DC, Wilmore JH, Bouchard C; HERITAGE family study. Age, sex, race, initial fitness, and response to training: the HERITAGE Family Study. *J Appl Physiol* 90: 1770–1776, 2001.
40. Sui X, Hecker SP, Lee IM, Church TS, Colabianchi N, Lee CD, Blair SN. A prospective study of cardiorespiratory fitness and risk of type 2 diabetes in women. *Diabetes Care* 31: 550–555, 2008.
41. Timmons JA, Knudsen S, Rankinen T, Koch LG, Sarzynski M, Jensen T, Keller P, Scheele C, Vollaard NB, Nielsen S, Akerstrom T, Macdougald OA, Jansson E, Greenhaff PL, Tarnopolsky MA, van Loon LJ, Pedersen BK, Sundberg CJ, Wahlestedt C, Britton SL, Bouchard C. Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *J Appl Physiol* 108: 1487–96, 2010.
42. United States Department of Health and Human Services. *Physical Activity and Health: a Report of the Surgeon General*. Atlanta, GA: U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1996.
43. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarrroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segrè AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Benediktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bengtsson Boström K, Bravenboer B, Bumpstead S, Burt NP, Charpentier G, Chines PS, Cornelis M, Couper DJ, Crawford G, Doney AS, Elliott KS, Elliott AL, Erdos MR, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnston PR, Jørgensen T, Kao WH, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lievever A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JR, Petersen AK, Platou C, Proença C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shraider P, Sigurdsson G, Sparso T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, Tichet J, Tuomi T, van Dam RMvan Haefen TW, van Herpt T, van Vliet-Ostapchouk JV, Bragi Walters G, Weedon MN, Wijmenga C, Witteman J, Bergman RN, Cauchi S, Collins FS, Gloy AL, Gyllenstein U, Hansen T, Hide WA, Hitman GA, Hofman A, Hunter DJ, Hveem K, Laakso M, Mohlke KL, Morris AD, Palmer CN, Pramstaller PP, Rudan I, Sijbrands E, Stein LD, Tuomilehto J, Uitterlinden A, Walker M, Wareham NJ, Watanabe RM, Abecasis GR, Boehm BO, Campbell H, Daly MJ, Hattersley AT, Hu FB, Meigs JB, Pankow JS, Pedersen O, Wichmann HE, Barroso I, Florez JC, Frayling TM, Groop L, Sladek R, Thorsteinsdottir U, Wilson JF, Illig T, Froguel P, van Duijn CM, Stefansson K, Altshuler D, Boehnke M, McCarthy MI; investigators; GIANT Consortium MAGIC. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 42: 579–589, 2010.
44. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 76: 887–893, 2005.

45. Wilmore JH, Stanforth PR, Gagnon J, Rice T, Mandel S, Leon AS, Rao DC, Skinner JS, Bouchard C. Heart rate and blood pressure changes with endurance training: the HERITAGE Family Study. *Med Sci Sports Exerc* 33: 107–116, 2001.
46. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42: 565–569, 2010.
47. Zou A, Lin Z, Humble M, Creech CD, Wagoner PK, Kraffe D, Jegla TJ, Wickenden AD. Distribution and functional properties of human KCNH8 (Elk1) potassium channels. *Am J Physiol Cell Physiol* 285: C1356–C1366, 2003.

