Do genetic variations alter the effects of exercise training on cardiovascular disease and can we identify the candidate variants now or in the future?

James M. Hagberg
Department of Kinesiology, School of Public Health, University of Maryland, College Park, Maryland
Submitted 4 February 2011; accepted in final form 9 May 2011

Hagberg JM. Do genetic variations alter the effects of exercise training on cardiovascular disease and can we identify the candidate variants now or in the future? J Appl Physiol 111: 916–928, 2011. First published May 12, 2011; doi:10.1152/japplphysiol.00153.2011.—Cardiovascular disease (CVD) and CVD risk factors are highly heritable, and numerous lines of evidence indicate they have a strong genetic basis. While there is nothing known about the interactive effects of genetics and exercise training on CVD itself, there is at least some literature addressing their interactive effect on CVD risk factors. There is some evidence indicating that CVD risk factor responses to exercise training are also heritable and, thus, may have a genetic basis. While roughly 100 studies have reported significant effects of genetic variants on CVD risk factor responses to exercise training, no definitive conclusions can be generated at the present time, because of the lack of consistent and replicated results and the small sample sizes evident in most studies. There is some evidence supporting “possible” candidate genes that may affect these responses to exercise training: APO E and CETP for plasma lipoprotein-lipid profiles; eNOS, ACE, EDN1, and GNB3 for blood pressure; PPARG for type 2 diabetes phenotypes; and FTO and BAR genes for obesity-related phenotypes. However, while genotyping technologies and statistical methods are advancing rapidly, the primary limitation in this field is the need to generate what in terms of exercise intervention studies would be almost incomprehensible sample sizes. Most recent diabetes, obesity, and blood pressure genetic studies have utilized populations of 10,000–250,000 subjects, which result in the necessary statistical power to detect the magnitude of effects that would probably be expected for the impact of an individual gene on CVD risk factor responses to exercise training. Thus at this time it is difficult to see how this field will advance in the future to the point where robust, consistent, and replicated data are available to address these issues. However, the results of recent large-scale genomewide association studies for baseline CVD risk factors may drive future hypothesis-driven exercise training intervention studies in smaller populations addressing the impact of specific genetic variants on well-defined physiological phenotypes.

It appears that cardiovascular disease (CVD) was initially proposed to be heritable and have a genetic basis in 1769 (141). Some of the first evidence supporting this hypothesis came from a 1930 study of a family with a three-generation history of sudden death (53). Then in 1955, Thomas and Cohen (141) reported that individuals were ~1.5 times more likely to have CVD if their siblings or parents had CVD. Since then numerous studies have reported heritability of 0.30–0.50 for various quantitative CVD measures (95). Thus these examples selected from a vast published literature clearly indicate that CVD is heritable and, therefore, may have a genetic basis. The next step is to identify the genes underlying this relationship.

CVD candidate gene studies

The first CVD candidate gene study appears to be in 1982 by Owerbach and co-workers (97) who, having previously identified an insulin gene variant, found it was associated with macroangiography in type 2 diabetes mellitus (T2DM) patients and healthy individuals. Later they found (81) this variant was 2–5 times more frequent in CVD patients than healthy individuals. In 1983 Menzel and co-workers (84) reported on the first variant that has remained as a candidate CVD gene. They found that apolipoprotein E (APO E) 2/3 genotype frequency was lower in CVD patients than healthy controls and con-
cluded that this allele may protect against the development of CVD. This report was rapidly followed by a number of studies, and a 1988 review (24) concluded that there was a small but growing literature supporting a cardioprotective role for the E2 allele.

The second variant investigated relative to CVD is the angiotensin-converting enzyme (ACE) I/D locus (21). This initial paper concluded that ACE genotype was a “potent” risk factor as across four European populations the DD genotype was 1.34 times more frequent in CVD cases than controls. This paper was followed by a flurry of publications and, although not having completely consistent results, the consensus was that ACE DD genotype individuals were at higher risk for CVD (13, 31, 82, 116).

Overall, the impact of genetic variants on CVD has remained a topic of major interest as indicated by an early 2011 PubMed search for genotype and CVD that yielded >22,000 citations. Thus there is a substantial literature relative to CVD candidate genes, and it continues to emphasize a strong genetic basis for CVD.

CVD GENOMEWIDE LINKAGE AND ASSOCIATION STUDIES

Additional methods for identifying genes underlying a phenotype are genomewide linkage (GWL) and association (GWA) studies. One strength of GWL/GWA studies is to potentially identify “novel” genes that would not a priori be considered candidate genes.

The first CVD GWL study was from Pajukanta et al. (99) in 2000 based on two independent Finnish cohorts, and they found two loci (2q21.1–22, Xq23–26) that were linked with premature CVD. Four CVD GWL studies were published in 2001–2002. Francke et al. (36) reported that 16p13ter was linked with premature CVD, Harrap and colleagues (47) found that 2q36–q37.3 was linked with acute coronary syndrome, Lange and co-workers (72) found linkages at 6p21.3 and 10q21.3 with coronary artery calcification, and Broeckel and co-workers (20) reported linkage on chr 14 with premature CVD. Thus, although these initial loci were not replicated across studies, they did provide further evidence of the genetic nature of CVD.

The first CVD GWA study was published in 2007 with a number of such studies published that year (74, 83, 94, 129, 152a). In perhaps the most exciting genetic results ever generated, all of these studies consistently identified CVD associations with 9p21. While the two most biologically plausible candidate genes [cyclin-dependent kinase inhibitor 2A (CDKN2A) and 2B (CDKN2B)] lie some distance from this locus, they do reside within the same linkage disequilibrium block. Numerous followup studies have replicated the finding that 9p21 is an important CVD locus (18, 68, 75, 128).

In a 2009 review Arking and Chakravarti (8) summarized the previous literature relative to GWA and candidate gene studies for CVD and listed 17 loci they concluded were significantly related to an elevated CVD risk, with each of them increasing an individual’s risk by 12–82%. Especially notable among the loci found to be associated with CVD were the CDKN2A, CDKN2B, proprotein convertase subtilisin/kexin type 9 (PCSK9), and low-density lipoprotein receptor (LDLR) genes.

INTERACTIVE RELATIONSHIPS BETWEEN EXERCISE AND GENETICS ON CVD

The next step then should be to summarize what is known about the interactive effects of genetics and exercise training on CVD. However, this issue does not appear to have been addressed in the available scientific literature. Thus the best we can do at this point is to summarize what is known about the interactive effect of genetics and exercise training on CVD risk factors, assuming that improving CVD risk factors would also “improve” CVD.

HERITABILITY OF CVD RISK FACTORS

Numerous studies have quantified the heritability of the four CVD risk factors that will be addressed in this review: plasma lipoprotein-lipids, blood pressure (BP), and T2DM- and obesity-related phenotypes. A 1974 twin study on the genetics of plasma lipoprotein-lipid levels (52) found heritability of 0.24–0.79 for various plasma lipoprotein-lipid levels. Three subsequent and larger studies found generally similar heritability for plasma total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglyceride (TG) levels (62, 89, 102). Numerous studies have found heritability of 0.20–0.50 for BP. For example, de Oliveira et al. (25) reported heritability of 0.26 for systolic and diastolic BP in 81 Brazilian families. More recently, Wang et al. (149) reported heritability of 0.59–0.75 for casual and ambulatory BP, and BP responses to a video game and a social stressor in Caucasians and African Americans. The first data on the heritability of diabetes were from 1933 by Pincus and White (105), who reported that 23% of diabetics, but only 10% of nondiabetics, had a family history of diabetes. In 1976 Lindsten et al. (79) reported that intravenous glucose tolerance test (IVGTT) glucose and insulin phenotypes had heritability of 0.38–0.72. For obesity-related phenotypes, in 1997 Rice and colleagues (119) found heritability of 0.47 for abdominal visceral fat and 0.55 for total fat mass. In 2010 Song et al. (132) found heritability of 0.60–0.71 for body mass index (BMI), waist circumference, and waist:hip ratio in Koreans. Thus these examples from a much larger literature clearly demonstrate that these four CVD risk factors are heritable. The next step is to identify genes that might underlie these phenotypes.

CANDIDATE GENE ASSOCIATION STUDIES OF CVD RISK FACTORS

The only viable option initially possible to address genetics were candidate gene studies. In this section, a brief historical context is established relative to candidate genes underlying our four CVD risk factors of interest. Numerous studies have investigated the effect of candidate genes on plasma lipoprotein-lipids, with a recent PubMed search for genotype and lipids yielding >14,000 citations and a very long list of candidate genes/single-nucleotide polymorphisms (SNPs). Only the results relative to two genes [APOE, cholesteryl ester transfer protein (CETP)] will be summarized here. The APOE gene was first shown to affect plasma lipoprotein-lipids in 1980 in a homozygous familial hypercholesterolemia case study (49). A 2007 meta-analysis (11) found that APO E2/3 carriers had 20 and 17 mg/dl lower LDL-C and
cholesterol levels, respectively, and ~3 mg/dl higher HDL-C levels than E3/4 individuals. These data plus numerous others indicate that APO E genotype is the strongest genetic predictor of lipoprotein-lipid levels.

The CETP gene was first investigated in 1985 in a familial hyperalphalipoproteinemia case study (67) with an HDL-C level of 301 mg/dl. A recent meta-analysis (142) summarized the impact of CETP genotype on lipoprotein-lipids and found that for each variant CETP allele present, HDL-C levels were 6% higher than in common homozygotes. Thus a common CETP TaqIB SNP [rs708272] also affects plasma HDL-C levels.

In 2004 Knoblauch et al. (65) investigated the effects of haplotypes in 13 genes and concluded that their 13 genes, which included APO E and CETP, accounted for 67% of the genetic variance in LDL-C levels and 58% for HDL-C. Thus at least in this study genes involved in lipoprotein-lipid metabolism appear to explain a substantial portion of the variation in plasma lipoprotein-lipid levels among individuals. However, a more recent and much larger GWAS study found that genetic variants they could identify accounted for only 10–12% of the total variance and 25–30% of the total genetic variance in plasma lipoprotein-lipids (139).

The first large-scale studies assessing candidate genes and BP investigated ACE genotype. In 1992/1993 two groups (46, 158) reported that the ACE I/D genotype associated with hypertension. In 1992 the M235T angiotensinogen (AGT) SNP became the second variant studied relative to BP when a strong association was reported between this variant and hypertension (59). This was followed in 1993/1994 by similar results from other groups (12, 48, 51). In addition to these candidate BP genes, there have been a huge number of other candidate genes/SNPs found to be associated with BP, as demonstrated by a recent PubMed search for BP/hypertension and genotype that generated ~120,000 citations.

In the first T2DM candidate gene study, Rotwein and co-workers (124) in 1981 found a ~2-fold higher prevalence of a large DNA insertion variant in the insulin gene promoter in T2DM patients compared with nondiabetics. This resulted in numerous follow-up studies with some finding (54, 98, 125) and others not finding altered distributions in diabetics (29, 66, 100). Since then a vast number of studies have investigated the association between candidate genes and T2DM phenotypes. In fact, a PubMed search in early 2011 for diabetes and genotype resulted in >11,000 citations. In terms of clinical impact, Weedon and co-workers (150) in 2006 reported that for each risk allele an individual had at three susceptibility loci (KCNJ11, TCF7L2, PPARG), their T2DM risk increased multiplicatively by 28%, with a person with all six risk alleles having a ~6-fold increased T2DM risk compared with those with no risk alleles.

The first study that assessed genotypes and obesity was in 1985 by Bouchard et al. (16). They typed alleles at an HLA system locus in 1,578 individuals but did not find any consistent associations. In the first significant results, a number of 1994–1996 studies indicated that leptin (LEP) variants were associated with obesity-related phenotypes (14, 23, 28, 117). In 1996 Bouchard and Perusse (15) published the first Human Obesity Gene Map reporting that to that point 10 genes had been associated with an obesity phenotype. Only a year later the number of genes had quadrupled to 40 (101)! In 2006 the last edition of this review (114) listed 127 candidate genes showing associations with obesity phenotypes. Furthermore, a recent PubMed search for obesity/body fat and genotype resulted in ~140,000 citations. Thus there is a huge body of literature addressing the role of genetic variation in determining obesity-related phenotypes.

**GENOMEWIDE LINKAGE AND ASSOCIATION STUDIES FOR CVD RISK FACTORS**

The first lipoprotein-lipids linkage study was in 1992; they assessed five markers in families with an “atherogenic lipoprotein phenotype” (92) and found linkage with an LDLR gene locus. In the first lipoprotein-lipids GWL study, Rainwater et al. (106) in 1999 identified two loci that were linked with small LDL particle levels with plausible candidate genes lying adjacent to these loci. In 2009 Hiura and coworkers (55) typed ~370,000 SNPs in 900 individuals and found 43 associated with HDL-C levels with the most significant being adjacent to the CETP gene.

A 2009 review by Hegele (50) concluded that eight common genetic variants affect plasma LDL-C levels including APOE, LDLR, APOB, and PCSK9. Hegele also listed 12 common variants that affect plasma TG levels, including APOA5, lipoprotein lipase (LPL), hepatic lipase (LIPC), and APOB. A 2010 review generated a list of ~20 genes that appeared to influence plasma HDL-C levels (152). Prominent among these genes were APOB, APOA1, APO E, ATP-binding cassette, subfamily A, member 1 (ABCA1), LIPC, endothelial lipase (LIPG), CETP, lecithin-cholesterol acyltransferase (LCAT), and LPL. Another 2010 study in ~100,000 individuals of European ancestry identified 95 loci, 59 of which were novel, that were significantly associated with different components of the plasma lipoprotein-lipid profile (139).

In the first BP GWL study, Rice et al. (123) typed 420 markers in 206 Quebec families and found 7 loci linked with BP. A 2009 study assessed ~800,000 SNPs in African Americans and the 10 best associations for BP and hypertension all had $P < 1.03 \times 10^{-5}$ (2). Another recent study assessed 2.5 million SNPs in ~34,000 Europeans and in replication cohorts totaling ~110,000 individuals (91), and found eight loci associated with BP at $P < 1 \times 10^{-8}$. They identified eight loci that were strongly and consistently associated with BP/hypertension with the most prominent and promising being the cytochrome P450, family 17, subfamily A, polypeptide 1 (CYP17A1, $P = 7 \times 10^{-24}$), cytochrome P450, family 1, subfamily A, polypeptide 2 (CYP1A2, $P = 1 \times 10^{-23}$), and methylenetetrahydrofolate reductase (MTHFR, $P = 2 \times 10^{-15}$) genes.

In the first T2DM linkage study, Elbein and coworkers (30) in 1995 typed markers in 19 candidate loci in ~450 Caucasians and found that a chromosome 17 locus was associated with T2DM. The first T2DM GWL studies were published in 1996 (45, 80). Hanis and co-workers (45) found one locus linked to T2DM in two Mexican-American cohorts; however, no linkage was found in non-Hispanic whites and Japanese. In a 2000 follow-up study, the calpain-10 (CAPN10) gene was positionally cloned to this locus and found to affect T2DM risk in three populations (57). The first T2DM GWA studies were published in 2007, with eight being published that year. These studies...
provided very strong evidence for T2DM risk genes, as each study found a number of strong associations and many of the loci replicated within and across studies. A 2009 review concluded that there were fairly strong and replicated data supporting the inclusion of four genes, CAPN10, PPARG, potassium inwardly-rectifying channel, subfamily J, member 11 (KCNJ11), and transcription factor 7-like 2 (TCF7L2), as candidate genes that alter an individual’s risk of developing T2DM (87). An even more recent 2010 study by Voight et al. (148) found 12 novel loci, over and above those previously identified, that were associated with T2DM susceptibility.

The first obesity GWL study was in 1997 with Norman and co-workers (76) identifying linkage at 11q21-q22 with percent body fat in Pima Indians. Within the next 5 years this initial study was followed by a number of GWL studies that identified linkages with obesity-related phenotypes (9, 33, 43, 58, 78, 145, 155, 160). Unfortunately, there was minimal replication in these initial studies, as the significantly linked loci were spread across 10 chromosomes. A 2010 very large GWA study concluded that there are 32 loci that have been associated with BMI or obesity (133). Notable among the genes listed are the fat mass and obesity-related (FTO), melanocortin 4 receptor (MC4R), and brain-derived neurotrophic factor (BDNF) genes, as well as a number of additional loci that lie near critical hypothalamic regulators of energy balance.

Thus GWL/GWA studies have identified several loci related to these four CVD risk factors, providing further evidence of their strong genetic basis.

VARIABILITY IN CVD DISEASE RISK FACTOR RESPONSES TO EXERCISE TRAINING

Although not widely appreciated, the responses of CVD risk factors to even highly standardized exercise training interventions vary substantially among individuals. Leon and HERITAGE co-workers (76) reported that their highest quartile of HDL-C responders had an 18% average increase in HDL-C levels with training; however, their lowest quartile of responders experienced a 9.3% average decrease in HDL-C levels with training. We have also found substantial interindividual variability in HDL-C responses to training (44). Unpublished data from our laboratory show that more than half of ~70 pre- and Stage 1 hypertensives reduced their BP with exercise training. However, ~15% of them reduced their BP by >10 mmHg with training, while another 10% increased their BP by this amount. Boule and HERITAGE co-workers (17) found that training significantly improved T2DM phenotypes, but 42% of their subjects did not change or decreased insulin sensitivity with training. For their other T2DM-related measures, ~50% of individuals did not change or had worse measures after training. Similar variability in obesity-related phenotype responses to training were also evident in HERITAGE (153). Thus there are clearly substantial interindividual differences in the responses of CVD risk factors to exercise training.

It has been proposed that being able to account for these interindividual differences in response to exercise training, or other lifestyle interventions or medications, based on genetic or other individual information could eventually evolve into a new era of “personalized medicine” (39). Furthermore, understanding the potential genetic basis for at least a portion of the interindividual variability in the physiological responses to exercise training will provide a strong rationale for future, more in-depth mechanistic studies of the cellular and molecular underpinnings of these responses. The next step is to determine if these responses are heritable.

HERITABILITY OF CVD RISK FACTOR RESPONSES TO EXERCISE TRAINING

Unfortunately, only HERITAGE has an appropriate sample size and study design to assess the heritability of responses to exercise training. In 2000, An and HERITAGE co-workers (5) reported that apo A-1 and B-100 responses to training were heritable in whites. In that same year, Hong and co-workers (56) reported heritability of 0.22 for TG changes with training in HERITAGE whites. Rice et al. (120) also found that seven of eight lipoprotein-lipid changes with training in blacks and whites had significant heritability of 0.25–0.64. Rice and co-workers found that resting BP changes with training had heritability of 0.14–0.24 (118). An and co-workers (6) reported heritability for four T2DM-related training response phenotypes to be ~0.10 in whites and 0.20–0.30 in blacks. On the other hand, three studies from HERITAGE have reported minimal heritability for obesity phenotype responses to training (4, 103, 121). Thus little is known about the heritability of CVD risk factor responses to training, but the available data indicate these responses are somewhat heritable, providing a rationale to identify their genetic basis.

CANDIDATE GENE STUDIES OF CVD RISK FACTOR RESPONSES TO EXERCISE TRAINING

This section will only review genes whose impact has been assessed relative to one of our four CVD risk factors in at least two studies, since clearly replication, at any level, is a major criterion underlying the validity of any research result, especially in the area of genetics where replication has been a critical issue.

Plasma lipoprotein-lipids. Our last Human Gene Map for Performance and Health-Related Fitness Phenotypes indicated there were 26 studies reporting significant genotype-dependent effects on lipoprotein-lipid responses to training (19). Since then another eight studies have been published. The two most-studied genes relative to training-induced lipoprotein-lipid changes are APOE and CETP, with 12 studies assessing the impact of APOE and 4 the effect of CETP.

Taimela et al. (138) in 1996 first provided evidence of a genotype-dependent effect of training on plasma lipoprotein-lipids. They found that physical activity (PA) and plasma lipoprotein-lipid levels were not related in APO E4/4, were moderately related in APO E3/4 and 3/3, and were strongly related in APO E2/3 males. St-Amand and coworkers (136) reported similar findings in another cross-sectional study. Since then a number of cross-sectional studies have reported significant, but not always consistent, effects of APO E genotype on the relationship between PA levels and lipoprotein-lipid levels.

In the first study of the genetics of lipoprotein-lipid responses to an exercise training intervention, in 1999 we found that APO E2 men improved HDL-C and HDL2-C more with training than APO E4 men (42). In 2004 Thompson et al. (143) found that lipoprotein-lipids improved with training in APOE
2/3 and 3/3, but not 3/4, individuals. Leon and HERITAGE co-workers (77) reported significant effects of APOE genotype on training-induced responses in 10 of 16 lipoprotein-lipid levels in whites, but in only 2 of 16 responses in blacks.

In 2002 we reported that CETP TaqI B1B2 heterozygotes improved HDL subfractions more with training than B1 homozygotes (154). In 2004 Mukherjee and Shetty (88) in a cross-sectional study reported that there was an interaction between CETP genotype and PA levels with active individuals exhibiting higher HDL-C levels than inactive individuals only in the B1B1 genotype group. In 2005 Ayyobi and co-workers (10) reported that B1B1 individuals improved lipoprotein-lipid levels more with training than B2 carriers. In 2007 Spielmann and co-workers (134) reported that A homozygote women at a different CETP locus (C-629A) improved HDL3-C and ApoA1 levels the most with training.

Only three other genes have been investigated in more than one study. The three LPL gene studies all assessed different LPL variants. For both the LIPC and LIPG genes, the two studies that assessed their impact each studied different variants. The remaining evidence for genotype effects on lipoprotein-lipid responses to training consists of single studies for each candidate gene (19).

**Blood pressure.** Our last Human Gene Map for Performance and Health-Related Fitness Phenotypes reported that 18 studies had found significant genotype-dependent effects for BP changes with exercise training (19). Since then another eight studies have addressed this issue. The three most studied genes relative to training-induced BP changes are endothelial nitric oxide synthase (eNOS), ACE, and AGT.

Three 2010 studies assessed the eNOS T-786C variant and BP responses to training. Zago and co-workers (157) reported that eNOS T-786C/C haplotype significantly influenced BP changes with training. Negrao et al. (90) found that exercise training increased forearm vascular conductance during handgrip exercise in T-786 homozygotes with no such changes evident in C allele carriers. Sponton and co-workers (135) reported T-786 homozygote women reduced BP more than C allele carriers with training. Two cross-sectional studies assessed the impact of different eNOS SNPs on the relationship between PA and BP (64, 146).

In 1999 in a small number of hypertensives we reported that ACE I allele carriers reduced or tended to reduce BP more with training than ACE D homozygotes (41). Zhang and co-workers (159) found that BP decreased with training in ACE I allele carriers, but not in ACE DD, hypertensives. Recently in a small population of obese adolescents Foschini and co-workers (35) reported that D allele carriers were more likely to reduce BP with exercise training; however, the training-induced BP reductions were similar among genotype groups.

For the three studies addressing AGT or AGT-related variants, one assessed the impact of AGT M235T on BP responses to endurance training, another the impact of the M235T variant on training-induced responses of BP during submaximal exercise, and the final study assessed the impact of two different AGT-related genes on BP changes with strength training (26, 110, 115).

In 2007 Rankinen and co-workers (109) found that endothelin-1 (EDN1) haplotype was associated with BP responses to training in HERITAGE. Then, in another independent cohort, they also found that two EDN1 SNPs interacted significantly with CV fitness to impact risk for developing hypertension, with both SNPs exerting their effects on BP only in low-fit individuals.

Rankinen and co-workers (111) also found that black G protein beta polypeptide 3 (GNB3) C825T CC women reduced BP more with training than TT genotype black women, although no relationships were evident in black men or whites. Grove and co-workers (40) also found that this variant interacted with PA levels and obesity to affect the prevalence of hypertension.

The remaining evidence for genotype-dependent effects on BP responses to exercise training consists of only single studies quantifying the effects of different candidate genes (19).

**T2DM phenotypes.** Our 2006–2007 update of the Human Gene Map for Performance and Health-Related Fitness Phenotypes (19), our 2010 Advances in Exercise, Fitness, and Performance Genomics (113), and the literature since then indicate there are ~30 papers relative to candidate genes affecting T2DM-related phenotype responses to training.

In 1997 in the first study addressing this issue, Sakane et al. (127) found that with a weight loss program that included exercise, obese T2DM beta adrenergic receptor (B3AR) Trp64Trp women improved fasting glucose and insulin resistance index more than Arg64 allele carriers. Kahara et al. (61) also found that exercise training in men decreased fasting glucose levels only in Trp homozygotes. Two other studies have addressed the effect of other AR gene variants on T2DM phenotype responses to training (60, 69).

The most frequently studied variants relative to T2DM phenotype responses to training are in the peroxisome proliferator-activated receptor (PPAR) signaling cascade. In terms of PPAR beta (PPARD), rs2267668 locus AA homozygotes improved insulin sensitivity more with training in two recent studies (137, 140). Both studies also found that PPARD rs1053049 TT homozygotes improved insulin sensitivity the most with training. These two variants are in substantial linkage disequilibrium (0.75–0.80). Stefan et al. (137) also found that the PPAR coactivator 1 alpha (PGC1A) Gly482Ser variant independently affected training-induced insulin sensitivity changes.

In some of the most consistent and exciting data relative to exercise training and genetics, all four exercise training studies investigating the PPAR gamma (PPARG) Pro12Ala variant found that Ala allele carriers had better T2DM-related phenotype responses to training. First Kahara et al. (61) reported that training decreased fasting insulin levels and insulin resistance more in Ala allele carriers. In 2005 two studies (1, 151) found greater T2DM-related phenotype responses to training in Ala carriers vs. Pro homozygotes. Finally, in 2010 Ruchat and HERITAGE co-workers (126) typed eight diabetes-susceptibility variants and found that only the PPARG Pro12Ala variant had a significant effect with, again, Ala carriers exhibiting the largest training-induced improvements in T2DM-related phenotypes. In addition, four cross-sectional studies have generally shown significant interactive effects between this variant and PA levels on T2DM-related phenotypes.

ACE is the only other gene that has been addressed in more than one study. In 2002 Dengel et al. (27) reported that a small number of older ACE II genotype individuals increased insulin sensitivity more with training than D allele carriers. However,
two more recent and larger studies did not find effects of ACE genotype on T2DM phenotype responses to training (35, 63).

The effects of the remaining genes on training-induced changes in T2DM phenotypes have only been assessed in single studies.

**Obesity phenotypes.** Our 2006–2007 update of the Human Gene Map for Performance and Health-Related Fitness Phenotypes (19) listed 14 studies assessing the genetics of obesity phenotype responses to training. Now there are ~20 papers addressing the interactive effects of genetics and exercise training on obesity-related phenotypes.

The variants studied most frequently relative to obesity-related phenotype responses to training lie within the AR pathway. In 2003 Garenc and co-workers (38) reported that beta-AR (B2AR) Arg16 homozygote obese white women reduced obesity-related phenotypes with training more than women carrying no or one Arg alleles. Also in 2003 Shiwaku and co-workers (130) reported that with a weight loss intervention that included exercise, Japanese women carrying the B3AR Trp64 allele improved their obesity-related phenotypes, while those carrying the Arg64 allele did not. In 2004 we reported that B2AR Gln27Glu, B3AR Trp64Arg, and alpha2b-AR (A2bAR) Gln12Glu and Glu12/Glu9 variants independently and interactively affected total body and trunk fat responses to training (104), although we could not replicate these findings in the larger HERITAGE cohort (unpublished results). Our group also found that B2AR Gln27Gln and A2bAR Glu12/Glu9 variants affected intramuscular fat changes with strength training (156).

In 2010 Rankinen and HERITAGE colleagues (112) reported that C allele carriers at the fat mass and obesity-associated (FTO) gene rs8050136 locus had threefold greater improvements in obesity phenotype responses to training than A homozygotes. However, a second report from this group on another independent cohort found that, if anything, A homozygotes responded better to training (85). In addition, two other studies found no effect of FTO genotype on obesity phenotype responses to lifestyle interventions that included increased PA levels (37, 73). A number of cross-sectional studies have relatively consistently indicated that PA levels alter the relationship between FTO genotype and obesity-related phenotypes (7, 107, 131, 147).

The impact of the ACE I/D variant on training-induced changes in obesity-related phenotypes has been assessed in four studies. In 1999 Montgomery and co-workers (86) reported that young British ACE D homozygotes reduced total body fat and subcutaneous skinfolds more with military training than ACE I allele carriers. However, three follow-up studies found no effect of ACE genotype on obesity-related phenotype responses to training in obese adolescents (35), Korean women (63), or healthy Danes with a family history of diabetes (96).

Two studies have assessed the impact of PPAR genes, with one finding significant effects of a PPARA variant on body weight reductions with exercise training (96), while the other found a significant effect of a PPARG variant on subcutaneous fat volume reductions with training (144).

Clearly, while some positive associations have been reported relative to candidate genetic variants and exercise training response phenotypes for these four CVD disease risk factors, substantially more studies are necessary before any definitive conclusions can be drawn. It is also critically important to keep in mind that many of these results were generated from studies on very small sample sizes and, therefore, may be false-positive results.

**GENOMICALLY LINKAGE AND ASSOCIATION STUDIES OF CVD RISK FACTOR RESPONSES TO EXERCISE TRAINING**

Two HERITAGE GWL studies have addressed plasma lipoprotein-lipid responses to exercise training. In 2005, although Feitosa et al. (32) reported strong linkages for baseline LDL, they found no linkages for LDL responses to exercise training. In 2006 they also found no significant linkages with TG changes with training (34).

In 2002 Rice and HERITAGE co-workers (122) performed the first GWL study for resting BP changes with training; however, they also found no strong linkages. In 2001 Rankinen and colleagues (108) found one locus linked to the change in submaximal exercise systolic BP with training in whites and another locus linked with submaximal exercise diastolic BP change in blacks.

The first GWL for T2DM-related phenotype responses to exercise training was by Lakka and HERITAGE coworkers in 2003 (71). They found significant linkage in whites between training-induced changes in fasting insulin and a LEPR gene locus. In whites they found a number of other loci linked with fasting insulin responses to training, but in blacks only one such linkage was found (71). Lakka et al. (70) extended these results by typing a common LEPR SNP and three in the LEPR receptor (LEPR) gene relative to training-induced changes in fasting and frequently sampled intravenous glucose tolerance test (FSIVGTT) glucose and insulin measures. In this study the LEPR gene did not affect T2DM phenotype responses to training. However, LEPR variants affected the training-induced changes in glucose metabolism in whites, but not blacks. In a second HERITAGE GWL study, An et al. (6) found a number of linkages with the training-induced changes in different FSIVGTT phenotypes, although these linkages differed between blacks and whites.

HERITAGE has published the only GWL study for changes in obesity-related phenotypes with exercise training (22). In whites they found significant linkages with training-induced changes in fat mass at 1q31, percent body fat at 18q21-q23, and BMI at 5q14-q21.

Thus only seven studies have assessed GWL for CVD risk factor responses to training, but only four found any significant linkages. Also, no GWA studies have been done for the responses of these phenotypes to exercise training.

**SUMMARY OF THE GENETIC BASIS FOR CVD RISK FACTOR RESPONSES TO EXERCISE TRAINING**

Clearly at this time, no consistent and replicated results are available on which to base definitive conclusions relative to candidate genes that underlie the responses of CVD risk factors to exercise training. There is some evidence supporting “possible” candidate genes that may affect these responses. APO E and CETP variants, which impact baseline plasma lipoprotein-lipid profiles, may also affect their responses to exercise training. Relative to BP, four genes that also affect baseline BP may impact BP response to training: eNOS, ACE, EDN1, and GNB3. There is fairly strong evidence that a PPARG variant...
that alters T2DM risk also affects T2DM-related phenotype responses to training. For obesity-related phenotypes, the FTO gene, which affects obesity risk, also appears, along with BAR system genes, to affect obesity-related phenotype responses to training. Since we have no large-scale hypothesis-generating GWA studies relative to exercise training response phenotypes at this time, our “source” for candidate genes to investigate relative to exercise training responses are those already shown to be associated with baseline levels of the phenotype of interest. Thus it is not at all surprising, and in fact pretty much must be the case, that candidate gene variants found to be at least somewhat related to exercise training response phenotypes are also associated with baseline levels of that same phenotype.

The inability to generate robust conclusions at this time is due to 1) the relatively small number of available studies, 2) the huge number of potential candidate genes and genetic variants that need to be investigated, 3) the numerous possible physiological, clinical, and pathological phenotypes that need to be assessed, and 4) the generally small sample sizes and inconsistency of the results of most studies to date.

Finally, it is also important to keep in mind the publication bias toward positive results, with negative data, those showing no association between a variant and a phenotype in this case, being difficult to publish. However, even if there truly is no relationship between a genetic variant and a phenotype, we would still expect to find statistically significant false-positive relationships. However, in this case, these significant relationships should have an equal chance of being published, but the significant findings should be equally distributed across all genotype groups. Thus there should be an equal number of published statistically significant results indicating that a given phenotype changes with training to a greater extent in the common homozygote, the rare homozygote, and the heterozygote groups. While our current results are nowhere close to consistent relative to the relationships between specific genotypes and the training response phenotypes discussed in this review, in no case that I could find were the significant published findings equally distributed across genotype groups! Thus, in my estimation, this provides some evidence that there are “real” genotype × exercise training response phenotype relationships out there, but that the four issues listed in the previous paragraph are markedly hindering our efforts to detect them.

THE FUTURE?

Clearly, substantially more data are necessary before strong conclusions can be generated relative to genes underlying CVD risk factor responses to exercise training. In addition, these future studies must address the issue of statistical power and the resultant sample sizes required to robustly detect effect sizes that will very rarely exceed 1% of the interindividual variation in CVD risk factor responses to training.

CONCLUSIONS

To conclude, I want to return to the questions posed in the title of this review, albeit now in the context of CVD risk factors rather than CVD. The first question posed in the title is: do genetic variations alter the effects of exercise training on CVD risk factors? I very firmly believe that the available evidence clearly indicates that the answer to this question is an unequivocal “yes”—genetic variations do influence the degree to which these four CVD risk factors respond to exercise training, given all our knowledge on the variability and heritability of responses, and all of the evidence from candidate gene and HERITAGE GWL studies. The second question posed in the title is: Can we identify the genetic variants that influence the responses of these CVD risk factors to exercise training now? I also very firmly believe that there is an unequivocal answer to this question based on the information summarized above and that answer is “no.” We may have some hints, but no genetecist would lend very much credence at all to the level of evidence that has been accumulated for any of these potential candidate loci, even the PPARG Pro12Ala variant which has by far the most consistent and replicated results relative to affecting a response to exercise training. The third question posed in the title is: “can”, or will we have the capacity in the future, to identify these candidate variants? Based on the rapid advances in genetic technologies, such as large SNP chips and very shortly the possibility of sequencing an entire genome for $1,000 or less, and the statistical analysis of such large to huge data sets, I think the answer again is an unequivocal “yes,” that now, or at the worst in the near future, we will have the capacity to identify the underlying candidate variants, at least in terms of the genetic technology and statistical aspects of such a study.

Now to the fourth question posed in the title: in the future, will we be able to actually identify the variants affecting the responses of CVD disease risk factors to exercise training? Here we come face-to-face with the overwhelming issue that looms over such studies—sample size! To put this in the context of current genetic studies, a 2010 GWA study investigated BMI in ~250,000 individuals (yes—a quarter of a million!) (133). They initially typed or imputed 2.8 million SNPs in ~125,000 individuals and identified 19 loci associated with BMI. They then typed these SNPs plus additional candidates for a total of 42, in another ~125,000 individuals and confirmed highly significant associations with 32 of these SNPs. However, in applied terms, their SNP with the highest association (P < 5 × 10−120) (yes—an exponent of −120) accounted for only 0.34% of the interindividual variance in BMI and the 32 significant SNPs altogether accounted for a total of only 1.45% of the interindividual variance in BMI.

Such sample sizes are virtually beyond comprehension for an exercise training study when considering the absolute necessity to measure highly standardized phenotypes and to also use a highly standardized exercise training intervention. Large-scale, long-term prospective population studies have been proposed as a more viable means of assessing the interactive effects of exercise training (or physical activity) and genetic variations on phenotypes similar to those addressed in this review. However, it is my opinion that while such alternatives may be appealing to some, the issue in my mind remains one of highly standardized interventions and highly standardized phenotypes. Clearly if some individuals undergoing exercise training or a physical activity program adhere substantially better than others, and thus are exposed to a greater exercise training stimulus, their responses to that training intervention will most likely be greater, but the variation in the response will be due to a different environmental stimulus, not different genetic backgrounds. One solution to this issue in a large
prospective population study is to account for such variability in the exercise or physical activity stimulus in the analyses; however, currently techniques are not available to validly and reliably quantify the potentially small differences in exercise duration, frequency, and intensity that could impact the training response phenotypes. Using such a design is also substantially confounded by the fact that the genetic variants identified to date that may affect exercise training response phenotypes also significantly affect the baseline levels of that phenotype.

At present the largest highly standardized exercise training intervention study is HERITAGE and they “only” studied ~750 individuals, substantially short of the sample size probably required to detect the undoubtedly small effects of a large number of genes on these outcomes. Thus, at present, even being an optimist, it is difficult to see how these sample size issues can be overcome to generate consistent and replicated evidence quantifying the effects of individual candidate genes/SNPs on these critically important CVD risk factor responses to exercise training. However, the results of a number of very recent and very large GWA studies have succeeded in identifying a large number of candidate gene loci that are significantly associated with baseline levels of the four CVD risk factors addressed in this review. Perhaps these gene loci can be used to drive future hypothesis-driven exercise training intervention studies in smaller, more manageable, populations that will address the impact of specific genetic variants on well-defined physiological phenotypes.

DISCLOSURES

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

REFERENCES

Review

GENETICS OF CVD RISK FACTOR RESPONSES TO EXERCISE TRAINING


926 GENETICS OF CVD RISK FACTORS TO EXERCISE TRAINING

Review

Perusse L, Chagnon YC, Dionne FT, Bouchard C.

Pincus G, White P.


