J. Appl. Physiol.
88: 551–559, 2000.

Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE Family Study*

CLAUDE BOUCHARD,[1] TUOMO RANKINEN,[1] YVON C. CHAGNON,[2] TREVA RICE,[3] LOUIS PÉRUSSE,[2] JACQUES GAGNON,[2] INGRID BORECKI,[3] PING AN,[3] ARTHUR S. LEON,[4] JAMES S. SKINNER,[5] JACK H. WILMORE,[6] MICHAEL PROVINCE,[3] AND D. C. RAO[3,7]
[1]*Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, Louisiana 70808-4124;* [2]*Physical Activity Sciences Laboratory, Laval University, Québec, Canada G1K 7P4;* [3]*Division of Biostatistics and* [7]*Departments of Genetics and Psychiatry, Washington University Medical School, St. Louis, Missouri 63110;* [4]*School of Kinesiology and Leisure Studies, University of Minnesota, Minneapolis, Minnesota 55455;* [5]*Department of Kinesiology, Indiana University, Bloomington, Indiana 47405; and* [6]*Department of Health and Kinesiology, Texas A&M University, College Station, Texas 77843-4243*

Bouchard, Claude, Tuomo Rankinen, Yvon C. Chagnon, Treva Rice, Louis Pérusse, Jacques Gagnon, Ingrid Borecki, Ping An, Arthur S. Leon, James S. Skinner, Jack H. Wilmore, Michael Province, and D. C. Rao. Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE Family Study. *J. Appl. Physiol.* 88: 551–559, 2000.—This study aimed to identify human genomic regions that are linked to maximal oxygen uptake ($\dot{V}O_{2max}$) in sedentary individuals or to the responsiveness of $\dot{V}O_{2max}$ to a standardized endurance training program. The results of a genomic scan based on 289 polymorphic markers covering all 22 pairs of autosomes performed on the Caucasian families of the HERITAGE Family Study are presented. The mean spacing of the markers was 11 cM, and a total of 99 families and 415 pairs of siblings were available for the study. $\dot{V}O_{2max}$ in the sedentary state was adjusted for the effects of age, sex, body mass, fat mass, and fat-free mass, whereas the $\dot{V}O_{2max}$ response was adjusted for age and baseline level of the phenotype. Two analytic strategies were used: a single-point linkage procedure using all available pairs of siblings (SIBPAL) and a multipoint variance components approach using all the family data (SEGPATH). Results indicate that linkages at *P* values of 0.01 and better are observed with markers on 4q, 8q, 11p, and 14q for $\dot{V}O_{2max}$ before training and with markers on 1p, 2p, 4q, 6p, and 11p for the change in $\dot{V}O_{2max}$ in response to a 20-wk standardized endurance training program. These chromosomal regions harbor many genes that may qualify as candidate genes for these quantitative traits. They should be investigated in this and other cohorts.

genetic markers; quantitative trait locus; linkage; candidate genes

MAXIMAL OXYGEN uptake ($\dot{V}O_{2max}$) varies considerably among sedentary adults. In a recent report (3), we have shown that there is significant familial aggregation for $\dot{V}O_{2max}$ in the sedentary state even when the data are adjusted for age, sex, body mass, and body composition. These observations were derived from the HERITAGE Family Study, and they indicate that the heritability of $\dot{V}O_{2max}$ among sedentary adults after adjustment for the above covariates could be as high as 50% of the total phenotypic variance, although this value is undoubtedly inflated by nongenetic familial factors.

Twin and other family studies have been published in the peer-reviewed literature on this topic over the past three decades (6, 12, 17–21, 23, 31). Despite the fact that these studies vary considerably in sample size, phenotype measurements, data adjustment procedures, and analytic strategies, they generally concluded that there is a significant genetic component to $\dot{V}O_{2max}$ in the untrained state.

Moreover, there is evidence that the trainability of $\dot{V}O_{2max}$ is characterized by a significant level of familial aggregation. For instance, members of the same pairs of identical twins are significantly more alike than unrelated individuals in the $\dot{V}O_{2max}$ increase following exposure to a standardized training program. This was confirmed by the results of three different experimental studies with identical twins (16, 26, 29). The findings of these studies are remarkably concordant: the intraclass correlations for the intrapair resemblance in the $\dot{V}O_{2max}$ changes with training range from 0.65 to 0.77. The *F* ratios of the between-pair variance in $\dot{V}O_{2max}$ gain to the within-pair variance are quite similar, with a range from 6 to 9 (4). Such results have been further supported by the findings from the HERITAGE Family Study (1). The $\dot{V}O_{2max}$ response to the standardized training regimen of the HERITAGE Family Study exhibited familial aggregation, with some families characterized by a high-trainability pattern and others by low responsiveness. The maximal heritability of $\dot{V}O_{2max}$ trainability among the HERITAGE families of Caucasians reached 47% (1).

Ultimately, the nature of the genetic effects on human variation in $\dot{V}O_{2max}$ in the untrained state and in response to training will have to be resolved at the gene and molecular level. A handful of studies have consid-

*** Original submission in response to a special call for papers on "Molecular and Cellular Basis of Exercise Adaptations."**

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

ered the role of a few candidate genes on $\dot{V}O_{2max}$ in sedentary adults or $\dot{V}O_{2max}$ responsiveness to training (7) with mixed results. Another approach is to focus on the detection of quantitative trait loci (QTL) affecting these phenotypes. In a preliminary paper based on this strategy, we have reported the results of a first screen of chromosome 22 using seven polymorphic markers (13). No evidence for a QTL affecting $\dot{V}O_{2max}$ in the sedentary state or the changes in $\dot{V}O_{2max}$ with training was uncovered.

Here, we report the results of a genomic scan, based on 289 markers covering all autosomes, performed on the Caucasian families of the HERITAGE Family Study with a view to identify the genomic regions encoding genes contributing to human heterogeneity in $\dot{V}O_{2max}$ in the untrained state and its trainability.

MATERIALS AND METHODS

Sample

The HERITAGE Family Study was designed to investigate the role of the genotype in cardiovascular, metabolic, and hormonal responses to aerobic exercise training and the contribution of regular exercise to changes in selected cardiovascular disease and diabetes risk factors. Five centers located at Indiana University, Pennington Biomedical Research Center, University of Minnesota, Texas A&M University, and Washington University are presently involved in the HERITAGE Family Study consortium. The study design, sample, and protocol have been described earlier (5).

A total of 481 individuals from 99 two-generation families of Caucasian descent (236 men, 245 women) were available for this study. The following criteria were applied to screen subjects for participation. First, individuals were required to be between the ages of 17 and 65 yr (17–40 yr of age for offspring and ≤65 yr of age for parents). Second, all participants were required to be sedentary at baseline. Third, individuals with a body mass index >40 kg/m^2 were excluded unless they were able to meet the demands of the exercise tests and exercise training program. Fourth, resting blood pressure levels could not exceed 159 mmHg for systolic and 99 mmHg for diastolic. Antihypertensive drug therapy was also a cause for exclusion. Participants were required to be in good physical health and able to complete the 20-wk exercise program. Further details about the study, including inclusion and exclusion criteria, can be found in Bouchard et al. (5). Informed written consent was obtained from all subjects. The study was approved by the Institutional Review Board of each participating institution.

Exercise Training Program

The training program was conducted on cycle ergometers (Universal Aerobicycle, Cedar Rapids, IA) interfaced with a Mednet computer system (Universal Gym Mednet, Cedar Rapids, IA) to control the power output of the ergometers so that constant training heart rates could be maintained. Subjects started training at a heart rate associated with 55% of the initial $\dot{V}O_{2max}$ for 30 min/day and gradually progressed to 75% of the heart rate of the initial $\dot{V}O_{2max}$ for 50 min/day at the end of 14 wk. They maintained this intensity and duration throughout the remaining 6 wk. Frequency was maintained at three sessions per week throughout the 20-wk training program. The power output of the cycle ergometer was adjusted automatically to the heart rate response of the subject at all times during all training sessions. All training sessions were supervised on site. A detailed description of the training program can be found elsewhere (5, 30).

Maximal Oxygen Uptake Measurements

Each individual was examined with a battery of measurements before and after the 20-wk standardized exercise program. Two maximal exercise tests designed to lead to $\dot{V}O_{2max}$ on a cycle ergometer 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 connected to a SensorMedics 2900 metabolic measurement cart. The tests were conducted at about the same time of day, with at least 48 h between the two tests. An electrocardiogram was used to monitor heart rate. Gas exchange variables [O_2 consumption ($\dot{V}O_2$), CO_2 production ($\dot{V}CO_2$), minute ventilation, and respiratory exchange ratio (RER)] were recorded as a rolling average of three 20-s intervals. The criteria for $\dot{V}O_{2max}$ were RER >1.1, plateau in $\dot{V}O_2$ (change of <100 ml/min in the last three 20-s intervals), and a heart rate within 10 beats/min of the maximal heart rate predicted for that age. All subjects achieved a $\dot{V}O_{2max}$ by at least one of these criteria in at least one of the two tests, both pre- and posttraining. In the first test, subjects exercised at a power output of 50 W for 3 min, followed by increases of 25 W each every 2 min until volitional exhaustion. For older, smaller, or less fit individuals, who were generally the older mothers among the family members, the test was started at 40 W, with increases of 10–20 W each every 2 min thereafter. In the second test, subjects exercised for ~10 min at an absolute (50 W) and then for ~10 min at a relative power output equivalent to 60% $\dot{V}O_{2max}$. They then exercised for 3 min at a relative power output that was 80% of their $\dot{V}O_{2max}$, after which resistance was increased to the highest power output attained in the first maximal test. If subjects were able to pedal after 2 min, power output was increased every 2 min thereafter until they reached volitional fatigue.

If both values were within 5% of each other, the average $\dot{V}O_{2max}$ from these two tests was taken as the $\dot{V}O_{2max}$ for that subject and used in this analysis. If they differed by more than 5%, the higher $\dot{V}O_{2max}$ value was used. The reproducibility of $\dot{V}O_{2max}$ in these subjects was examined and was characterized by an intraclass correlation coefficient of 0.97 for repeated tests with a coefficient of variation of 5% and no difference among clinical centers (3, 30). The $\dot{V}O_{2max}$ response was defined as the difference (ml O_2/min) between posttraining $\dot{V}O_{2max}$ and baseline $\dot{V}O_{2max}$ (i.e., $\dot{V}O_{2max}$ response = posttraining $\dot{V}O_{2max}$ minus baseline $\dot{V}O_{2max}$).

Molecular Studies

Genomic DNA was prepared from permanent lymphoblastoid cells by the proteinase K and phenol/chloroform technique. DNA was dialyzed four times against 10 mM Tris-1 mM EDTA (pH 8.0) buffer for 6 h at 4°C, and ethanol was precipitated.

Choice of markers. Microsatellite markers (di-, tri-, and tetranucleotide repeat) were selected from different sources but mainly from the Marshfield panel version 8a (http://www.marshmed.org/genetics) and the Location Database (LDB) maps (January 1999 version) from Southampton, UK (http://cedar.genetics.soton.ac.uk). The LDB summary map was also used to look for candidate genes 10 cM on each side of the markers yielding the strongest results. The mean spacing for the 289 markers was 11 cM with a range from 0.2 to 51 cM.

PCR conditions. PCR conditions and genotyping have been described in detail elsewhere (9). Briefly, PCR reactions were conducted using 250 ng of genomic DNA, 0.1 pmol of the forward primer coupled to the infrared tag IRD800 or IRD700

(LICOR), and 0.4 pmol of the reverse untagged primer (Research Genetics), 125 μM dNTPs, and 0.3 units *Taq* polymerase (Perkin-Elmer or Pharmacia) in PCR buffer (100 mM Tris·HCl, pH 8.3, 15 mM $MgCl_2$, 0.5 M KCl, and 0.01% gelatin) for a final volume of 10 μl. PCR cycles followed a two-step procedure: 1 cycle at 93°C for 5 min, 10 cycles at 93°C for 20 s and 57°C for 60 s, and 24 cycles at 93°C for 20 s and 52°C for 60 s. Annealing temperature was defined according to the melting temperature of the primers used and ranged between 52 and 67°C, with a difference of 5°C being maintained between the two annealing temperatures.

Marker analysis. Automatic infrared DNA sequencers from LICOR were used to detect the PCR products. One microliter of each of the 52 samples, including up to five different markers, spaced with standards at each of the four samples, was applied to the gel. Standards were produced by PCR from a PUC19 plasmid using a tagged M13 forward primer (LICOR) and corresponding untagged reverse primer for each standard. The 18-cm long gel was run for 1.5–2 h. At the end of the run, an electronic image of the gel was produced and used for genotyping.

Genotyping. Automatic genotyping was performed using the computer software SAGA (Rick McIndoe, Roger Bumgarner, and Russ Welti, University of Washington, Seattle, WA; LICOR). In this process, sample and standard lanes were automatically found, the different markers were located on the gel, bands for each sample were identified, and genotyping was done simultaneously for a given marker on all the subjects of the study. After a manual edition of the typing, results were exported directly to a local dBase IV database (GENEMARK) in which a procedure had been developed to check for misinheritance (incompatibilities among the genotypes) within nuclear families and extended pedigrees. Subjects with incompatibilities were excluded from the database by GENEMARK and retyped completely for the incompatible markers, i.e., from the PCR reaction to the genotyping. Less than 5% of the subjects had to be retyped (mainly because of insufficient DNA amplification).

Data Adjustment

Two $\dot{V}O_{2max}$ phenotypes are considered. $\dot{V}O_{2max}$ at baseline was analyzed after adjustment for the effects of age (up to a cubic polynomial), sex, body mass, fat mass, and fat-free mass using stepwise multiple regression procedures, separately in each of the four sex-by-generation groups, and retaining only those terms that were significant at the 5% level. Fat mass and fat-free mass were estimated from an underwater weighing assessment of body density as recently described (33). The resulting squared residuals were similarly regressed on another polynomial in age (to adjust for heterocedacticity). The phenotype used in the genetic analysis was thus adjusted for relevant covariate effects in both the mean and variance. The phenotype was standardized to a mean of zero and a standard deviation of one. The percentage of variance accounted for by these covariates was quite large, ranging from 42.2% to 57.6%.

The response of $\dot{V}O_{2max}$ to training was adjusted for age and baseline $\dot{V}O_{2max}$ using a stepwise multiple regression procedure. Briefly, the response variable was regressed on up to a cubic polynomial in age plus baseline value within four sex-by-generation groups (fathers, mothers, sons, and daughters). Only terms significant at the 5% level were retained. The resulting squared residuals were similarly adjusted for effects on the variance; the final adjusted phenotype was standardized to a mean of zero and a standard deviation of one. Significant terms (percentage of variance) in each of the sex-by-generation groups for the $\dot{V}O_{2max}$ response phenotype were seen only in sons (9.2%) and daughters (5.0%). SAS (version 6.08) for PC was used for the regression analysis.

Statistical Analysis

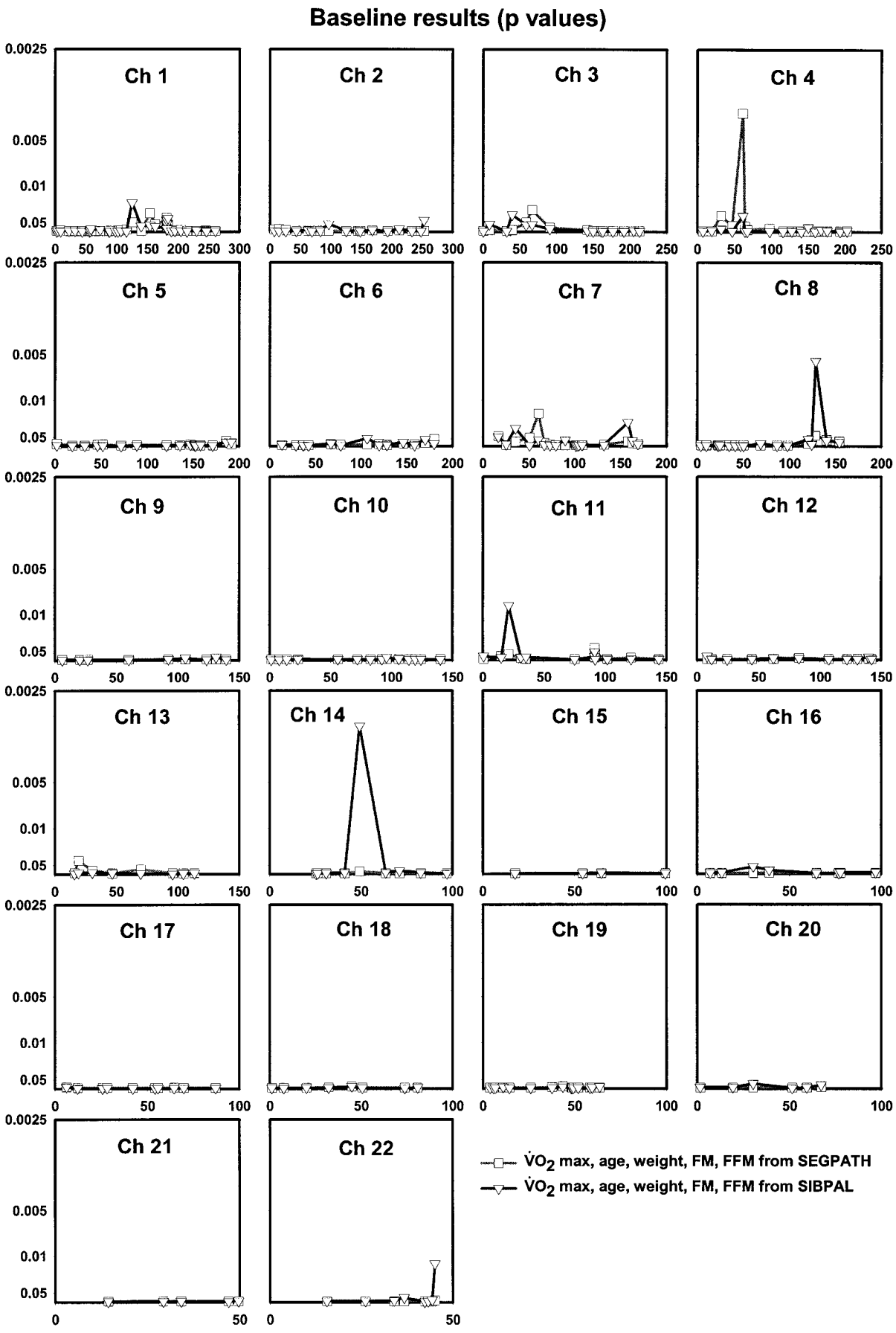
First, single-point linkage analysis using all available sibling pairs was performed on these two phenotypes (SIBPAL; SAGE version 3.0). Second, analyses were carried out using the multipoint variance components analytic strategy implemented in SEGPATH (25). This is an extension of a path model for analysis of family resemblance (24) in which correlations among family members are modeled as functions of the allele sharing at a marker locus and a residual component including genetic as well as shared environmental effects. In addition to the trait locus heritability (h_g^2) and the residual heritability (h_r^2), the spouse correlation and excess sibling resemblance beyond that predicted under the genetic model were estimated by maximum likelihood. Linkage was tested by the likelihood ratio test of the null hypothesis $h_g^2 = 0$. Multipoint estimation of identical by descent was used in the SEGPATH analysis, whereas a single-point approach was implemented in SIBPAL. The number of families available for the SEGPATH analysis was 99. The SIBPAL linkage analyses were undertaken with 415 pairs of siblings at baseline and 327 for the response to training. Suggestive evidence of linkage was defined as $P < 0.01$, whereas weaker but potentially useful evidence of linkage was identified as $0.01 < P < 0.05$.

RESULTS

General trends in the linkage results are depicted in Figs. 1 and 2 for the 22 autosomes. In Fig. 1, the evidence for linkage (*P* values) from SEGPATH and SIBPAL is illustrated for the baseline $\dot{V}O_{2max}$ (adjusted for age, body weight, fat mass, and fat-free mass within each sex). Suggestive evidence of linkage can be seen for chromosomes 4, 8, 11, and 14 ($P < 0.01$). The training response (adjusted for age and baseline level by sex) results are shown in Fig. 2. Again, suggestive evidence of linkage is observed for chromosomes 1, 2, 4, 6, and 11.

The details of all suggestive linkages for the pretraining $\dot{V}O_{2max}$ are summarized in Table 1. The region with the highest significance level with SEGPATH is located on the long arm of chromosome 4 (4q12) at map position 61.658 cM. Three additional regions for baseline $\dot{V}O_{2max}$ were identified with SIBPAL. These regions are located on chromosome 14q21.3 at map position 49.255 cM, chromosome 8q24.12 (128.157 cM), and chromosome 11p15.1 (21.174cM). Moreover, 11 regions on seven chromosomes (chromosomes 1, 3, 4, 7, 8, 11, 13) with SEGPATH and 10 markers on six chromosomes (chromosomes 1, 2, 3, 4, 7, 22) with SIBPAL showed weaker but possibly useful evidence of linkage ($0.01 < P < 0.05$) with $\dot{V}O_{2max}$ in the sedentary state (Table 2).

Multipoint analyses (SEGPATH) revealed two regions and single-point analyses (SIBPAL) revealed three regions with suggestive evidence of linkage with the $\dot{V}O_{2max}$ training response. The multipoint signals are located on chromosomes 4 (4q26, 127.793 cM) and 6 (6p21.33, 28.788 cM), whereas the single-point linkages are seen with markers on chromosomes 11 (11p14.1, 31.272 cM), 1 (1p11.2, 125.032 cM), and 2 (2p16.1, 62.602 cM). In addition, weaker but potentially useful linkages ($0.01 < P < 0.05$) were found with SEGPATH


Baseline results (p values)
Ch 1
Ch 2
Ch 3
Ch 4
Ch 5
Ch 6
Ch 7
Ch 8
Ch 9
Ch 10
Ch 11
Ch 12
Ch 13
Ch 14
Ch 15
Ch 16
Ch 17
Ch 18
Ch 19
Ch 20
Ch 21
Ch 22
0.0025
0.005
0.01
0.05
VO2 max, age, weight, FM, FFM from SEGPATH
VO2 max, age, weight, FM, FFM from SIBPAL

for regions located on chromosomes 3, 6, 9, and 13 plus for 15 markers on chromosomes 1, 3, 5, 6, 7, 11, 13, 15, 19, and 21 with SIBPAL (Table 3).

DISCUSSION

From the quantitative genetic evidence gathered in twin and family studies, it is justified to search for QTLs and, eventually, genes and mutations pertaining to $\dot{V}O_{2max}$ in the sedentary state and its responsiveness to training. The genetic dissection of these two phenotypes and their determinants will require a wide array of designs and technologies. At this time, only one linkage study has been reported in humans, but it has dealt with a limited number of markers on one chromosome only (13). Other studies have focused on candidate genes encoded in the nuclear DNA and on a few mitochondrial DNA sequences.

In the same population, we have earlier reported a maximal heritability of ~50% for the baseline $\dot{V}O_{2max}$ data (3). On the other hand, the maximal heritability estimate of the $\dot{V}O_{2max}$ response to training reached 47% for the same sample of families with 2.5 times more variance between families than within families. Thus the familial factors underlying $\dot{V}O_{2max}$ in sedentary families are quantitatively similar to those underlying its response to exercise training. However, although they are quantitatively about the same, the familial and genetic factors underlying the two phenotypes appear to be different, as indicated by the noncomparability of the linkage results for baseline $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ response. From the earlier observations on identical twins and nuclear families plus the recent data on the HERITAGE Family Study cohort, we previously concluded that $\dot{V}O_{2max}$ in the sedentary state and its trainability are highly familial with a significant genetic component (1, 3). It should therefore be possible to identify the genes and mutations responsible for the variability in both phenotypes. However, it must be recognized at the outset that it is unlikely that a single gene or very few loci will be sufficient to define the genetic component of $\dot{V}O_{2max}$ and trainability.

A few reports on the topic of genetic markers and performance have appeared over the past 30 years or so. The study of elite performance by means of genetic markers was first conducted during the 1968 Olympic Games in Mexico City (11). The purpose of that study was to test if there was any association between participation in the Olympic Games as an athlete and allelic variation in single-gene blood systems. Results indicated that participation in the 1968 Olympic Games was not associated with allelic variations in red blood cell antigens or enzyme variants of red blood cells. A second effort was carried out during the 1976 Olympic Games in Montreal (8, 10). No significant differences between Caucasian elite endurance athletes and controls were observed for genetic markers in red blood cell antigens and four red blood cell enzymes.

Since then, a handful of papers have dealt with the potential contribution of a few candidate genes. For instance, three studies have reported an association between the insertion allele of the angiotensin-converting enzyme (*ACE*) insertion/deletion (I/D) polymorphism and indicators of performance (14, 15, 22). The *ACE* locus is encoded on 17q23. In the present study, no evidence of linkage was observed on chromosome 17 for either baseline $\dot{V}O_{2max}$ or its responsiveness to training. Therefore, it is not yet clear whether the *ACE* I/D polymorphism is truly of importance for physical performance in general or endurance performance in particular.

Skeletal muscle-specific creatine kinase (*CKMM*) is a legitimate candidate gene (19q13.22) to investigate in relation to endurance performance. Because *CKMM* activity level is two times greater in type II (fast-twitch) than in type I fibers (34), a low *CKMM* activity level is typical of the skeletal muscle of endurance athletes. An early study indicated that a *CKMM* protein charge variant was weakly associated with the ability to perform a 90-min endurance test (2). In addition, research on transgenic mice indicates that a low *CKMM* activity is associated with improved skeletal muscle resistance to fatigue (32). More recently, a sibling pair linkage study has shown a weak genetic linkage between the *CKMM* locus and changes in $\dot{V}O_{2max}$ (age, sex, and pretraining $\dot{V}O_{2max}$ adjusted) in the HERITAGE Family Study (28). This linkage was suggested again in the present study with a *P* value of 0.026 (Table 3). Moreover, we have reported a significant association between the *CKMM* genotype and the $\dot{V}O_{2max}$ response to 20 wk of endurance training in both parents and adult offspring of the HERITAGE Family Study (27). One-third of all homozygotes for the less-frequent allele (*CKMM Nco* I polymorphism in the 3′ untranslated region) were observed in the low-responder group (lowest decile of response), whereas this genotype was not seen in any high responders (upper decile of response). The *CKMM* genotype accounted for ~10% of the variance of $\dot{V}O_{2max}$ response.

The chromosomal regions showing suggestive linkages with $\dot{V}O_{2max}$ in the sedentary state or in response to endurance training in the present paper encode several potential candidate genes. The marker D4S3248 on chromosome 4, which was the strongest signal with the baseline $\dot{V}O_{2max}$ in multipoint analyses, is located close (0.2 cM) to the β-sarcoglycan gene. β-Sarcoglycan is part of the dystrophin-glycoprotein complex, which acts as a structural link between the cytoskeleton of muscle and the extracellular matrix. It is thought to confer stability to the sarcolemma and protect muscle cells from contraction-induced damage. A gene of another dystrophin-associated protein, syntrophin β-1, located 5.9 cM from the D8S592 marker, showed suggestive single-point linkage evidence with the baseline $\dot{V}O_{2max}$. In addition, the genes encoding

Fig. 1. Overview of the linkage results for the 22 autosomes (Ch1 to Ch22) based on SEGPATH and SIBPAL for maximal O_2 uptake ($\dot{V}O_{2max}$) in the sedentary state. Map distances in centimorgans (cM) are shown on the abscissa, and *P* values for the strength of the linkage relationships are shown on the ordinates. FM, fat mass; FFM, fat-free mass.

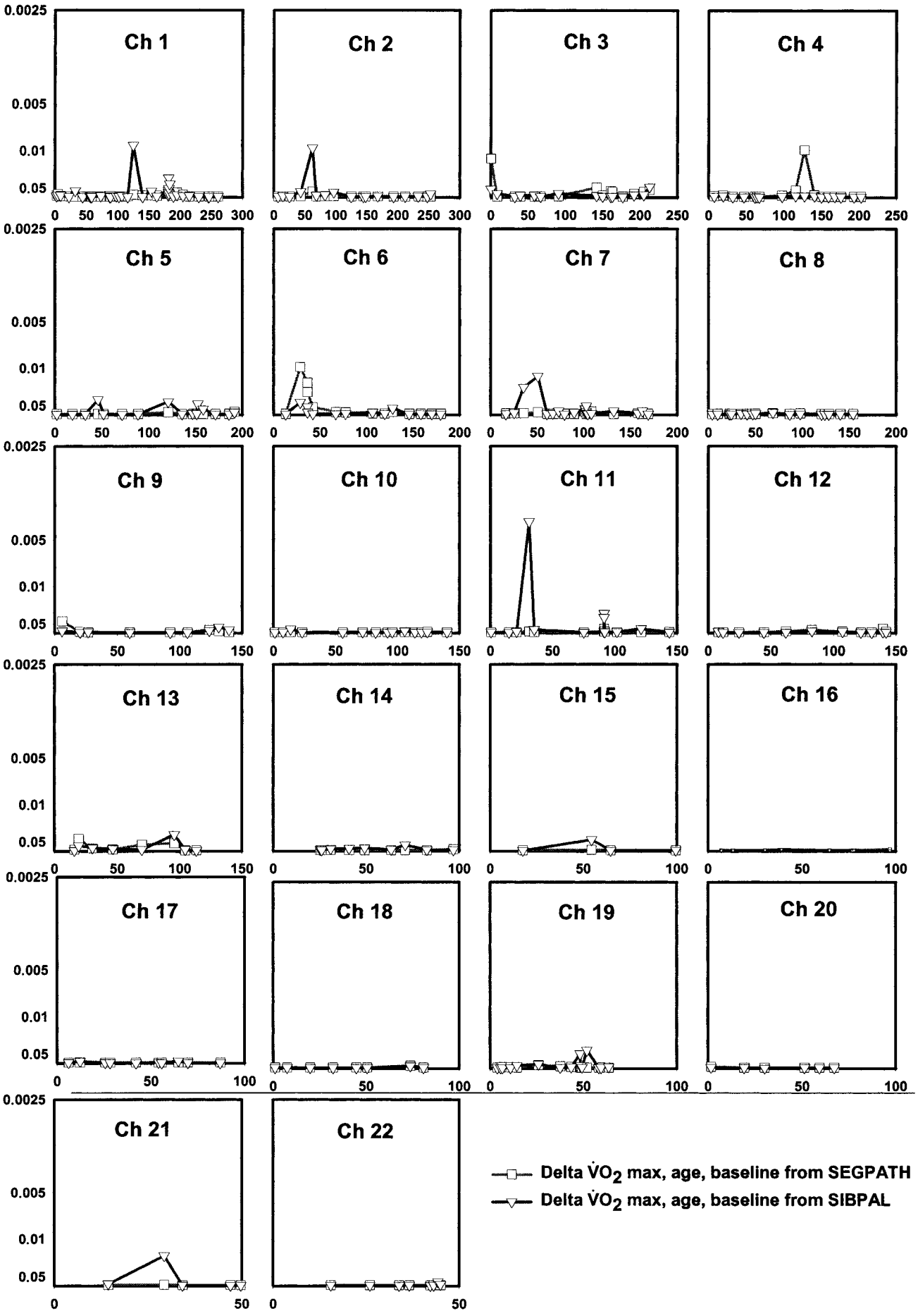

Training response results (p values)
Ch 1
Ch 2
Ch 3
Ch 4
Ch 5
Ch 6
Ch 7
Ch 8
Ch 9
Ch 10
Ch 11
Ch 12
Ch 13
Ch 14
Ch 15
Ch 16
Ch 17
Ch 18
Ch 19
Ch 20
Ch 21
Ch 22
0.0025
0.005
0.01
0.05
Delta $\dot{V}O_2$ max, age, baseline from SEGPATH
Delta $\dot{V}O_2$ max, age, baseline from SIBPAL

Table 1. *Summary of the suggestive linkages ($P < 0.01$) with $\dot{V}O_{2max}$ in the sedentary state and in response to endurance training in the HERITAGE Family Study*

P Value	Chromosome	Marker	Map Position, cM
	Sedentary state		
SEGPATH			
0.0038	4q12	D4S3248	61.658
SIBPAL			
0.0054	8q24.12	D8S592	128.157
0.0084	11p15.1	*SUR*	21.174
0.0031	14q21.3	D14S587	49.255
	Training response		
SEGPATH			
0.0098	4q26	FABP2	127.793
0.0098	6p21.33	D6S2439	28.788
SIBPAL			
0.0090	1p11.2	D1S534	125.032
0.0095	2p16.1	D2S2739	62.602
0.0042	11p14.1	ATA34E08	31.272

$\dot{V}O_{2max}$, maximal oxygen uptake; cM, centimorgans.

γ-sarcoglycan (13q12.11), dystrophin-associated glycoprotein 1 (3p21.31), and lamin A/C (1q21.2) are located within 2.3 to 6.2 cM from markers, showing weak but potentially useful linkages with $\dot{V}O_{2max}$ in the sedentary state. The marker (D14S587) showing the strongest evidence of linkage with the baseline $\dot{V}O_{2max}$ in SIBPAL is located in the vicinity of the liver glycogen phosphorylase (0.6 cM) and the GTP cyclohydrolase I (1.8 cM) genes. GTP cyclohydrolase I is a rate-limiting enzyme in the synthesis of tetrahydrobiopterin, an essential cofactor of nitric oxide synthase.

A restriction fragment length polymorphism marker on chromosome 11p15.1 within the sulfonylurea receptor (*SUR*) gene exhibited one of the most significant linkages with $\dot{V}O_{2max}$ in the sedentary state. *SUR* is expressed in pancreatic β-cells where it forms ATP-sensitive potassium channels together with an inward rectifier potassium channel member, Kir6.2, and is thereby involved in the regulation of insulin secretion. One cannot conclude from this linkage result that *SUR* itself has any effect on $\dot{V}O_{2max}$ variability. An alternative explanation could be that the *SUR* marker is a surrogate for a relationship with another gene such as Kir6.2 (*KCNJ11*), which is located close to the *SUR* locus. Unlike *SUR*, Kir6.2 is expressed in several tissues, including heart and skeletal muscle, where it plays a role in the coupling of cell metabolism to membrane potential.

Potential candidate genes located close to markers with suggestive linkages to $\dot{V}O_{2max}$ training response include the voltage-gated potassium channel gene (*KCNA4*) on chromosome 11p14.1 (0.1 cM from the marker), the pancreatic colipase (*CLPS*, 0.8 cM) gene and the hemochromatosis locus (*HFE*, 1.3 cM) on chromosome 6p21.33, the fatty-acid binding protein 2 (*FABP2*, 0 cM) gene and the long QT syndrome 4 locus (*LQT4*, 4.0 cM) on chromosome 4q26, the calmodulin 2 (*CALM2*, 0.5 cM) and the calcineurin B (*PPP3R1*, 3.4 cM) genes on chromosome 2p16.1, and the 3-β-hydrox-

Table 2. *Summary of the potentially useful linkages ($0.01 < P < 0.05$) with $\dot{V}O_{2max}$ in the sedentary state in the HERITAGE Family Study*

P Value	Chromosome	Marker	Map Position, cM
SEGPATH			
0.047	1p11.2	D1S534	125.032
0.025	1q21.2	S100A1	154.041
0.033	1q24.1	ATP1B1	181.137
0.021	3p21.1	D3S1766	66.879
0.048	3p21.31	D3S1447	58.763
0.028	4p14	D4S2397	32.246
0.014	7p11.2	IGFBP1	60.720
0.047	7p21.3	D7S513	16.740
0.043	8q24.12	D8S592	128.157
0.038	11q14.1	UCP2	91.588
0.033	13q12.11	D13S787	19.370
SIBPAL			
0.016	1p11.2	D1S534	125.032
0.039	1q24.1	ATP1A2	181.720
0.035	1q24.2	D1S1677	183.512
0.039	2q37.3	D2S2968	253.130
0.027	3p22.1	D3S1768	39.702
0.028	4q12	D4S3248	61.658
0.026	7p14.3	D7S2541	35.660
0.049	7p21.3	D7S513	16.740
0.019	7q36.1	D7S2195	158.000
0.012	22q13.2	D22S274	45.091

Table 3. *Summary of the potentially useful linkages ($0.01 < P < 0.05$) with $\dot{V}O_{2max}$ response to endurance training in the HERITAGE Family Study*

P Value	Chromosome	Marker	Map Position, cM
SEGPATH			
0.012	3p26.3	D3S2387	0.335
0.047	3q21.3	D3S3023	142.000
0.014	6p21.31	TNF	37.805
0.020	6p21.31	HSPA1A	36.701
0.039	9p24.1	GATA62F03	6.084
0.038	13q12.11	D13S787	19.370
SIBPAL			
0.024	1q24.1	ATP1A2	181.720
0.035	1q24.2	D1S1677	183.512
0.044	3q29	D3S1311	213.050
0.031	5p12	D5S1470	46.124
0.036	5q22.2	D5S1453	120.870
0.043	5q31.3	D5S436	152.500
0.038	6p21.33	D6S2439	28.788
0.012	7p12.3	IGFBP3	50.892
0.017	7p14.3	D7S2541	35.660
0.024	11q14.1	UCP2	91.588
0.028	13q32	D13S793	95.754
0.041	15q21.1	D15S659	54.511
0.032	19q13.2	TGFB1	48.290
0.026	19q13.32	CKM	52.052
0.015	21q21.3	D21S1437	29.270

Fig. 2. Overview of the linkage results for the 22 autosomes based on SEGPATH and SIBPAL for the $\dot{V}O_{2max}$ response to a standardized training program. Map distances in cM are shown on the abscissa, and *P* values for the strength of the linkage relationships are shown on the ordinates.

ysteroid dehydrogenase (*HSD3B1*, 0.1 cM) and the cardiac muscle calsequestrin (*CASQ2*, 5.5 cM) genes on chromosome 1p11.2. These genes are involved in cardiac contractility (*KCNA4*, *LQT4*), long-chain fatty acid absorption (*CLPS*, *FABP2*), calcium homeostasis and signaling in skeletal and cardiac muscle (*CALM2*, *PPP3R1*, *CASQ2*), and steroid hormone synthesis (*HSD3B1*). Although there is no direct evidence to support the notion that they are involved in human trainability, one could hypothesize that they contribute to interindividual variation in training response.

There was not a high degree of concordance between the single point and multipoint linkage results in the present study. The differences in the findings are likely explained by the relatively small sample sizes available for such studies, the marker spacing, which is relatively large in some chromosomal areas, and perhaps misspecification of the location of markers on the genetic map. The latter two factors could have had influences on the multipoint linkage results, particularly if errors of location exist in the current map, but would not have affected the single-point findings. Moreover, it should be recognized that, even though a panel of 289 markers is reasonably large for a first genomic screen, a more dense and accurate map will likely yield more and perhaps stronger signals in the future.

In summary, multipoint and single point linkage studies performed on the Caucasian families of the HERITAGE Family Study have revealed some indications of linkages with $\dot{V}O_{2max}$ in the sedentary state and with its responsiveness to training. None of the linkages is very strong. However, several linkages were found at the 0.01 level and better with markers on 4q, 8q, 11p, and 14q for $\dot{V}O_{2max}$ before training and with markers on 1p, 2p, 4q, 6p, and 11p for the gain in $\dot{V}O_{2max}$ in response to 20 wk of a standardized endurance training program. These chromosomal regions encode several candidate genes that should be investigated further.

Thanks are expressed to all the coprincipal investigators, investigators, coinvestigators, local project coordinators, research assistants, laboratory technicians, and secretaries who have contributed to the study and to Diane Drolet, MSc, for assistance with the preparation of this manuscript.

The HERITAGE Family Study is supported by the National Heart, Lung, and Blood Institute through the following grants: HL-45670 (C. Bouchard, principal investigator), HL-47323 (A. S. Leon, principal investigator), HL-47317 (D. C. Rao, principal investigator), HL-47327 (J. S. Skinner, principal investigator), and HL-47321 (J. H. Wilmore, principal investigator). A. S. Leon is partially supported by the Henry L. Taylor Professorship in Exercise Science and Health Enchancement.

Address for reprint requests and other correspondence: C. Bouchard, Pennington Biomedical Research Center, Louisiana State Univ., 6400 Perkins Rd., Baton Rouge, LA 70808-4124 (E-mail: bouchac@pbrc.edu).

Received 31 August 1999; accepted in final form 29 October 1999.

REFERENCES

1. **Bouchard, C., P. An, T. Rice, J. S. Skinner, J. H. Wilmore, J. Gagnon, L. Pérusse, A. S. Leon, and D. C. Rao.** Familial aggregation of $\dot{V}O_{2max}$ response to exercise training: results from the HERITAGE Family Study. *J. Appl. Physiol.* 87: 1003–1008, 1999.
2. **Bouchard, C., M. Chagnon, M. C. Thibault, M. R. Boulay, M. Marcotte, C. Côté, and J. A. Simoneau.** Muscle genetic variants and relationship with performance and trainability. *Med. Sci. Sports Exerc.* 21: 71–77, 1989.
3. **Bouchard, C., E. W. Daw, T. Rice, L. Pérusse, J. Gagnon, M. A. Province, A. S. Leon, D. C. Rao, J. S. Skinner, and J. H. Wilmore.** Familial resemblance for $\dot{V}O_{2max}$ in the sedentary state: the HERITAGE Family Study. *Med. Sci. Sports. Exerc.* 30: 252–258, 1998.
4. **Bouchard, C., F. T. Dionne, J. A. Simoneau, and M. R. Boulay.** Genetics of aerobic and anaerobic performances. *Exerc. Sport Sci. Rev.* 20: 27–58, 1992.
5. **Bouchard, C., A. S. Leon, D. C. Rao, J. S. Skinner, J. H. Wilmore, and J. Gagnon.** The HERITAGE Family Study. Aims, design, and measurement protocol. *Med. Sci. Sports. Exerc.* 27: 721–729, 1995.
6. **Bouchard, C., R. Lesage, G. Lortie, J. A. Simoneau, P. Hamel, M. R. Boulay, L. Pérusse, G. Thériault, and C. Leblanc.** Aerobic performance in brothers, dizygotic and monozygotic twins. *Med. Sci. Sports Exerc.* 18: 639–646, 1986.
7. **Bouchard, C., B. Wolfarth, M. A. Rivera, J. Gagnon, and J. A. Simoneau.** Genetic determinants of endurance performance. *Endurance in Sport* (2nd ed.), edited by R. J. Shephard. Oxford, UK: Blackwell Science. In press.
8. **Chagnon, Y. C., C. Allard, and C. Bouchard.** Red blood cell genetic variation in Olympic endurance athletes. *J. Sports Sci.* 2: 121–129, 1984.
9. **Chagnon, Y. C., S. Roy, M. Chagnon, M. Lacaille, C. Leblanc, and C. Bouchard.** High-throughput genotyping using infrared automatic Li-Cor DNA sequencers in the study of the obesity and co-morbidities genes. *LiCor Application Note 500*. Lincoln, NE; Li-Cor, 1998.
10. **Couture, L., M. Chagnon, C. Allard, and C. Bouchard.** More on red blood cell genetic variation in Olympic athletes. *Can. J. Appl. Sport Sci.* 11: 16–18, 1986.
11. **DeGaray, A., L. Levine, and J. E. L. Carter.** *Genetic and Anthropological Studies of Olympic Athletes.* New York: Academic, 1974.
12. **Fagard, R., E. Bielen, and A. Amery.** Heritability of aerobic power and anaerobic energy generation during exercise. *J. Appl. Physiol.* 70: 352–362, 1991.
13. **Gagnon, J., M. A. Ho-Kim, Y. C. Chagnon, L. Pérusse, F. T. Dionne, A. S. Leon, D. C. Rao, J. S. Skinner, J. H. Wilmore, and C. Bouchard.** Absence of linkage between $\dot{V}O_{2max}$ and its response to training with markers spanning chromosome 22. *Med. Sci. Sports Exerc.* 29: 1448–1453, 1997.
14. **Gayagay, G., B. Yu, B. Hambly, T. Boston, A. Hahn, D. S. Celermajer, and R. J. Trent.** Elite endurance athletes and the ACE I allele: the role of genes in athletic performance. *Hum. Genet.* 103: 48–50, 1998.
15. **Hagberg, J. M., R. E. Ferrell, S. D. McCole, K. R. Wilung, and G. E. Moore.** $\dot{V}O_{2max}$ is associated with ACE genotype in postmenopausal women. *J. Appl. Physiol.* 85: 1842–1846, 1998.
16. **Hamel, P., J. A. Simoneau, G. Lortie, M. R. Boulay, and C. Bouchard.** Heredity and muscle adaptation to endurance training. *Med. Sci. Sports. Exerc.* 18: 690–696, 1986.
17. **Klissouras, V.** Heritability of adaptive variation. *J. Appl. Physiol.* 31: 338–344, 1971.
18. **Klissouras, V., F. Pirnay, and J. M. Petit.** Adaptation to maximal effort: genetics and age. *J. Appl. Physiol.* 35: 288–293, 1973.
19. **Lesage, R., J. A. Simoneau, J. Jobin, J. Leblanc, and C. Bouchard.** Familial resemblance in maximal heart rate, blood lactate and aerobic power. *Hum. Hered.* 35: 182–189, 1985.
20. **Lortie, G., C. Bouchard, C. Leblanc, A. Tremblay, J. A. Simoneau, G. Thériault, and J. P. Savoie.** Familial similarity in aerobic power. *Hum. Biol.* 54: 801–812, 1982.
21. **Maes, H. H. M., G. P. Beunen, R. F. Vlietinck, M. C. Neale, M. Thomis, B. Vanden Eynde, R. Lysens, J. Simons, and C. Derom.** Inheritance of physical fitness in 10-yr-old twins and their parents. *Med. Sci. Sports Exerc.* 28: 1479–1491, 1996.
22. **Montgomery, H. E., R. Marshall, H. Hemingway, S. Myerson, P. Clarkson, C. Dollery, M. Hayward, D. E. Holliman,**

M. Jubb, M. World, E. L. Thomas, A. E. Brynes, N. Saeed, M. Barnard, J. D. Bell, K. Prasad, M. Rayson, P. J. Talmud, and S. E. Humphries. Human gene for physical performance. *Nature* 393: 221–222, 1998.

23. **Montoye, H. J., and R. Gayle.** Familial relationships in maximal oxygen uptake. *Hum. Biol.* 50: 241–249, 1978.
24. **Province, M. A., and D. C. Rao.** General purpose model and a computer program for combined segregation and path analysis (SEGPATH): automatically creating computer programs from symbolic language model specifications. *Genet. Epidemiol.* 12: 203–219, 1995.
25. **Province, M. A., T. Rice, I. B. Borecki, C. Gu, and D. C. Rao.** A multivariate and multilocus variance components approach using structural relationships to assess quantitative trait linkage via SEGPATH. *Genet. Epidemiol.* In press.
26. **Prud'homme, D., C. Bouchard, C. Leblanc, F. Landry, and E. Fontaine.** Sensitivity of maximal aerobic power to training is genotype-dependent. *Med. Sci. Sports. Exerc.* 16: 489–493, 1984.
27. **Rivera, M. A., F. T. Dionne, J. A. Simoneau, L. Pérusse, M. Chagnon, Y. C. Chagnon, J. Gagnon, A. S. Leon, D. C. Rao, J. S. Skinner, J. H. Wilmore, and C. Bouchard.** Muscle-specific creatine kinase gene polymorphism and $\dot{V}O_{2max}$ in the HERITAGE Family Study. *Med. Sci. Sports Exerc.* 29: 1311–1317, 1997.
28. **Rivera, M. A., L. Pérusse, J. A. Simoneau, J. Gagnon, F. T. Dionne, A. S. Leon, J. S. Skinner, J. H. Wilmore, M. Province, D. C. Rao, and C. Bouchard.** Linkage between a muscle-specific CK gene marker and $\dot{V}O_{2max}$ in the HERITAGE Family Study. *Med. Sci. Sports Exerc.* 31: 698–701, 1999.
29. **Simoneau, J. A., G. Lortie, M. R. Boulay, M. Marcotte, M. C. Thibault, and C. Bouchard.** Inheritance of human skeletal muscle and anaerobic capacity adaptation to high intensity intermittent training. *Int. J. Sports Med.* 7: 167–171, 1986.
30. **Skinner, J. S., K. M. Wilmore, A. Jaskólska, A. Jaskólski, E. W. Daw, T. Rice, J. Gagnon, A. S. Leon, J. H. Wilmore, D. C. Rao, and C. Bouchard.** Reproducibility of maximal exercise test data in the HERITAGE Family Study. *Med. Sci. Sports. Exerc.* 31: 1623–1628, 1999.
31. **Sundet, J. M., P. Magnus, and K. Tambs.** The heritability of maximal aerobic power: a study of Norwegian twins. *Scand. J. Med. Sci. Sports.* 4: 181–185, 1994.
32. **Van Deursen, J., A. Heerschap, F. Oerlemans, W. Ruitenbeek, P. Jap, H. ter Laak, and B. Wieringa.** Skeletal muscles of mice deficient in muscle creatine kinase lack burst activity. *Cell* 74: 621–631, 1993.
33. **Wilmore, J. H., P. R. Stanforth, M. A. Domenick, J. Gagnon, E. W. Daw, A. S. Leon, D. C. Rao, J. S. Skinner, and C. Bouchard.** Reproducibility of anthropometric and body composition measurements: the HERITAGE Family Study. *Int. J. Obesity* 21: 297–303, 1997.
34. **Yamashita, K., and T. Yoshioka.** Profiles of creatine kinase isoenzyme compositions in single muscle fibres of different types. *J. Muscle Res. Cell Motil.* 12: 37–44, 1991.