the heritage family study

Aims, design, and measurement protocol

CLAUDE BOUCHARD, ARTHUR S. LEON, D. C. RAO, JAMES S. SKINNER, JACK H. WILMORE, and JACQUES GAGNON

Physical Activity Sciences Laboratory, Laval University, Québec, G1K 7P4, CANADA
School of Kinesiology and Leisure Studies, University of Minnesota, Minneapolis, MN 55455
Division of Biostatistics and Departments of Genetics and Psychiatry, Washington University Medical School, St. Louis, MO 63110
Exercise and Sport Research Institute, Arizona State University, Tempe, AZ 85287
Department of Kinesiology and Health Education, The University of Texas at Austin, Texas, TX 78712

ABSTRACT

BOUCHARD, C., A. S. LEON, D. C. RAO, J. S. SKINNER, J. H. WILMORE, and J. GAGNON. Aims, design, and measurement protocol, *Med. Sci. Sports Exerc.*, Vol 27, No. 5, pp. 721–729, 1995. The HERITAGE family study (*HE*alth, *RI*sk factors, exercise *T*raining *A*nd *GE*netics) will document the role of the genotype in the cardiovascular, metabolic, and hormonal responses to aerobic exercise training. A consortium of five universities in the United States and Canada are involved in carrying out the study. A total of 90 Caucasian families and 40 African-American families with both parents and three or more biological adult offspring are being recruited, tested, exercise-trained in the laboratory with the same program for 20 wk, and re-tested. Oxygen uptake, respiratory exchange ratio, blood pressure, heart rate, cardiac output, blood lactate, glucose, and free-fatty acids are measured during exercise, and maximal oxygen uptake is determined before and after training. Plasma lipids, lipoproteins and apoproteins, glucose and insulin response to an intravenous glucose load, plasma sex steroids and glucocorticoids, and body fat and fat distribution are assessed. Dietary and activity habits and other life style components are assessed by questionnaires, prior to, during, and after training. A variety of genetic analyses will be undertaken, including heritability studies and major gene effects, for each phenotype and its response to regular exercise. Cell lines are established, and DNA sequence variation at a variety of molecular markers will be determined for association and linkage studies.

EXERCISE, TRAINING, INDIVIDUAL DIFFERENCES, GENETICS

0195-9131/95/2705-0721$3.00/0
MEDICINE AND SCIENCE IN SPORTS AND EXERCISE®
Copyright © 1995 by the American College of Sports Medicine

Submitted for publication March 1994.
Accepted for publication October 1994.

It is commonly recognized that a physically active life style is associated with a decreased risk for a variety of morbid conditions. The conclusion of a recent International Consensus Conference on the relationship between physical activity, fitness, and health was that there is ample evidence from experimental studies, as well as cross-sectional and prospective data from large-scale epidemiological studies, that regular exercise improves health-related fitness and has important positive effects on general wellness, morbidity, and mortality (9). However, the consensus document specified that additional studies were needed to improve our understanding of the contribution of inherited factors in determining changes in risk profiles in response to life style modifications, such as those seen when individuals engage in regular exercise. From the population viewpoint, regular exercise can potentially reduce the risk associated with the incidence of coronary artery disease, hypertension, type II diabetes mellitus, osteoporosis, obesity, and other degenerative conditions. Alterations in risk profile brought about by regular exercise are thought to be mediated by favorable effects on such factors as plasma lipoproteins, insulin sensitivity, body fat content and distribution, particularly upper body fat and abdominal visceral fat, blood pressure, steroid and glucocorticoid hormonal profile, cardiac stroke volume, and oxygen transport capacity.

Since cardiovascular disease and non-insulin-dependent diabetes account for over half of the deaths and a

large fraction of the health care costs in the United States, there is considerable interest in understanding the role of regular exercise on risk factors related to these diseases, as well as the individual variation in the response to regular exercise. A large body of research clearly indicates that maximal oxygen uptake ($\dot{V}O_{2max}$), cardiac output, and the metabolic pathways related to tissue substrate availability are all biological characteristics that can favorably adapt to exercise training (23,36,42). For instance, $\dot{V}O_{2max}$ typically increases by about 20–30% after several months of training (31,36). The activity of key enzymes of skeletal muscle oxidative potential often increases by 50% and, at times, doubles pre-training values (24,25,43,46).

While few experiments have been designed to document the importance of variation in the response to regular exercise, those which looked for these individual differences have found them (6–8). For instance, in one experiment, Lortie et al. (31) trained 24 sedentary young adults for 20 wk. The cycle ergometer training program was the same for each subject and fully monitored. While the average $\dot{V}O_{2max}$ rose 0.6 $l \cdot min^{-1}$ or 26%, gains ranged from 7 to 87%. Large individual differences in sensitivity to regular exercise have also been reported for endurance performance (18,31), cardiorespiratory adaptation to exercise (39), and changes in skeletal muscle enzyme activities (18,45,46), plasma lipoproteins, and apoproteins (14,15), adipose tissue metabolism (13,44), and insulin response to a glucose load (48). Individual differences in response to training for these biological characteristics typically range from zero to about 100% of the pre-exercise values.

What factors contribute to such variation in response to regular exercise? Briefly, age and gender of subjects, as well as prior exercise experience, did not seem to contribute significantly to individual differences in trainability in 18–30-yr olds. The major causes of variation in endurance training response were pre-training level of the phenotype and undetermined genetic characteristics in several separate experiments conducted with sets of identical twins (8). These experiments suggest that there are high, low, and non-responder genotypes in the response of $\dot{V}O_{2max}$ to regular exercise. Although this phenomenon has been documented primarily in young adults of both sexes, it is probably also present in middle-aged and older individuals, as well as for most of the cardiovascular disease and diabetes risk factors. Moreover, the problem has been investigated only with identical twins (8). The need for a more detailed study of the phenomenon in parents and their offspring was thus apparent.

AIMS OF THE HERITAGE STUDY

The overall objective of the HERITAGE project is *to study the role of the genotype in cardiovascular, metabolic, and hormonal responses to aerobic exercise training and the contribution of regular exercise to changes in several cardiovascular disease and diabetes risk factors.* It addresses the following major aims.

1. To extend our understanding of the genetic epidemiology of the following selected risk factors, with emphasis on changes observed with regular exercise: blood lipids, lipoproteins, apoproteins, and lipolytic enzymes; glucose tolerance and insulin sensitivity; systolic and diastolic blood pressure at rest and during exercise; body weight, total body fat, and regional fat distribution, including quantification of abdominal visceral fat; and steroid and glucocorticoid hormone levels. The issues to be resolved will be related to the heritability of each phenotype and its response to exercise training, the contribution of a specific paternal or maternal effect, sex-limited effects, major gene effects, and related segregation patterns.
2. To extend our understanding of the genetic epidemiology of aerobic exercise tolerance, its functional determinants, and their response to training. Major dependent variables include $\dot{V}O_{2max}$, cardiac output, stroke volume, and adaptation to submaximal exercise and training. This will allow the determination of potential mechanisms responsible for the large variation in physiological response to a standardized endurance exercise training program.
3. To investigate a number of physiological and epidemiological issues independent of potential familial aggregation such as: the role of regular exercise on various risk factors; differences between genders, between young and middle-aged adults, and between Caucasian and African-American subjects; and effects of other life style components (particularly smoking, nutrient intake, alcohol intake, sleep habits) on various phenotypes.
4. Our ultimate objective is to perform association and linkage studies between risk factor and performance phenotypes, especially as they relate to responses to exercise training, with a panel of candidate genes and other genetic markers.

This research should contribute significant new information on the genetic basis of adaptation to exercise training and concomitant changes in cardiovascular disease and diabetes risk factors. It also has important implications for understanding human variations in blood lipids and lipoproteins, glucose tolerance and insulin action, blood pressure, body fatness and fat distribution, and their associations with plasma sex and glucocorticoid hormone profiles in response to training.

DESIGN AND SAMPLING

The study design calls for the recruitment of 90 two-generational nuclear families of Caucasian descent and

40 families of African-American ancestry, each with both biological parents and at least three biological children. However, due to the socio-demographic characteristics of African-American families, they often tend to be uniparental (mostly the mother and her children from multiple marriages), and tend not to have many children in the required age range who are full sibs. Therefore, some African-American families will involve fewer than five subjects. We expect that about half of the African-American families will each have five or more members. The sample is expected to comprise a total of about 650 individuals. Families will be selected such that all participating members are within 17–65 yr of age (age of children: 17–40; age of parents: 65 and less), "sedentary" for 3 months prior to beginning the study, and essentially in good health. The participating individuals will be studied before and again after a 20-wk standardized training program, providing the opportunity to assess the familiality of response to regular exercise.

A control group of families is not included in the design for several reasons. First, there are several existing lines of evidence indicating that the implementation of a regular exercise program in previously sedentary individuals produces appreciable changes in fitness and risk factor variables on a group basis. The goal is not to replicate these findings, but to characterize the interindividual variability and address the genetic etiology of such responses. We seek to explain why some individuals respond well, while others show little or no response, and determine whether the degree of response is familial and genetic. Second, the mean magnitude of change in the measured variables as a consequence of the exercise training protocol is expected to be large, as compared to any changes occurring naturally over the same 20-wk period. Third, under the assumption that naturally occurring changes over 20 wk are negligible relative to the responses to training, the pre-training measurements provide approximate "control" levels. Finally, the enormous cost of obtaining data on control families is not justified vis-à-vis the stated goal, as it provides no information on the response of certain genotypes to an active life style. Thus, there is a certain economy in the study design without controls, in that individuals act as their own controls and information is maximized with respect to the specific aims of the research.

Recruitment of eligible families is based on extensive publicity and advertisement efforts. Combinations of campus advertisements, newspaper ads, radio and television ads, and community and church contacts are used. Respondents are initially screened on the phone, followed by a more extensive screening done at the clinic. Families eligible as of then are evaluated and measured during multiple visits to the clinic to determine final eligibility. The study has been approved by the competent Internal Review Board of each participating institution and written informed consent is obtained from each subject.

EXCLUSION CRITERIA

The exclusionary criterion based on age (outside 17–65 yr) was chosen so as to give the greatest range possible, while avoiding the complications of maturation at the low end and aging at the high end. The exclusion on the basis of health is based upon the ethical contraindications regarding maximal exercise testing in previously sedentary subjects.

Activity level. To assess the effects of exercise training accurately, all participants are required to be "sedentary" at baseline. Sedentary is defined as no regular physical activity over the previous 3 months. More specifically, these include activities lasting 30 min or more, involving an energy expenditure of 7 METS or more (based on American Heart Association criteria) for subjects 50 yr of age and over, and 8 METS for those below 50 yr, and occurring more than once a week. Families wishing to participate with some nonsedentary members can be reconsidered, providing that the nonsedentary individual(s) "detrain" for at least 3 months. The 3-month requirement will ensure that the individuals are essentially untrained with respect to cardiovascular fitness at the onset of the study; this is similar to the common "drug-free" period required for many clinical trials involving pharmacological agents.

Body weight standards. Body mass index or BMI (weight in kilograms divided by height in meters squared) is less than 40 kg/m^2 because of metabolic abnormalities and difficulty in exercising that are commonly associated with extreme obesity. Exceptions are accepted only with sufficient clinical justification on a case by case basis.

Blood pressure standards. The upper limits of acceptable blood pressure (BP) levels are a systolic BP of 159 mm Hg and a diastolic BP of 99 mm Hg for two out of three readings in the sitting position after at least 5 min of rest. Individuals on diuretic or antihypertensive drugs at the initial interview are permitted to enter the study if they are free of hypertensive complications, their personal physician permits them to discontinue their medication(s), and their blood pressure level meets the above criteria after at least 3 months off medication. The rationale for admitting people with mild hypertension (i.e., systolic BP 140–159 mm Hg and/or diastolic BP 90–99 mm Hg) without complications is that exercise training is reported to reduce elevated levels of both systolic and diastolic BP by about 10 mm Hg on the average (16). Subjects unable to be removed from their antihypertensive treatment are excluded from the study.

Absence of significant medical conditions and diseases. A detailed medical history and physical examination are conducted by a physician, or a nurse prac-

titioner under the supervision of a physician after screening. Subjects with suspicious symptoms or suggestive medical histories must have their personal physicians provide additional medical information, test results, and hospital records prior to inclusion or exclusion. A past history and/or physical or laboratory finding of medical conditions defined below require exclusion from the study.

Definite or possible coronary heart disease, including a positive exercise test or other clinically significant cardiovascular conditions. Other exclusion criteria include: significant chronic or recurrent respiratory conditions; a chronic or recurrent gastrointestinal problem; clinically significant urinary tract or genital diseases, disorders, or conditions; significant neuromuscular, neurological, or psychiatric conditions; major musculoskeletal problems interfering with ability to walk or cycle; major autoimmune or collagen vascular diseases; malignancies in the past 5 yr with the exception of skin cancer therapeutically controlled; uncontrolled endocrine and metabolic disorders or diabetes requiring insulin or oral hypoglycemic therapy, uncontrolled thyroid, parathyroid, adrenal, or hypothalamopituitary disorders, hypercholesterolemia with history of blood cholesterol level $\geq$350 mg·dl^{-1} or hypertriglyceridemia with a fasting blood triglyceride level $\geq$500 mg·dl^{-1} or the use of lipid-lowering drugs, hematologic disorders, and any other condition or disease that is life-threatening or that can interfere with or be aggravated by cycle exercise. Moreover, participants cannot have been a blood donor for the previous 6 wk or cannot be a blood donor during the course of the study. The Manual of Procedures of the HERITAGE study provides a detailed listing and definition of all exclusion criteria.

MEASUREMENT PROTOCOL

The study was approved by each participating institution's review board for including human subjects in research. Written informed consent is obtained from each study participant. Each subject receives $1,000 in incremental payments for successful completion of the study, i.e., after completing the 20-week training program plus the pre- and post-training battery of tests.

Health screening includes a health history, physical examination, a resting electrocardiogram (ECG) and an exercise test with ECG monitoring. Subjects also complete at baseline a health habit questionnaire, including smoking and alcohol consumption habits; the ARIC-Baecke Physical Activity Questionnaire (1,27); the Willett Food Frequency Questionnaire to assess usual food nutrient patterns (50); the Minnesota Eating Pattern Assessment Tool (EPAT) to evaluate dietary fat sources (26,35); a menstrual history; and a detailed family history questionnaire. The health habit and EPAT questionnaires are repeated at the midpoint of training (week 10), at which time participants are counseled again not to change baseline health habits. They are administered again at the end of training.

Anthropometric measurements before and after training include standing height, body weight, and a series of eight skinfold measurements using Harpenden calipers (20). Circumferences of the upper arm, waist, and hip (buttocks) are also obtained (11). Underwater weighing is performed in the post-absorptive state before and after training to determine body density, fat mass, fat-free mass and relative body fat (2). A correction is made for residual lung volume by the oxygen dilution principle (51,52). Abdominal visceral fat and abdominal subcutaneous fat is quantified before and after training by the method of Sjöström et al. (47). This technique involves computed axial tomography (CT scan) at the level of the disk between lumbar vertebrae 4 and 5 using the same protocol at all centers and the scans are sent to Laval University for review.

Resting blood pressure measurements are made twice prior to the start of exercise training and at 24 and 72 h post-training. Subjects are tested before 11:00 a.m. in the post-absorptive state with no caffeine-containing beverages and tobacco products for at least 2 h prior to measurements. Measurements are performed in a quiet room at neutral ambient temperature (24–25°C) with the lights dimmed and subjects rested for at least 5 min in a reclining chair with legs elevated and the chair's back support reclined at about 45° from the ground. Blood pressure is determined using a properly fitted cuff connected to a Colin STBP-780 automated unit. Ear phones are worn by the technicians during measurements to confirm blood pressure values.

Three exercise tests are administered prior to training and three additional tests at the conclusion of the training period. All tests are conducted on a stationary cycle ergometer in the sitting position. Each Clinical Center uses the same cycle ergometer (Ergo-Metrics 800S from SensorMedics). The first test at each time is to establish the participant's $\dot{V}O_{2max}$, as well as the normality of the exercise electrocardiogram in the pre-training test. During the second test of each battery, participants exercise at 50 W and at 60% of his/her $\dot{V}O_{2max}$ determined in the initial test for approximately 8 min at each power output to determine steady-state $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, RER, systolic and diastolic blood pressure, heart rate, cardiac output, and stroke volume. During the third and final test of each battery, the participant repeats the 50 W and power output at 60% of $\dot{V}O_{2max}$ of the second test, after which the power output is increased to 80% $\dot{V}O_{2max}$ for 3 min, and then continues to increase until the subject reaches exhaustion. The same variables are monitored during the second and third tests. In addition, a venous catheter is inserted in the left arm to obtain blood samples at rest, during exercise at 50 W, 60 and 80% $\dot{V}O_{2max}$, and immediately upon completion of the maximal test. Blood

samples are analyzed for glucose, free-fatty acids, lactate and total proteins. Metabolic measurements and cardiac output (CO_2 rebreathing technique) are determined using a SensorMedics 2900 metabolic cart, blood pressure with a Colin STBP-780 automated blood pressure monitor, and heart rate from the electrocardiogram. Stroke volume is derived from cardiac output and heart rate.

To measure plasma lipid and lipoprotein levels, blood samples are collected from an antecubital vein into vacutainer tubes containing EDTA. Samples are taken in the morning after a 12 h fast while the subjects are in a supine position. Cholesterol (C) and triglycerides (TG) levels are determined in plasma and in lipoprotein fractions by enzymatic methods using the Technicon RA-1000 analyzer. Plasma very low density lipoprotein (VLDL) (density < 1.006 g·ml^{-1}) are isolated by ultracentrifugation (21) and the high density lipoprotein (HDL) fraction obtained after precipitation of low density lipoprotein (LDL) in the infranatant (density > 1.006 g/ml) with heparin and $MnCl_2$ (10). The C and TG content of the infranatant fraction are measured before and after the precipitation step. Apoprotein (Apo) B concentration is measured in plasma and the infranatant (LDL-Apo B) by the rocket immunoelectrophoretic method of Laurell (29). Apo A-1 concentration is also measured in the infranatant fraction. The concentrations of LDL-C, LDL-TG, and VLDL-Apo B are obtained by difference. The cholesterol content of HDL_2 and HDL_3 subfractions prepared by precipitation method (17) is also determined. Apo E phenotype is measured by an isoelectrofocusing method. LPL and H-TGL activities are measured in plasma obtained from 12-h fasted subjects, 10 min after i.v. injection of heparin (60 IU·kg^{-1} body weight). Enzyme activities are measured by a modification of the method of Nilsson-Ehle and Ekman (33).

An intravenous glucose tolerance test is performed in the morning after an overnight fast. Blood samples are collected through a venous catheter from an antecubital vein for a total of 16 time points over 3 h to determine plasma glucose, insulin and connecting peptide (C-peptide) concentrations. Plasma glucose is enzymatically measured (41), whereas plasma insulin is measured by radioimmunoassay with polyethylene glycol separation (12). C-peptide is measured by a modification of the method of Heding (22) using polyethylene glycol precipitation (12). A colorimetric micromethod (34) is used to measure free fatty acids.

After extraction, measurement of serum steroid levels are performed by specific radioimmunoassays (3). The following steroids are assayed: androstenedione, testosterone, dihydrotestosterone, androsterone glucuronide, androstane-3α,17β-diol glucuronide, pregnenolone fatty acid esters, and dehydroepiandrosterone and its fatty acid. Progesterone, 17-hydroxyprogesterone, cortisol, aldosterone, estradiol, and dehydroepiandrosterone sulfate are determined using Diagnostic Product Corporation kits. Sex hormone binding globulin is also determined using a commercial kit.

Table 1. Overview of the 20-wk training program.

Weeks	Frequency (sessions/wk)	Intensity (% $\dot{V}O_{2max}$)	Duration (min/session)[a]
2	3	55	30
2	3	55	35
2	3	65	35
2	3	65	40
2	3	70	40
2	3	70	45
2	3	75	45
6	3	75	50

[a] Does not include 5-min warm-up or 3-min cool-down.

Permanent lymphoblastoid cell lines are established for each individual of the cohort to insure a continuous source of DNA. Given the wealth of performance, physiological, metabolic, and health data that will be gathered, such an approach is of the utmost importance. Lymphoblastoid cell lines are obtained by transformation of human lymphocytes with the Epstein-Barr virus (EBV). Such cell lines grow well in culture, have an infinite life span and present a chromosomal stability over the years (32). The procedure requires the isolation of monocyte cells, the transformation with the EBV and the cryopreservation of the transformed cell lines. Cell lines are also thawed occasionally to check cell viability.

TRAINING PROGRAM

Each family member is trained on a cycle ergometer, three times a week, for 20 wk using the same standardized training protocol in each of the four Clinical Centers. Training intensity is adjusted for individual differences in $\dot{V}O_{2max}$. The program is summarized in Table 1. The intensity and/or duration of the training program is adjusted each 2 wk, such that the subjects are working at the heart rate associated with 75% $\dot{V}O_{2max}$ for 50 min during the last 6 wk; this allows an adequate total energy turnover for a sufficient time period, increasing the chances to induce changes in many of the variables under investigation. The power output of the cycle ergometer is adjusted automatically to the heart rate response of the subject at all times during all training sessions. All training sessions are supervised on site.

Subjects must complete the required 60 training sessions within a maximum of 21 wk. They cannot exercise more than four sessions per week or less than one per week and cannot get ahead nor fall behind by more than two sessions. Adherence is monitored several times per week, and if a subject is falling behind, a plan is developed to bring him or her back on schedule as soon as possible.

The study personnel were centrally trained on all aspects of recruitment, measurement, and training using a specially prepared manual of procedures. Reproducibility studies are performed from time to time, and the reliabil-

ity coefficients, both within and among centers, are examined. Brief retraining efforts were undertaken early in the study when they were deemed necessary. These actions are designed to maximize similarity among the four centers with respect to recruitment, data collection and standardization of training protocol.

DATA ENTRY AND MANAGEMENT

A distributed data entry system has been developed for HERITAGE using SAS as a data management tool. The system was developed by the Data Coordinating Center in St. Louis. Each Clinical Center enters the data on its own commputer (PC) and sends it to the Data Coordinating Center on diskette, where it is collated into a master database. As errors are discovered, reports are generated back to the Clinical Centers, where the errors are resolved and the revised data re-sent to the Data Coordinating Center.

Each data collection center has a 486 PC licensed for SAS/PC and can run data entry programs written in SAS. These programs make particular use of the full screen product (FSP) and application facility (AF), and each data entry form is its own SAS dataset, with one observation per subject/visit. Heavy use of the Screen Control Language of PROC FSEDIT allows error verification and cross-checking at the time of data entry. Double data entry is employed as one of the quality assurance measures. At the Data Coordinating Center, Data Manager designs and maintains the data entry/checking programs, maintains the collated database, generates reports, and provides long-distance support for the local data entry personnel at each Clinical Center.

DATA ANALYSIS PLAN

The data analysis plan presented here provides an overview of the types of data analyses that are planned. Racial differences will be investigated by analyzing data on African-American families and on Caucasian families separately whenever possible and necessary. We remain alert to the possibility of modifying these plans as other pertinent information becomes available, such as from analysis of partial data.

The sample size for the Caucasian families was determined based on power studies for the types of data analyses discussed below. Additional boost in power can be achieved by pooling data over races to test certain hypotheses; thus we can potentially use all 130 families at one time. We expect that same "factors" are involved in both races that determine one's response to training; only the frequencies and/or effects may be different. For example, same genes may be involved with different gene frequencies in the two races. So long as such differences are accommodated within the data analyses, pooling over races raises no particular methodological problems. In fact, the sample size for African-Americans will permit such combined analyses without losing power on account of estimating some race-specific parameters.

Epidemiological models of response to exercise training. Because this study is an environmental intervention involving a host of measurements on 650 individuals, a natural approach is purely epidemiological, e.g., "Which factors measured at baseline best predict response to exercise?" Such an approach defers the deeper question of the underlying genetic and environmental causes of the responses and simply seeks an observational assessment. For this purpose, models will be developed for predicting the post-training levels using the pre-training levels as an additional covariate. This will allow higher order interaction terms to be entertained in the models, which might involve the pre-training levels of the response variable itself.

Path analysis. We propose to analyze the familial aggregation of many variables using the methods of path analysis, which enables a resolution of genetic, familial environmental, and random environmental effects (40). This is accomplished by defining the covariance among relatives, and in the case of multivariate analysis, the covariance among phenotypes both within individuals and among relatives, in terms of parameters of a model of genetic and environmental effects. For many of the pre-training variables in the performance domain, the familial aggregation of which is not well known, univariate analyses of one variable at a time will provide that information. Although these analyses will be interesting contributions in their own right, they will be equally useful in guiding additional analyses using complex multivariate models. Using multivariate models (19,49), it is possible to test whether the same genetic influences are acting on multiple measures. For example, the question of whether genetic influences on a cardiovascular risk variable in sedentary individuals (pre-training) are identical to the influences on the variable in trained individuals can be modeled and explicitly tested. Similarly, it is possible to assess in detail relationships between the mechanism underlying the change in a variable in response to exercise training and its baseline level. Finally, some of these models will also enable investigations of temporal trends in family resemblance (37).

Segregation analysis. Segregation analysis will be carried out to investigate possible major gene effects on a variety of phenotypes such as, for example, change in a certain risk factor (4,5,28). Underlying genetic models assume that the phenotype results from the joint and independent effects of an autosomal locus and residual multifactorial effects. Segregation at the major locus is modelled using general transmission probabilities that are the probabilities of the genotypes AA, Aa, and aa transmitting the A allele; Mendelian transmission holds when these probabilities are 1, ½, and 0, respectively. Test of the major gene hypothesis involves contrasting

the model in which the major gene parameters are set to zero against the general model in which those parameters are estimated. For those variables exhibiting evidence of a major gene effect, tests on transmission probabilities will be performed to protect against an incorrect inference in the presence of environmentally induced major effects. More recent models also permit combined path and segregation analysis (30,38).

For variables that exhibit significant changes in response to exercise training, complex segregation analysis of the change between the pre- and post-training values will be carried out. The primary purpose of these analyses is to investigate the nature of familial effects, if any, on the observed changes, with special reference to additional major gene effects. This is expected to suggest whether specific genotypes are associated with the ability to respond to exercise training. A dichotomy of "responders" and "nonresponders," defined in terms of whether there is a marked change between the pre- and post-training values, may be useful. For some of the more informative variables, analyzing the segregation of the "responder" status in families may provide important information.

Association and linkage studies. Associations between dichotomized response variables and a host of genetic markers will be tested. This approach serves two purposes. There is a possibility of identifying alleles associated with a higher/lower mean response in polygenically (multifactorially) determined variables. In addition, a preliminary notion of potential linkage relationships can be ascertained for those variables influenced by major genes.

If evidence of a single gene affecting the change in any variable after exercise training is found, then an attempt to implicate a particular gene will be made via linkage analysis with likely candidate-gene markers. Candidate genes can be identified on the basis of biological considerations, or in light of the associations revealed in the preliminary screening. Linkage analyses will be carried out using standard computer programs such as LINKAGE and LIPED. Lod scores, which denote measures of the relative odds that two loci are linked at particular recombination fractions vs unlinked, will be calculated.

Heterogeneity among the four centers. The proposed sample size is deemed necessary to yield enough power for hypothesis testing and for obtaining reasonably accurate parameter estimates. The complexity of this study makes it impossible to achieve the sample size in any single center within a reasonable period of time. That is why four centers will generate the data over a period of about 5 yr. We expect all four centers to contribute equal numbers of families; however, the racial composition at each center will vary. The multiple criteria for inclusion in the study should make the four subsamples of Caucasians and the subsamples of African-Americans rather homogeneous. However, we are well aware that heterogeneity among the four centers can be a potential problem. Although the sample size is insufficient to analyze each center's data separately, some steps will be taken to address this issue. Our plan calls for a detailed analysis only of the combined data set within each race. Once a certain parsimonious model is developed for the combined data set, we will find out if, within the context of such a model, specific components are different in the four subsamples. For example, when using the bivariate path model, we will investigate if the genetic effect on the response variable is different in the four groups. This will be done by estimating only the genetic effect separately in each group, and by fixing all other components of the model at the overall solution. In the case of segregation analysis, we can find out if the displacement between the homozygous means is different in the four groups. Similar investigations will be performed on other components of the models that seem interesting and important from a population heterogeneity point of view.

PROGRESS TO DATE

After several months of planning, protocol development, and standardization of procedures, data collection was initiated late in February 1993. At the time of writing this manuscript, more than 45 families have completed the full protocol and about 30 more were enrolled in the study. We believe that HERITAGE will be able to achieve its goal of completing 130 families, thus creating a unique resource for the investigation of the role of biological individuality in the adaptation to regular exercise.

The HERITAGE study is supported by the National Heart, Lung and Blood Institute through the following grants: HL45670 (C. Bouchard, PI), HL47323 (A. S. Leon, PI), HL47317 (D. C. Rao, PI), HL47327 (J. S. Skinner, PI), and HL47321 (J. H. Wilmore, PI).

Thanks are expressed to all the co-principal investigators, investigators, co-investigators, local project coordinators, research assistants, laboratory technicians, and secretaries who are contributing to the study. Finally, the entire HERITAGE consortium is very thankful to those hard-working participating families whose involvement alone demonstrates the feasibility of this study.

The HERITAGE study research consortium consists of the following.

Consortium Steering Committee: Claude Bouchard, Ph.D., Laval University, Chairperson; Arthur S. Leon, M.D., University of Minnesota; D. C. Rao, Ph.D., Washington University; James S. Skinner, Ph.D., Arizona State University; Jack H. Wilmore, Ph.D., The University of Texas at Austin; Jacques Gagnon, Ph.D., Laval University, Project Director; Jean Paul Albert, M.B.A., Laval University, Administrator.

Advisory Board: Elizabeth Barrett-Connor, M.D., University of California, San Diego; Jean Davignon, M.D., Clinical Research Institute of Montreal; E. Randy Eichner, B.A., M.D., University of Oklahoma; Robert C. Elston, M.A., Ph.D., Louisiana State University; William L. Haskell, Ph.D., Stanford University.

Laval University Consortium Coordinating Center and Clinical Center: Come S. Bouchard, M.Sc.; Marcelle Lareau, M.Sc., Local Project Coodinator; Claude Leblanc, M.Sc.; Gilles Lortie, M.D., Ph.D.; Isabel Mercier, M.Sc.; Louis Pérusse, Ph.D.; Denis Prud'homme, M.D., M.Sc.; Germain Thériault, M.D.; Angelo Tremblay, Ph.D.

University of Minnesota Clinical Center: Fernando S. Branco, M.D.; Erik J. Ekstrom, R.D.; William V. Mendoza, M.D.; Robert C. Serfass, Ph.D.; Ava J. Walker, Ph.D., Local Project Coordinator.

Washington University Data Coordinating Center: Ingrid B. Borecki, Ph.D.; Harry Cheng, M.Sc.; John O. Holloszy, M.D.; Zhaohai Li, Ph.D.; Laura E. Mitchell, Ph.D.; Michael A. Province, Ph.D.; Treva Rice, Ph.D.; Kenneth B. Schechtman, Ph.D.; George P. Vogler, Ph.D.

Arizona State University Clinical Center: Kelly Enders, R.N.; Anna A. Jaskolska, Ph.D.; Artur J. Jaskolski, Ph.D.; Kristine M. Wilmore, M.A., Local Project Coordinator.

The University of Texas at Austin Clinical Center: David W. Hayes, M.D.; Paul J. Roach, M.D.; Philip R. Stanforth, M.Sc., Local Project Coordinator.

Laval University Core Laboratories. Diabetes Research Unit: Yolande Montreuil, R.N.; André Nadeau, M.D., Ph.D., Director; Gilles Tancrède, M.Sc. *DNA and Cell Line Unit:* Claude Bouchard, Ph.D.; France T. Dionne, Ph.D., Director; Marie-Christine Thibault, Ph.D. *Lipid Research Center:* Jean-Pierre Després, Ph.D.; Paul Lupien, M.D., Ph.D., Director; Sital Moorjani, Ph.D.; *Steroid Laboratory:* Alain Bélanger, Ph.D., Director; Simon Caron, M.Sc.; Fernand Labrie, M.D., Ph.D.; Line Lavoie, M.Sc.

Address for correspondence: Claude Bouchard, Ph.D., Physical Activity Sciences Laboratory, PEPS, Laval University, Ste-Foy, Québec, G1K 7P4, CANADA.

REFERENCES

1. Baecke, J. A. H., J. Burema, and J. E. R. Frijters. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am. J. Nutr.* 36:936–942, 1982.
2. Behnke, A. R. and J. H. Wilmore. *Evaluation and Regulation of Body Build and Composition.* Englewood Cliffs, NJ, Prentice-Hall, 1974, pp. 21–27.
3. Belanger, A., S. Caron, and D. Picard. Simultaneous radioimmunoassay of progestins, androgens and estrogens in rat testis. *J. Steroid Biochem.* 13:185–190, 1980.
4. Bonney, G. E. On the statistical determination of major gene mechanisms in continuous human traits: regressive models. *Am. J. Med. Genet.* 18:731–749, 1984.
5. Borecki, I. B., G. M. Lathrop, G. R. Bonney, J. Yaouanq, and D. C. Rao. Combined segregation and linkage analysis of genetic hemochromatosis using affection status, serum iron, and HLA. *Am. J. Hum. Gen.* 47:542–550, 1990.
6. Bouchard, C. Genetics of aerobic power and capacity. In: *Sport and Human Genetics*, R. M. Malina and C. Bouchard (Eds.). Champaign, IL: Human Kinetics Publishers, 1986, pp. 59–88.
7. Bouchard, C., M. R. Boulay, J. A. Simoneau, G. Lortie, and L. Perusse. Heredity and trainability of aerobic and anaerobic performances. An update. *Sports Med.* 5:69–73, 1988.
8. Bouchard, C., F. T. Dionne, J. A. Simoneau, and M. R. Boulay. Genetics of aerobic and anaerobic performances. *Exerc. Sport Sci. Rev.* 10:27–58, 1992.
9. Bouchard, C., R. J. Shephard, and T. Stephens (Eds.). *Physical Activity, Fitness, and Health: Consensus Statement.* Champaign, IL: Human Kinetics Publishers, 1993, pp. 102.
10. Burstein, M. and J. Samaille. Sur un dosage rapide du cholestérol lié aux β-lipoprotéines du sérum. *J. Clin. Chim. Acta* 5:609–610, 1960.
11. Callaway, C. W., W. C. Chumlea, C. Bouchard, et al. Circumferences. In: *Anthropometric Standardization Reference Manual*, T. G. Lohman and A. F. Roche (Eds.). Champaign, IL: Human Kinetics Publishers, 1988, pp. 39–54.
12. Desbuquois, B. and G. D. Aurbach. Use of polyethylene glycol to separate free and antibody-bound peptide hormones in radioimmunoassays. *J. Clin. Endocrinol. Metab.* 37:732–738, 1971.
13. Despres, J. P., C. Bouchard, R. Savard, D. Prud'homme, L. Bukowiecki, and G. Theriault. Adaptive changes to training in adipose tissue lipolysis are genotype dependent. *Int. J. Obes.* 8:87–95, 1984.
14. Despres, J. P., C. Bouchard, R. Savard, A. Tremblay, and C. Allard. Lack of relationship between changes in adiposity and plasma lipids following endurance training. *Arteriosclerosis* 54: 135–143, 1985.
15. Despres, J. P., S. Moorjani, A. Tremblay, et al. Heredity and changes in plasma lipids and lipoproteins after short-term exercise training in men. *Arteriosclerosis* 8:402–409, 1988.
16. Fagard, R. H. and C. M. Tipton. Physical activity, fitness, and hypertension. In: *Physical Activity, Fitness, and Health. International Proceedings and Consensus Statement*, C. Bouchard, R. J. Shephard, and T. Stephens (Eds.). Champain, IL: Human Kinetics Publishers, 1994, pp. 633–654.
17. Gidez, L. I., G. J. Miller, M. Burstein, S. Slage, and H. H. Eder. Separation and quantitation of subclasses of human plasma high density lipoproteins by a simple precipitation procedure. *J. Lipid Res.* 23:1206–1223, 1982.
18. 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.
19. Hanis, C. L., C. F. Sing, W. R. Clarke, and H. G. Schrott. Multivariate models for human genetic analysis: aggregation, and tracking of systolic blood pressure and weight. *Am. J. Hum. Genet.* 36:1196–1210, 1983.
20. Harrison, G. G., E. R. Buskirk, J. E. Lindsay Carter, et al. Skinfold thicknesses and measurement technique. In: *Anthropometric Standardization Reference Manual*, T. G. Lohman and A. F. Roche (Eds.). Champaign, IL: Human Kinetics Publishers, 1988, pp. 55–80.
21. Havel, R. J., H. Eder, and H. F. Bragdow. The distribution and chemical composition of ultracentrifugally separated lipoproteins in human serum. *J. Clin. Invest.* 34:1345–1353, 1955.
22. Heding, L. G. Radioimmunological determination of human C-peptide in serum. *Diabetologia* 11:541–548, 1975.
23. Holloszy, J. O. Biochemical adaptations in muscle. Effects of exercise on mitochondrial oxygen uptake and respiratory enzyme activity in skeletal muscle. *J. Biol. Chem.* 242:2278–2282, 1967.
24. Holloszy, J. O. and E. F. Coyle. Adaptation of skeletal muscle to endurance exercise and their metabolic consequences. *J. Appl. Physiol.* 56:831–838, 1984.
25. Howald, H. Training-induced morphological and functional changes in skeletal muscle. *Int. J. Sports Med.* 3:1–12, 1982.
26. Hunninghake, D. B., R. M. Mullis, M. L. Brekke, J. R. Peters, and E. S. Quiter. A simpletool for the assessment of dietary fat intake in clinical populations. *Circulation, Abstracts of the 60th Scientific Sessions* 76(Suppl. IV):35, 1987.
27. Jacobs, D. R., Jr., B. E. Ainsworth, T. J. Hartman, and A. S. Leon. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med. Sci. Sports Exerc.* 25:81–91, 1992.
28. Lalouel, J. M., D. C. Rao, N. E. Morton, and R. C. Elston. A unified model for complex segregation analysis. *Am. J. Hum. Genet.* 35:816–826, 1983.
29. Laurell, C. B. Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. *Anal. Biochem.* 15:45–52, 1966.
30. Li, Z., G. E. Bonney, G. M. Lathrop, and D. C. Rao. Genetic analysis combining path analysis with regressive models: the TAU model of multifactorial transmission. *Hum. Hered.* 44:305–311, 1994.
31. Lortie, G., J. A. Simoneau, P. Hamel, M. R. Boulay, F. Landry, and C. Bouchard. Responses of maximal aerobic power and capacity to aerobic training. *Int. J. Sports Med.* 5:232–236, 1984.
32. Nilsson, K. Establishment of permanent human lymphoblastoid cell lines *in vitro*. In: *In Vitro Methods in Cell-Mediated and Tumor Immunity*, B. R. Bloom and J. R. David (Eds.). New York: Academic Press, 1976, pp. 713–735.
33. Nilsson-Ehle, P. and R. Ekman. Specific assays for lipoprotein

lipase activities of post-heparin plasma. In: *Protides of Biological Fluids*, H. Peeters (Ed.). Oxford: Pergamon Press, 1978, pp. 243–246.
34. NOMA, A., H. OKABE, and M. KITA. A new colorimetric microdetermination of free fatty acids in serum. *Clin. Chim. Acta* 43:317–320, 1973.
35. PETERS, J. R., M. J. BREKKE, E. S. QUITER, R. M. MULLIS, M. L. BREKKE, and D. B. HUNNINGHAKE. Eating pattern assessment tool: a simple instrument for assessing dietary fat. *J. Am. Diet. Assoc.* 94:1008–1013, 1994.
36. POLLOCK, M. L. The quantification of endurance training programs. *Exerc. Sport Sci. Rev.* 1:155–188, 1973.
37. PROVINCE, M. A. and D. C. RAO. A new model for the resolution of cultural and biological inheritance in the presence of temporal trends: Application to systolic blood pressure. *Genet. Epidemiol.* 2:363–374, 1985.
38. PROVINCE, M. A. and D. C. RAO. A general purpose model and a computer program for combined segregation and path analysis (SEGPATH): automatically creating computer programs form symbolic language model specifications. *Genet. Epidemiol.* (in press).
39. 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.
40. RAO, D. C. Application of path analysis in human genetics. In: *Multivariate Analysis VI*, P. R. Krishnaish (Ed.). Amsterdam: North-Holland Publishing Co., 1985, pp. 467–484.
41. RICHTERICH, R. and H. DAUWALDER. Zur Bestimmung der Plasmaglukose-Konzentration mit der Hexokinase-Glucose-6-Phosphat-Dehydrogenase Methode. *Schweiz. Med. Wochenschr.* 101:615–618, 1971.
42. SALTIN, B., G. BLOMQVIST, J. H. MITCHELL, R. L. JOHNSON, K. WILDENTHAL, and C. B. CHAPMAN. Response to exercise after bed rest and after training. *Circulation* 38:1–68, 1968.
43. SALTIN, B. and P. GOLLNICK. Skeletal muscle adaptability: significance for metabolism and performance. In: *Skeletal Muscle. Handbook of Physiology*, L. D. Peachy, R. H. Adrian, and S. R. Geiger (Eds.). Bethesda, MD: American Physiological Society, 1983, pp. 555–631.
44. SAVARD, R., J. P. DESPRES, M. MARCOTTE, and C. BOUCHARD. Endurance training and glucose conversion into triglycerides in human fat cells. *J. Appl. Physiol.* 58:230–235, 1985.
45. 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.
46. SIMONEAU, J. A., G. LORTIE, M. R. BOULAY, M. MARCOTTE, M. C. THIBAULT, and C. BOUCHARD. Effects of two high-intensity intermittent training programs interspaced by detraining on human skeletal muscle and performance. *Eur. J. Appl. Physiol.* 56:516–521, 1987.
47. SJOSTROM, L., H. KVIST, A. CERDERBLAD, and U. TYLEN. Determination of the total adipose tissue volume in women by computed tomography: comparison with 40K and tritium techniques. *Am. J. Physiol.* 250:736–745, 1986.
48. TREMBLAY, A., E. T. POEHLMAN, A. NADEAU, L. PERUSSE, and C. BOUCHARD. Is the response of plasma glucose and insulin to short-term exercise-training genetically determined? *Horm. Metab. Res.* 19:65–67, 1987.
49. VOGLER, G. P., D. C. RAO, P. M. LASKARZEWSKI, C. J. GLUECK, and J. M. RUSSELL. Multivariate analysis of lipoprotein cholesterol fractions. *Am. J. Epidemiol.* 125:706–719, 1987.
50. WILLETT, W. C., L. SAMPSON, M. STAMPER, et al. Reproducibility and validity of a semi-quantitative food frequency questionnaire. *Am. J. Epidemiol.* 122:51–65, 1985.
51. WILMORE, J. H. A simplified method for determination of residual lung volume. *J. Appl. Physiol.* 27:96–100, 1969.
52. WILMORE, J. H., P. A. VODAK, R. B. PARR, R. M. GIRANDOLA, and J. E. BILLING. Further simplification of a method for determining residual lung volume. *Med. Sci. Sports* 12:216–218, 1980.