Spanish genetic admixture is associated with larger \( \dot{V}O_{2\text{max}} \) decrement from sea level to 4,338 m in Peruvian Quechua

Tom D. Brutsaert,1 Esteban J. Parra,2 Mark D. Shriver,3 Alfredo Gamboa,4 Jose-Antonio Palacios,4 María Rivera,4 Ivette Rodríguez,3,5 and Fabiola León-Velarde4

1Department of Anthropology, The University at Albany, State University of New York, Albany, New York 12222; 2Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada L5L 1C6; 3Department of Anthropology, The Pennsylvania State University, University Park, Pennsylvania 16804; 4Departamento de Ciencias Biológicas y Fisiológicas, Universidad Peruana Cayetano Heredia, San Martin de Porras, Peru; and 5Instituto Boliviano de Biología de Altura, Casilla 717, La Paz, Bolivia

Submitted 27 November 2002; accepted in final form 7 April 2003

Brutsaert, Tom D., Esteban J. Parra, Mark D. Shriver, Alfredo Gamboa, Jose-Antonio Palacios, María Rivera, Ivette Rodríguez, and Fabiola León-Velarde. Spanish genetic admixture is associated with larger \( \dot{V}O_{2\text{max}} \) decrement from sea level to 4,338 m in Peruvian Quechua. J Appl Physiol 95: 519–528, 2003. First published April 11, 2003; 10.1152/japplphysiol.01088.2002.—Quechua in the Andes may be genetically adapted to altitude and able to resist decrements in maximal \( O_2 \) consumption in hypoxia. This hypothesis was tested via repeated measures of \( \dot{V}O_{2\text{max}} \) (sea level vs. 4,338 m) in 30 men of mixed Spanish and Quechua origins. Individual genetic admixture level (%Spanish ancestry) was estimated by using ancestry-informative DNA markers. Genetic admixture explained a significant proportion of the variability in \( \Delta \dot{V}O_{2\text{max}} \) after control for covariate effects, including sea level \( \dot{V}O_{2\text{max}} \) and the decrement in arterial \( O_2 \) saturation measured at \( \dot{V}O_{2\text{max}} \) (\( \Delta \dot{SpO}_{2\text{max}} \) \( R^2 \) for admixture and covariate effects \( \sim \) 0.80). The genetic effect reflected a main effect of admixture on \( \Delta \dot{V}O_{2\text{max}} \) \( (P = 0.041) \) and an interaction between admixture and \( \Delta \dot{SpO}_{2\text{max}} \) \( (P = 0.018) \). Admixture predicted \( \Delta \dot{V}O_{2\text{max}} \) only in subjects with a large \( \Delta \dot{SpO}_{2\text{max}} \) \( (P = 0.031) \). In such subjects, \( \Delta \dot{V}O_{2\text{max}} \) was 12–18% larger in a subgroup of subjects with high vs. low Spanish ancestry, with least squares mean values (±SE) of 739 ± 71 vs. 606 ± 68 ml/min, respectively. A trend for interaction \( (P = 0.095) \) was also noted between admixture and the decrease in ventilatory threshold at 4,338 m. As previously, admixture predicted \( \Delta \dot{V}O_{2\text{max}} \) only in subjects with a large decrease in ventilatory threshold. These findings suggest that the genetic effect on \( \Delta \dot{V}O_{2\text{max}} \) depends on a subject’s aerobic fitness. Genetic effects may be more important (or easier to detect) in athletic subjects who are more likely to show gas-exchange impairment during exercise. The results of this study are consistent with the evolutionary hypothesis and point to a better gas-exchange system in Quechua.

deoxyribonucleic acid; genetic markers; aerobic performance; Andes; hypoxia; altitude

NATIVE SOUTH AMERICAN Quechua, and a few other native Andean ethnic groups, are thought to be descendant from groups who first reached the highland Andes ~10,000 years ago (31). Given the extremes of altitude that are now permanently inhabited in the region [up to 5,200 m (50)], it seems reasonable to hypothesize that these populations derive in part from groups who experienced natural selection in the past favoring superior oxygen transport phenotypes. If so, one functional consequence of genetic adaptation might be an ability (in current populations) to limit normal impairments in oxygen uptake and/or utilization that occurs during strenuous exercise in hypoxia. Indeed, many previous studies have made this argument, pointing to the relatively high maximal oxygen consumption \( \dot{V}O_2; \dot{V}O_{2\text{max}} \) in ml·min\(^{-1}\)·kg\(^{-1}\) in hypoxia of Andean study groups (3, 11, 14, 29, 34, 35), or suggesting that such groups, compared with lowland groups, experience only a small decrement in \( \dot{V}O_{2\text{max}} \) when exposed to hypoxia (2, 11, 14, 23, 45, 47, 49). Regarding the latter, only a few studies have directly measured the \( \Delta \dot{V}O_{2\text{max}} \) in such groups using a repeated-measures design (2, 23, 47), and these report \( \dot{V}O_{2\text{max}} \) decrements in Andean natives that are between ~30 and 80% of the decrement seen in lowland comparison groups. In this regard, the smaller \( \dot{V}O_{2\text{max}} \) decrement in Quechua may reflect an integrated functional response to hypobaric hypoxia that involves multiple physiological and biochemical systems (22). However, the reported differences in the magnitude of the \( \dot{V}O_{2\text{max}} \) decrement may be exaggerated. In two of the studies cited above, the lowland comparison groups were trained athletes, and aerobic fitness has a well-known positive effect on the \( \dot{V}O_{2\text{max}} \) decrement (12, 17, 28, 30, 32, 41, 42, 44).

In the present study, we assessed \( \dot{V}O_{2\text{max}} \) decrement via a repeated-measures design in a large group \( (n = 30) \) of young Peruvian men of mixed Quechua and Spanish ancestry who were born and raised in Lima, Peru (sea level). These subjects were first measured in...
Lima and then were transported to Cerro de Pasco, Peru (4,338 m) for measurement after ~12 h of exposure to hypobaric hypoxia. The main objective of our study was to assess the influence of Quechua vs. Spanish genetic admixture on the VO₂ max decrement within this study group, controlling for variation in aerobic fitness that may impact the magnitude of decrement experienced. Our focus on individual genetic admixture level as a study-independent variable represents a new research strategy to detect the effects of genetic adaptation on physiological phenotypes. The approach is made possible by new molecular genetic techniques that give admixture estimates for individuals descendant from two or more parental populations. Admixture in the Andes is historically well documented and began ~500 yr ago with the contact of Andean Native American, Spanish, and, to a lesser extent, West African groups. Thus the approach represents an alternative to the more typical study comparing an Andean control group with a lowland “control” group. Such studies ignore the reality of Andean population history and may be confounded by many factors, including unknown admixture levels in the so-called native group.

To estimate the extent of Spanish ancestry (admixture) for each individual within the sample, we used a panel of 22 ancestry-informative DNA genetic markers. This revealed an admixture range within the sample from <1 to 64% Spanish ancestry. From this, we hypothesized a positive relationship between the extent of Spanish ancestry and the magnitude of the VO₂ max decrement, as would be expected if natural selection had favored hypoxia-tolerant phenotypes in past Quechua populations. Importantly, for the approach to be useful, the individual genetic markers need not be associated (i.e., linked) with whatever genes determine exercise capacity in Quechua. The markers are simply used to produce a probability estimate of the proportionate ancestry of an individual. As a construct that reflects ancestry (and not specific genes or markers), the admixture estimate may be associated with a specific physiological phenotype, even if the individual genetic markers are not. In this regard, the approach may be seen as a first step to identify physiological phenotypes for further genetic study.

### MATERIALS AND METHODS

#### Subject Selection Criteria and Study Populations

The subjects for this study were young men (18–35 yr), who gave written, informed consent according to guidelines approved by the Institutional Review Boards at the University at Albany, State University of New York, and the Universidad Cayetano Heredia, Lima, Peru. Subjects were identified as nonsmokers and screened via a brief clinical history and medical examination for conditions contraindicating participation in the study protocols, including chronic obstructive respiratory diseases, cardiovascular disease, and renal disease. At screening, a venous blood sample was drawn from the antecubital vein, and Hb concentration (g/dl) was immediately determined by a Hemocue blood Hb analyzer (Abelholm, Sweden). Subjects with Hb <13.4 g/dl were considered anemic and were excluded from the study.

The subjects were recruited from within a specific district of Lima, Peru (Barrios Altos district, population ~150,000). In this district, ~10% of inhabitants are recent down-migrants from highland Peru. Individuals accepted into the study were born in Lima or near sea level, and both sets of their parents and grandparents were born at an altitude >3,000 m. Thus all study subjects were first- or second-generation down-migrants. No specific attempt was made to recruit subjects based on surname composition or skin reflectance measures, both of which have been used in the past to assess ancestry. The majority of subjects described themselves as “Peruvians,” but acknowledged both their Quechua and Spanish origins.

#### Study Design

Subjects were recruited into the study during the last 2 wk of July 2001 in Lima, Peru. Each subject completed baseline studies in Lima requiring ~4 h of participation. These subjects were studied again within 2 wk at 4,338 m in the town of Cerro de Pasco, Peru. Cerro de Pasco is a 6- to 10-h bus ride from Lima on paved road. The first ~4–6 h of the trip involve a steady gain in altitude to a high mountain pass (~4,800 m). The road then descends to the Peruvian Altiplano (3,600–4,300 m) for the next 3–4 h of the trip. Subjects arrived in Cerro de Pasco from Lima in groups of four to five per day over a 2-wk period and rested in the laboratory for 2–4 h before studies were initiated. Thus these subjects were studied after 10–12 h of acute exposure to hypobaric hypoxia. Of 32 subjects measured in Lima, two were not measured in Cerro de Pasco. One was unable to make the trip for personal reasons, and the other was diagnosed with acute mountain sickness on arrival to Cerro de Pasco.

#### Anthropometry and Pulmonary Function

Standard anthropometry was performed on each subject by the same investigator. Measurements included height, weight, and skinfolds at subescapular, suprailiac, biceps, and triceps sites. Body density was calculated according to age and sex-specific equations given by Durnin and Womersley (10). Kashiwazaki et al. (26) have tested the validity of a number of reference equations against doubly labeled water measures of body composition in Bolivian Aymara and concluded that the reference equations above are the best available for use in Andean native populations. Percent body fat was calculated from the Siri equation. Pulmonary function was assessed on each subject in Cerro de Pasco by using a VS400 volumetric spirometer (Puritan-Bennett, Mallinckrodt, Hazelwood, MO), calibrated daily with a 3-liter calibration syringe. Each subject performed a maximal inspiration, followed immediately by a forced maximal expiration while in a standing position. From this procedure, the forced vital capacity (FVC) and forced expiratory volume made in 1 s were determined based on the best of at least two efforts. FVC and forced expiratory volume in 1 s measures were corrected for BTPS.

#### Admixture Rate

**Genetic markers.** Individual Spanish admixture proportion was estimated by using a panel of 22 informative genetic markers (MID-575, TSC1102055, WI-11153, MID-52, SGC30610, WI-17163, WI-9231, WI-4019, WI-11909, D11S429, TYR-192, DRD2 TagD, DRD2 BcI, WI-14319, CYP19, PV92, WI-7423, CKM, MID-161, MID-93, FY, and...
F13B). The first 20 markers were selected because they show high-frequency differences (>30%) between Native American and Spanish populations, which are a priory the main parental populations in this sample. We also included in the panel two markers (FY and F13B) providing information on African ancestry. Details of these markers, including allele frequencies in all parental populations, DNA sequences, exact positions of single-nucleotide polymorphisms (SNPs), and the PCR primers and amplification conditions used are available from the dbSNP database (www.ncbi.nlm.nih.gov/SNP) under the submitter handle PSU-ANTH.

**Genotyping.** SNPs and insertion and deletion markers were scored by a melting-curve assay, in which the target sequence containing the SNP is amplified by PCR by using a mismatched primer where necessary to create an artificial restriction site polymorphism. PCR products were digested with a restriction enzyme, and the resulting restriction fragment-length polymorphisms are scored by their melting curves in a Hybaid DASH machine (ThermoHybaid). More details about this genotyping method can be found in Akery et al. (1). PCR products of the PV92 Alu insertion polymorphism were separated by conventional SSCP gel electrophoresis.

**Admixture estimation.** Usually, in samples from admixed populations, admixture is estimated by using the frequencies of samples representing the contributing parental populations as a contrasting reference. In this case, although there is information on the allele frequencies for these markers in several European populations, no such information is available for the Quechua parental population. Thus, to estimate admixture, we have used a strategy in which information on all parental frequencies is not required. We have used the program STRUCTURE, developed by Jonathan Pritchard, to infer admixture proportions in the samples from Lima and Cerro de Pasco. This program has been designed to infer the presence of genetic structure and to estimate the admixture proportions in samples with unknown genetic structure. This program can be used to estimate the number of subpopulations present in a sample and to assign individuals to each of those subpopulations, including estimates of individual admixture from each subpopulation. To estimate admixture in our sample, we prepared an input file, which included the genotype data for the 32 samples from Lima, 39 samples of highland-born subjects from Cerro de Pasco (not reported here), and 72 additional samples from Spain. We then ran the program STRUCTURE, with K = 2 as the predefined setting for the number of populations, using 30,000 iterations for the burn-in period and 70,000 additional iterations to obtain parameter estimates. The output file provides an estimate of the European and Native American ancestry for each individual in the sample. We ran the program several times, with consistent results. More information about the program STRUCTURE can be found in Pritchard et al. (38). Using STRUCTURE in this manner has been shown to result in individual admixture estimates that are highly correlated with estimates made by using maximum likelihood and parental allele frequencies from both parental populations.

The estimates of individual admixture range from 0 (Quechua) to 1 (Spanish) and reflect the proportionate contribution of two different population histories to the genetic makeup of an individual.

**Exercise**

Identical protocols to measure VO_{2 max} (l/min) were administered in Lima and Cerro de Pasco. To begin, VO_{2} was measured at rest (5 min) with the subject seated. After resting measurements, VO_{2 max} was measured on a mechanically braked Monarch 818e research ergometer. Subjects started with a workload of 1.0-kg resistance at 60 rpm, and resistance was incremented by 0.5 kg at constant rpm every 3 min until subject volitional fatigue. Subjects were given verbal encouragement, and VO_{2 max} was defined as the highest level of VO_{2} averaged over the final minute of the test, concomitant with at least one of the following: a nonlinear increase in exercise ventilation, resulting in a respiratory exchange ratio (RE) > 1.10, a plateau in the VO_{2}-work rate relationship, or a maximal heart rate (HR) within 10% of the age-predicted maximum.

During VO_{2} testing, subjects breathed through a low-resistance breathing valve, and expired ventilation (Ve, l/min BTPS), as well as the fractional concentrations of O_{2} and CO_{2} in expired air, was processed by a Parvo-medics True Max metabolic measuring system (Sandy, UT) to produce 1-min-interval calculations of VO_{2}, carbon dioxide production (VCO_{2}), the RER, and the ventilatory equivalents for oxygen and carbon dioxide (Ve/VO_{2} and Ve/VCO_{2}, respectively). Gas analyzers were calibrated with standard gases before each exercise test. The pneumotach used to measure ventilatory flow was calibrated and gel electrified before each test with a 3-liter calibration syringe. HR was continuously monitored via telemetry (Polar Electric Oy, Sweden) interfaced with the metabolic measuring system. Arterial oxygen saturation by pulse oximetry (Sp_{O_{2}}) was continuously monitored by an Ohmeda 5740 pulse oximeter by using a finger-tip sensor (subjects were instructed not to grip with that finger). The pulse oximeter signal was acquired by an REM/400M data-acquisition system (CB Sciences) and recorded every 15 s during VO_{2} measurements. Ventilatory threshold (Ve{\text{threshold}}) was determined graphically from the Ve, Ve/VO_{2}, and Ve/VCO_{2}, according to the method of Caiozzo et al. (7).

**Statistics**

All variables were evaluated for normality by using the Kolmogorov-Smirnov test against a standard normal distribution by using the Lilliefors two-tail probability. The admixture variable was transformed by the natural logarithm (log_{e}(admixture)) to achieve normality for statistical testing purposes. Differences between exercise response variables in Lima and Cerro de Pasco were evaluated by paired Student’s t-test, whereas differences between subgroups established within the data set were evaluated by t-test for independent samples. The general linear model procedure from SYSTAT version 5.1a (Macintosh) or version 9.0 (personal computer) was used to construct univariate and multivariate models predicting the \Delta VO_{2 max} (decrement) from sea level to 4,338 m. Such analyses are potentially problematic because change (\Delta) depends on the initial or baseline value. Baseline, in this regard, is the value achieved at sea level for a given measure. To address this issue, we have adopted the statistical approach recommended by a number of authors (21, 25). That is, when analyzing a \Delta-dependent variable, the baseline value must be entered as a covariate, despite the obvious mathematical relationship between dependent and independent variables. Similarly, a baseline value must be entered as a covariate (control) when analyzing a \Delta-independent variable, despite the obvious mathematical collinearity between the baseline and \Delta-independent variables. This procedure results in a nonbiased removal of baseline effects. In the present context, it allowed the evaluation of the effect of admixture on VO_{2 max} decrement, independent of the level of VO_{2 max} achieved at sea level.

Our analytic strategy to detect genetic variance in \Delta VO_{2 max} was as follows. First, multivariate models were

J Appl Physiol • VOL 95 • AUGUST 2003 • www.jap.org
constructed that controlled for baseline and other important covariate effects on $\Delta V_O_2\_m_a_x$. Then, genetic main effects were evaluated by the addition of the log($e$) admixture variable. Last, interaction effects between log($e$) admixture and relevant covariates were evaluated.

Values are expressed as means ± SD, unless otherwise indicated. Statistical significance criteria was $P \leq 0.05$ for all tests.

RESULTS

Subject Characteristics

Subject characteristics are given in Table 1. Compared with US National Health and Nutrition Examination Surveys reference standards, these subjects fell below the 5th percentile for stature by age, and near the 15th and 50th percentiles for weight by age and stature by weight, respectively. In this respect, they are typical of both highland- and lowland-born healthy Andean native men. Compared with a highland-born control population from Cerro de Pasco, where exercise was subjectively unpleasant, they were slightly taller and fatter and had 15% smaller FVC, as expected (comparative data not shown). The average admixture rate was ~10% Spanish ancestry and ranged from <1 to 64%.

Maximal exercise response data are given in Table 2. These subjects were well motivated to perform the exercise test, and all achieved the criteria for a true $V_O_2\_m_a_x$ during exercise at sea level. Even in Cerro de Pasco, where exercise was subjectively unpleasant, most subjects showed nonlinear increases in $V_e$ near the end of the $V_O_2\_m_a_x$ test, resulting in a high RER, and a majority (21 out of 30 subjects) had maximal HR levels within 10% of their age-predicted maximum.

The $V_O_2\_m_a_x$ decreased from sea level to 4,338 m by ~19% ($P < 0.001$). Maximal HR was significantly lower at 4,338 m compared with sea level. $V_e$ $B_T P_S$ was unchanged, but the $V_e$ $B_T P_S$/$V_O_2$ was significantly higher at 4,338 m. The RER was unchanged from sea level to 4,338 m. The mean $V_e$ $thres$, expressed in liters per minute of $V_O_2$, was significantly lower at 4,338 m compared with sea level, i.e., 2.27 vs. 1.64 l/min ($P < 0.01$). $V_e$ $thres$ was also significantly lower at altitude when expressed as a percentage of $V_O_2\_m_a_x$, i.e., 74.1 vs. 66.1% of $V_O_2\_m_a_x$ at sea level and 4,338 m, respectively.

Table 3 gives the matrix of correlation coefficients between variables that were used in subsequent multivariate analyses to model the effect of genetic admixture on $\Delta V_O_2\_m_a_x$. Variables that were not significantly related to $\Delta V_O_2\_m_a_x$, and thus not included in multivariate analyses, included the fat-free mass and Hb concentration. Sea level $V_O_2\_m_a_x$, sea level $V_e$ $thres$, $\Delta S_P_O_2$ measured at $V_O_2\_m_a_x$, $\Delta S_P_O_2\_E$ and $\Delta V_e$ $thres$ were all strongly correlated to $\Delta V_O_2\_m_a_x$, i.e., correlation coefficients between 0.59 and 0.81. Figures 1 and 2 show the relationship between $\Delta V_O_2\_m_a_x$ and sea level $V_O_2\_m_a_x$ and $\Delta S_P_O_2\_m_a_x$, respectively, as both relationships are central to the multivariate analyses that follow. Figure 1 reveals minimal $V_O_2\_m_a_x$ decrement for individuals with <2.5 l/min sea level $V_O_2\_m_a_x$, but substantially larger decrements (~25%) for individuals approaching 4.0 l/min sea level $V_O_2\_m_a_x$. Similarly, Fig. 2 reveals only minimal $V_O_2\_m_a_x$ decrements for individuals with <10% decrease in $S_P_O_2\_m_a_x$ and larger $V_O_2\_m_a_x$ decrements for individuals decreasing $S_P_O_2\_m_a_x$ by >20%.

Table 4 gives results of multivariate analyses modeling the decrease in $V_O_2\_m_a_x$ from sea level to 4,338 m. Model 1 (Table 4) is the simplest model and again demonstrates that the majority of the variability in $\Delta V_O_2\_m_a_x$ is explained by sea level $V_O_2\_m_a_x$ ($\beta = 0.516$, $R^2 = 0.669$, $P < 0.001$). Model building continued with the addition of the $\Delta S_P_O_2$ variable, controlling for the sea level value of $S_P_O_2\_m_a_x$ (model 2, Table 4). This model reveals a significant independent association of $\Delta S_P_O_2\_m_a_x$ on $\Delta V_O_2\_m_a_x$ ($\beta = 0.022$, $P = 0.033$), increasing the $R^2$ over model 1 by 0.058. The addition of log($e$) admixture (model 3, Table 4) reveals a significant main effect of log($e$) admixture ($\beta = -0.205, P = 0.041$) and a significant interaction of log($e$) admixture with $\Delta S_P_O_2$ ($\beta = 0.014, P = 0.018$).

The nature of this interaction is best understood by viewing plots that show the $\Delta V_O_2\_m_a_x$ as a function of admixture and $\Delta S_P_O_2$ in various study subgroups (Figs. 3 and 4). Subgroups were created by splitting the

### Table 1. Subject characteristics

| Age, yr | 25.3 ± 4.2 |
| Weight, kg | 76.3 ± 61.4 |
| Body fat, % | 21.8 ± 6.2 |
| Fat-free mass, kg | 49.2 ± 5.6 |
| Hemoglobin, g/dl | 14.9 ± 0.9 |
| FVC, liters BTPS | 4.86 ± 0.50 |
| FEV1, liters BTPS | 4.39 ± 0.60 |
| FEV1/FVC | 0.91 ± 0.09 |
| Admixture, % | 9.5 ± 14.2 |

Values are means ± SD; n = 30 subjects completing the study.

### Table 2. Maximal exercise response in Lima (sea level) and in Cerro de Pasco, Peru (4,338 m)

<table>
<thead>
<tr>
<th></th>
<th>Lima (Sea Level)</th>
<th>Cerro de Pasco (4,338 m)</th>
<th>Δ, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_O_2_m_a_x$ l/min</td>
<td>3.08 ± 0.51</td>
<td>2.48 ± 0.32</td>
<td>-19.5*</td>
</tr>
<tr>
<td>ml·min⁻¹·kg⁻¹</td>
<td>47.4 ± 9.4</td>
<td>38.3 ± 5.9</td>
<td>-19.2*</td>
</tr>
<tr>
<td>$V_e$ BTPS, l/min</td>
<td>116.7 ± 20.1</td>
<td>120.2 ± 20.4</td>
<td>+3</td>
</tr>
<tr>
<td>HR, beats/min</td>
<td>189 ± 8</td>
<td>181 ± 9</td>
<td>-4.2*</td>
</tr>
<tr>
<td>RER</td>
<td>1.13 ± 0.05</td>
<td>1.12 ± 0.06</td>
<td>-1.0</td>
</tr>
<tr>
<td>$S_P_O_2$ max, %</td>
<td>96.9 ± 2.0</td>
<td>79.1 ± 4.4</td>
<td>-18.3*</td>
</tr>
<tr>
<td>$S_P_O_2_E$</td>
<td>2.72 ± 0.44</td>
<td>1.64 ± 0.32</td>
<td>-27.8*</td>
</tr>
<tr>
<td>$S_P_O_2_max$, %</td>
<td>74.1 ± 0.1</td>
<td>66.1 ± 0.1</td>
<td>-10.8*</td>
</tr>
</tbody>
</table>

Values are means ± SD, Δ (%), percent decrease (−) or increase (+) from sea level to 4,338 m; $V_O_2\_m_a_x$, maximal oxygen consumption; $V_e$, expired ventilation; HR, heart rate; RER, respiratory exchange ratio; $V_e$ BTPS/$V_O_2$, ventilatory equivalent for oxygen; $S_P_O_2$ max, arterial O2 saturation at $V_O_2\_m_a_x$ by pulse oximetry; $V_e$ $thres$, ventilatory threshold. *Significant difference: sea level vs. 4,338 m, paired $t$-test, $P < 0.01$. |
data sample above and below the median of admixture and ΔSpO2 max. The ΔV02 max residuals were used as the dependent variable in these plots from the regression of ΔV02 max on both sea level V02 max and sea level SpO2 max. This allows a nonbiased assessment of the individual deviation from the overall group mean ΔV02 max, adjusting for baseline (i.e., sea level) effects. Thus positive and negative values specify individuals with larger and smaller than average decrements in V02 max, respectively, independent of the sea level V02 max and the baseline SpO2 max. In fact, control for baseline SpO2 max (sea level SpO2 max) makes little difference in this regard, because sea level saturation was uniformly high and showed little variation relative to the SpO2 max observed in Cerro de Pasco. Nevertheless, as described in MATERIALS AND METHODS, control for baseline is the correct approach to view the effect of admixture or ΔSpO2 max on ΔV02 max.

Figure 3 shows the relationship between the V02 max residuals and log(e) admixture in subgroups with the smallest (Fig. 3A) and largest (Fig. 3B) decreases in SpO2 max. Log(e) admixture was not related to ΔV02 max in the subgroup with the smallest ΔSpO2 max (mean decrease 14.7 percentage points), but significantly correlated to ΔV02 max in the subgroup with the largest ΔSpO2 max (mean decrease 20.8 percentage points) (R = 0.557, P = 0.031). Whereas subgroup analysis allows the visualization of this interaction effect, the transformation of a continuous variable (ΔSpO2 max) into a dichotomous variable results in a loss of information. Thus the reader should refer back to model 3 (Table 2) when gauging the true strength of association between admixture and ΔV02 max.

Alternately, the interaction between admixture and ΔSpO2 max may be seen in Fig. 4, which shows the relationship between the ΔV02 max residuals and ΔSpO2 max in subgroups with the lowest (Fig. 4A) and highest (Fig. 4B) levels of genetic admixture. ΔSpO2 max was not related to ΔV02 max in the subgroup with the lowest genetic admixture (mean 1.1% European genetic influence), but was highly significantly correlated to ΔV02 max in the high-admixture subgroup (mean 18.4% European genetic influence) (R = 0.652, P = 0.008).

To address the potential for confounding or spurious correlation, it is important to establish the general similarity between subgroups used in the analyses above. Comparative data in Table 5 show no significant differences for potential confounding variables between subgroups, with the exception of a significantly higher relative V02 max (ml·min⁻¹·kg⁻¹) at sea level in

Table 3. Correlation matrix for variables used in multivariate analyses of admixture effect on ΔV02 max

<table>
<thead>
<tr>
<th>ΔV02 max</th>
<th>SL-V02 max</th>
<th>ΔSpO2 max</th>
<th>SL-SpO2 max</th>
<th>ΔVthresh</th>
<th>SL-Vthresh</th>
<th>Log(e) Admixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000*</td>
<td>0.818*</td>
<td>0.588*</td>
<td>-0.270</td>
<td>0.687*</td>
<td>-0.044</td>
<td>1.000*</td>
</tr>
<tr>
<td>0.818*</td>
<td>1.000*</td>
<td>0.502*</td>
<td>0.399*</td>
<td>0.322*</td>
<td>-0.183</td>
<td>0.502*</td>
</tr>
<tr>
<td>0.588*</td>
<td>0.502*</td>
<td>1.000*</td>
<td>0.529*</td>
<td>0.522*</td>
<td>-0.238</td>
<td>1.000*</td>
</tr>
<tr>
<td>-0.270</td>
<td>0.399*</td>
<td>0.529*</td>
<td>-0.298</td>
<td>-0.265</td>
<td>-0.073</td>
<td>0.265</td>
</tr>
<tr>
<td>0.687*</td>
<td>0.322*</td>
<td>0.522*</td>
<td>-0.298</td>
<td>1.000*</td>
<td>0.029</td>
<td>1.000*</td>
</tr>
<tr>
<td>-0.044</td>
<td>-0.183</td>
<td>-0.238</td>
<td>-0.073</td>
<td>0.029</td>
<td>0.057</td>
<td>1.000*</td>
</tr>
<tr>
<td>1.000*</td>
<td>0.502*</td>
<td>1.000*</td>
<td>0.557</td>
<td>0.399*</td>
<td>1.000*</td>
<td>0.557</td>
</tr>
</tbody>
</table>

Δ, Decrement from sea level (SL) to 4,338 m; log(e) admixture, natural logarithm of genetic admixture level (%). Correlation significant: *P < 0.01, †P < 0.05.
the subgroup of subjects with the largest decrease in \( \Delta \text{SpO}_2 \text{max} \) from sea level to 4,338 m (A; \( R = 0.038 \), not significant) but are positively related to genetic admixture in the subgroup of subjects with the largest \( \Delta \text{SpO}_2 \text{max} \) (B; \( R = 0.557 \), \( P = 0.031 \)). Smallest and largest \( \Delta \text{SpO}_2 \text{max} \) subgroups represent individuals below and above the median \( \Delta \text{SpO}_2 \text{max} \). \( \Delta \text{Vo}_2 \text{max} \) residuals are from the regression of \( \Delta \text{Vo}_2 \text{max} \) on sea level \( \text{Vo}_2 \text{max} \) and \( \text{SpO}_2 \text{max} \). Residual values above and below zero represent individuals above and below the mean \( \Delta \text{Vo}_2 \text{max} \) after controlling for baseline (sea level) effects (see text for details).

Log(e), natural logarithm.

**Table 4. Covariance models explaining variability in \( \Delta \text{Vo}_2 \text{max} \)**

<table>
<thead>
<tr>
<th>Model</th>
<th>( \beta )</th>
<th>( P ) value</th>
<th>( \beta )</th>
<th>( P ) value</th>
<th>( \beta )</th>
<th>( P ) value</th>
<th>( \beta )</th>
<th>( P ) value</th>
<th>( \beta )</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>-0.991</td>
<td>1.138</td>
<td>2.814</td>
<td>-0.855</td>
<td>-0.783</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.516</td>
<td>&lt;0.001</td>
<td>0.405</td>
<td>&lt;0.001</td>
<td>0.419</td>
<td>&lt;0.001</td>
<td>0.300</td>
<td>&lt;0.001</td>
<td>0.448</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Model 3</td>
<td>-0.022</td>
<td>0.218</td>
<td>-0.036</td>
<td>0.044</td>
<td>0.016</td>
<td>0.906</td>
<td>-0.076</td>
<td>0.587</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4</td>
<td>0.222</td>
<td>0.033</td>
<td>-0.001</td>
<td>0.946</td>
<td>0.297</td>
<td>0.006</td>
<td>0.242</td>
<td>0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 5</td>
<td>0.014</td>
<td>0.018</td>
<td>0.014</td>
<td>0.018</td>
<td>0.078</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MSE, mean squared error.

The association between \( \Delta \text{Vo}_2 \text{max} \) and \( \Delta \text{Ve}_\text{thresh} \) is only evident in individuals with high Spanish admixture, but not evident in individuals with low Spanish admixture.

Additional multivariate models were run, controlling for \( \Delta \text{SpO}_2 \text{max} \) and \( \Delta \text{Ve}_\text{thresh} \) simultaneously. These models do not change the basic results presented in models 3 and 5 (Table 4). That is, the interaction between log(e) admixture and \( \Delta \text{SpO}_2 \text{max} \) remained significant after control for \( \Delta \text{Ve}_\text{thresh} \) (\( \beta = 0.011 \), \( P = 0.037 \)). Similarly, the interaction between log(e) admixture and \( \Delta \text{Ve}_\text{thresh} \) remained nonsignificant after control for \( \Delta \text{SpO}_2 \text{max} \), although the \( P \) value once again reached a relatively low level (\( \beta = 0.086 \), \( P = 0.114 \)). A final model testing for the three-way interaction among log(e) admixture, \( \Delta \text{Ve}_\text{thresh} \), and \( \Delta \text{SpO}_2 \text{max} \) revealed a marginally significant interaction in this regard (\( \beta = 0.004 \), \( P = 0.06 \)). Interpretation of this interaction is similar to the interpretation given above for two-way interaction. That is, the admixture effect is evident in individuals with large decrements in \( \text{SpO}_2 \text{max} \) and \( \text{Ve}_\text{thresh} \).
but absent in individuals showing only small decreases in SpO2 max and \( V_{\text{E}}_{\text{thresh}} \) from sea level to 4,338 m.

**DISCUSSION**

This study shows an association between Spanish admixture level and the \( V_{O2\max} \) decrement in a subset of subjects showing larger than average altitude-related decreases in SpO2 max. In these subjects, high Spanish ancestry was associated with larger \( \Delta V_{O2\max} \), consistent with the general hypothesis that Quechua natives of the highland Andes are adapted to high altitude based on population-specific genetic factors that have arisen as a consequence of natural selection. The strength of this evolutionary inference depends on the novel research design and on the choice of \( \Delta V_{O2\max} \) as the study-dependent variable. A repeated-measures design, which exploits intragroup variability in genetic admixture, has clear advantages over the traditional comparative approach i.e., Andean native vs. lowland “control.” Comparisons between groups may be confounded by many factors, including unknown levels of admixture in the study populations (4). Regarding the \( \Delta V_{O2\max} \), this phenotype has long been considered an important marker of hypoxia sensitivity or tolerance (2, 23, 47).

The study approach depends on the well-documented (and different) population history of exposure to hypobaric hypoxia between Quechua and Spanish. In essence, the admixture estimate is a construct describing the proportionate contribution of alternate population histories to the genetic makeup of an individual. This is an important point because the association between individual admixture and \( \Delta V_{O2\max} \) is to be interpreted relative to population history, not relative to the specific genetic markers that were used to derive the estimates. In other words, for the approach to work, it is not necessary that there be a direct linkage between the genetic markers and the \( \Delta V_{O2\max} \) phenotype, especially as the markers used represent only a small fraction of the total genome. Our laboratory has previously reported that admixture can show a significant correlation with specific phenotypes, even if the markers informative for admixture are not physically linked to the trait in question (36, 43). The process of admix-

**Table 5. Subgroup comparisons for selected variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lowest Admixture</th>
<th>Highest Admixture</th>
<th>Smallest ( \Delta SpO2_{max} )</th>
<th>Largest ( \Delta SpO2_{max} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>24.0 ± 4.2</td>
<td>27.0 ± 4.6</td>
<td>25.6 ± 4.2</td>
<td>25.4 ± 4.3</td>
</tr>
<tr>
<td>Height, cm</td>
<td>163.8 ± 5.6</td>
<td>164.8 ± 6.6</td>
<td>164.8 ± 7.0</td>
<td>163.9 ± 4.7</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>64.6 ± 8.9</td>
<td>66.8 ± 12.5</td>
<td>64.4 ± 11.8</td>
<td>67.7 ± 7.9</td>
</tr>
<tr>
<td>Body fat, %</td>
<td>21.3 ± 5.6</td>
<td>23.2 ± 6.7</td>
<td>21.1 ± 7.0</td>
<td>23.4 ± 5.2</td>
</tr>
<tr>
<td>Admixture, %</td>
<td>1.1 ± 0.4</td>
<td>18.4 ± 16.4*</td>
<td>13.1 ± 18.9</td>
<td>6.4 ± 7.2</td>
</tr>
<tr>
<td>SL-( V_{O2\max} ), l/min</td>
<td>3.24 ± 0.46</td>
<td>2.93 ± 0.54</td>
<td>2.93 ± 0.59</td>
<td>3.24 ± 0.40</td>
</tr>
<tr>
<td>SL-( V_{O2\max} ), ml·min(^{-1})·kg(^{-1})</td>
<td>50.5 ± 9.9</td>
<td>44.4 ± 7.3*</td>
<td>45.8 ± 6.9</td>
<td>49.0 ± 9.32</td>
</tr>
<tr>
<td>( V_{O2\max} ), 4,338 m, l/min</td>
<td>2.59 ± 0.30</td>
<td>2.38 ± 0.31</td>
<td>2.47 ± 0.38</td>
<td>2.50 ± 0.25</td>
</tr>
<tr>
<td>SL-( V_{O2\max} ), %</td>
<td>96.7 ± 1.9</td>
<td>97.0 ± 2.2</td>
<td>96.9 ± 2.3</td>
<td>96.8 ± 1.9</td>
</tr>
<tr>
<td>( SpO2_{max} ), 4,338 m, %</td>
<td>78.0 ± 4.3</td>
<td>80.1 ± 4.5</td>
<td>82.2 ± 2.6</td>
<td>75.9 ± 3.6†</td>
</tr>
<tr>
<td>SL-( V_{E\text{thresh}} ), l/min</td>
<td>2.30 ± 0.39</td>
<td>2.25 ± 0.51</td>
<td>2.10 ± 0.39</td>
<td>2.45 ± 0.45†</td>
</tr>
<tr>
<td>( V_{E\text{thresh}} ), 4,338 m, l/min</td>
<td>1.63 ± 0.38</td>
<td>1.65 ± 0.25</td>
<td>1.70 ± 0.39</td>
<td>1.59 ± 0.23</td>
</tr>
</tbody>
</table>

Values are means ± SD. Highest and lowest admixture subgroups are the highest and lowest 50th percentiles of genetic admixture rate. Smallest and largest \( \Delta SpO2_{max} \) subgroups are highest and lowest 50th percentiles of \( SpO2_{max} \) decrement from sea level to 4,338 m. *Significantly different from lowest admixture subgroup, \( P \leq 0.05 \). †Significantly different from smallest \( \Delta SpO2_{max} \) subgroup, \( P \leq 0.05 \).
The Association Between Admixture and $\Delta V_{O2_{max}}$

By itself, genetic admixture level was not related to variability in $\Delta V_{O2_{max}}$. However, after control for both sea level $V_{O2_{max}}$ and $\Delta SpO2_{max}$, strong associations between genetic admixture and $\Delta V_{O2_{max}}$ were revealed. Covariate control is an important step because, in this and previous studies (12, 17, 28, 30, 32, 41, 42, 44), the majority of the variability in the $\Delta V_{O2_{max}}$ was explained by the sea level $V_{O2_{max}}$. This effect is impressive as aerobically fit subjects typically lose two to four times more of their sea level $V_{O2_{max}}$ and show decrements in aerobic performance at even modest altitudes, i.e., <600 m (17). Similarly, this and other studies (12, 32, 41, 42, 44) demonstrate that much of the variability in the $\Delta V_{O2_{max}}$ is explained by the $\Delta SpO2_{max}$. Sea level $V_{O2_{max}}$ and $\Delta SpO2_{max}$ together explained nearly 73% of the variability in the $\Delta V_{O2_{max}}$ within our study sample. Of the remaining variability, another 7% was explained by genetic factors assessed by the admixture variable ($P = 0.041$) and an admixture-$\Delta SpO2_{max}$ interaction effect ($P = 0.018$, model 3, Table 4). Thus the effect of admixture on $\Delta V_{O2_{max}}$ depends on the magnitude of $\Delta SpO2_{max}$. Admixture effects were large between subjects with large $\Delta SpO2_{max}$ (Fig. 3B), but not detectable between individuals who showed only modest decreases in $SpO2_{max}$ (Fig. 3A). In the subset of individuals showing larger than average $\Delta SpO2_{max}$, adjusted mean values for $V_{O2_{max}}$ (ml/min) were 18% larger in the highest vs. lowest subgroups of Spanish ancestry, i.e., least squares mean values, adjusted for sea level $V_{O2_{max}}$, were ~739 and ~606 ml/min, respectively.

One interpretation of this finding is that genetic effects are more important (i.e., easier to detect) across the range of admixture when subjects are aerobically fit. This makes sense because aerobically fit individuals are, a priori, more hypoxia sensitive, given their larger $\Delta V_{O2_{max}}$ (see references cited above). In addition, athletes vs. nonathletes show a wider alveolararterial partial pressure difference during exercise and show impairments in gas exchange, even in normoxia (9, 40). These differences suggest a pulmonary limitation to exercise in athletes via ventilation-perfusion mismatch and/or diffusion limitation, both of which increase as problems with increasing exercise intensity and/or hypoxia (15, 19). Thus better pulmonary function in Quechua may be the basis for part of the admixture effect described, particularly as Andean natives are characterized by large lungs (6, 13, 18, 24) and high-pulmonary diffusion capacities (39, 46). One previous study is consistent with this hypothesis, showing higher arterial oxygen saturation during exercise in Andean natives compared with a highland-born control group of lowland ancestry (5).

This interpretation is further supported by the $V_{E_{\text{threshold}}}$ data. In this and previous studies (28, 41), $\Delta V_{E_{\text{threshold}}}$ explained a large proportion of the variability in $\Delta V_{O2_{max}}$. We also noted a trend for interaction (admixture-$\Delta V_{E_{\text{threshold}}}$) that paralleled the admixture-$\Delta SpO2_{max}$ interaction ($P = 0.095$). That is, the admixture effect tended to be evident only in individuals who showed a large decrease in $V_{E_{\text{threshold}}}$. This may be a reflection of pulmonary limitation and a decrease in arterial $O2$ saturation in the most fit subjects, as experimental studies demonstrate an increase in blood lactate concentration when arterial saturation is lowered (27). In fact, $\Delta V_{E_{\text{threshold}}}$ and $\Delta SpO2_{max}$ were strongly correlated ($R = 0.522$, $P < 0.01$, Table 3).

Alternately, the relationship between $\Delta V_{O2_{max}}$ and $\Delta SpO2_{max}$ may also be seen to depend on the genetic admixture level (Fig. 4). Theoretically, $\Delta V_{O2_{max}}$ and $\Delta SpO2_{max}$ should show a positive linear relationship (as in Fig. 4B) because of the shape of the Hb-O2 equilibrium curve (OEC). This has been demonstrated by Ferretti et al. (12), who show nonlinear relationships between the $\Delta V_{O2_{max}}$ and the fractional concentration of inspired oxygen from 0.1 (hypoxia) to 0.3 (hyperoxia) and also between the arterial saturation measured at $V_{O2_{max}}$ and fractional concentration of inspired oxygen. The nonlinearity in both relationships mirrors the OEC, and, consequently, a plot of $\Delta V_{O2_{max}}$ vs. arterial saturation (or $\Delta SpO2_{max}$) yields a linear plot. Interestingly, both aerobically fit and sedentary subjects fall on the same line (12), although aerobically fit subjects have both larger $\Delta V_{O2_{max}}$ and $\Delta SpO2_{max}$. This suggests that athletes vs. nonathletes operate on the steep part of the OEC at maximal exercise and, as a consequence, have little margin to increase oxygen flow conductance when oxygen levels decrease (12, 17, 30, 44). The absence of a positive relationship between $\Delta V_{O2_{max}}$ and $\Delta SpO2_{max}$ in low-admixture subjects is thus unexpected (Fig. 4A).

In Quechua, it may be that variability in $\Delta V_{O2_{max}}$ is not dependent on the OEC or that arterial desaturation is not a major factor driving the $\Delta V_{O2_{max}}$. If so, then other factors independent of $O2$ content may be implicated, including the cardiac output, the peripheral $O2$ diffusion (capillary to mitochondria), or the mitochondrial oxidative capacity itself. However, this study was not designed to address these specific possibilities. Also, $V_{O2_{max}}$ has been described as an integrated functional response variable, not defined (or limited) by a single factor, but rather set by the interaction of multiple $O2$ transport conductances in the lungs, circulatory system, and skeletal muscle (48). In this sense, it might be more realistic to consider the possibility that genetic adaptation in Quechua is the
result of an integrated evolutionary response involving multiple interacting components of the O₂ transport chain.

**Limitations of the Present Study**

The estimates of individual admixture were based on 22 genetic markers. These markers were selected to estimate admixture because they differed greatly in the frequency of specific alleles between Native American and European populations. They have been validated by genotyping diverse samples of Native American populations (Mayan, Southwestern Native Americans, Native Mexican, and Aymara) and European populations (Europe, Germany, and European Americans). Although this panel of markers is sufficient to obtain a precise estimate of group mean admixture (typically, with a standard error <3%), the application of these markers does not provide as precise an estimate of admixture at the individual level. The issue of quantifying the exact error around these estimates is a complicated one, and standard errors are not constant across the range of admixture. For example, standard errors are larger for individuals with high Spanish admixture. We estimate that, to obtain a precision of <3% in the individual admixture estimates, at least 80–100 ancestry-informative markers would be required. Thus it is important to mention that the estimates of individual admixture for each person show a wide confidence interval. A partial solution to this problem is to ensure a large sample size with a wide range of admixture across the study sample, as we have done here.

One weakness of this study is that arterial oxygen saturation was measured via pulse oximetry rather than from blood sampled during exercise. Fortunately, the study partially compensates for error in the SpO₂ measure via a relatively large sample size. A large sample ensures accurate group mean values, assuming that there is no bias in the measurement. While a number of studies demonstrate the general validity and accuracy of SpO₂ measures at rest and during exercise vs. blood-gas measures (20, 33, 41, 51), some studies suggest bias, especially during maximal exercise (37, 52). This external bias, or the idea that the “true” saturation cannot be measured via pulse oximetry at maximal exercise, is a minor problem in the present context. A larger potential problem would be the presence of internal bias, or the idea that measurement validity depends on conditions internal to the study. For example, it would be problematic if SpO₂ validity depended on altitude or fitness level. There is no indication that this is the case, and, moreover, the repeated-measures approach minimizes some of the potential problems in this regard. Nevertheless, blood-gas studies in Andean natives will be necessary to confirm the results of this study and to further explore the specific mechanisms that explain hypoxia tolerance in Quechua natives.

**Summary**

The results of this study are consistent with the hypothesis that Quechua are genetically adapted to hypobaric hypoxia. We have employed a novel research strategy that is based on estimating the individual genetic admixture level within a specific study group using a panel of ancestry-informative genetic markers. We demonstrate an association between the extent of Spanish ancestry and an important functional phenotype (ΔVO₂ max) that may be considered to define hypoxia tolerance. Whereas all comparative studies have some inherent problems that limit evolutionary inference (4, 16), the admixture approach applied here holds great promise. This is particularly true because the library of ancestry-informative genetic markers is expected to grow in the future, leading to increased precision in admixture estimation.

We thank the volunteers who gave their time, sweat, and DNA for this research.

**DISCLOSURES**

This work was supported in part by National Science Foundation Grant BCS-0129377 (to T. D. Brutsaert) and National Human Genome Research Institute Grant HG-002154 (to M. D. Shriver).

**REFERENCES**


