Quantitative Genetic Analysis of Arterial Oxygen Saturation in Tibetan Highlanders

CYNTHIA M. BEALL,[1] KINGMAN P. STROHL,[2] JOHN BLANGERO,[3] SARAH WILLIAMS-BLANGERO,[3] MICHAEL J. DECKER,[4] GARY M. BRITTENHAM,[5] AND MELVYN C. GOLDSTEIN[1]

Abstract This study was designed to test the hypothesis that genetic differences inferred from biological kinship relationships among individuals contribute to individual variation in percentage of oxygen saturation of arterial hemoglobin (Sa_{O_2}) in a high-altitude native population. Sa_{O_2} data were obtained by pulse oximetry from 354 nonpregnant, healthy Tibetan residents of Pen-Dri, two rural agropastoral villages at 3800–4065 m altitude in Lhasa Municipal District, Tibet Autonomous Region, China. Statistical analyses of these data from 46 pedigrees tested the hypothesis of a significant genetic contribution to Sa_{O_2} variation. The average Sa_{O_2} was 89.4 ± 0.2%, with a range of 76–97%. Additive genetic effects account for 44% of the interindividual phenotypic variation in Sa_{O_2} in the sample. Complex segregation analysis and variance decomposition analysis determined that 21% of the total phenotypic variation could be explained by a major gene influencing Sa_{O_2}. Homozygotes for the low-Sa_{O_2} allele have a mean Sa_{O_2} of 83.6%, whereas heterozygotes and homozygotes for the high-Sa_{O_2} allele have means of 87.6% and 88.3%, respectively. This confirms findings in another Tibetan sample and extends the known geographic distribution of the major gene. These results suggest the hypothesis that individuals with the dominant allele for higher Sa_{O_2} have a selective advantage in their high-altitude hypoxic environment.

The variation in percentage of oxygen saturation of arterial hemoglobin (Sa_{O_2}) of healthy people at high altitude is evidence of differences in adaptive response to hypobaric hypoxia (Beall et al. 1994; Niermeyer et al. 1995; Reeves et al. 1993). If genetic factors account for a substantial proportion of

[1] Department of Anthropology, Case Western Reserve University, Cleveland, OH 44106-7125.
[2] VA Medical Center 111j(W), Cleveland, OH 44106.
[3] Southwest Foundation for Biomedical Research, San Antonio, TX 78227-0147.
[4] Department of Anatomy, Case Western Reserve University, Cleveland, OH 44106.
[5] Division of Hematology/Oncology, Metrohealth Hospital, Cleveland, OH 44109.

Human Biology, October 1997, v. 69, no. 5, pp. 597–604.

KEY WORDS: MAJOR GENE, HIGH-ALTITUDE HYPOXIA, TIBET, ADAPTATION, HERITABILITY

Sa_{O_2} variation at high altitude, then there are implications for understanding the operation of natural selection in resident human populations, the reasons for variation in successful acclimatization among sojourners to high altitude, and the variation in clinical characteristics among sea-level patients stressed by hypoxia-inducing illnesses. Two lines of evidence support the hypothesis of genetic influence on this quantitative physiological trait. Beall et al. (1994), in a community study at the extreme altitudes of 4850–5450 m in Phala, Tibet Autonomous Region, reported Sa_{O_2} values ranging from 71% to 91% and found that additive genetic effects accounted for 41% of the total phenotypic variation independent of age, sex, and body size. Niermeyer et al. (1995), in a hospital study at 3658 m in Lhasa, Tibet Autonomous Region, reported higher Sa_{O_2} values among Tibetan infants compared with Han Chinese infants and suggested that the Tibetan advantage is explained by selection for heritable characteristics enhancing Sa_{O_2}. Here, we report the results of a study conducted at 3800–4065 m altitude to test the hypothesis of significant genetic basis for Sa_{O_2} variation in a second Tibetan community at slightly lower altitude to confirm the original finding and extend it to a larger geographic area.

Materials and Methods

Sample. The study site was Pen-Dri, two rural agropastoral villages with a population of 773 ethnic Tibetans living at 3800–4065 m altitude in Lhasa Municipal District, Tibet Autonomous Region, China. The barometric pressure was 479 torr. Each household was contacted during May-June 1993 or October-November 1993. Household members and their biological relatives 9 years and older were asked to participate in the study. Ninety-six percent of the 127 households contributed 1–7 participants each. Four hundred twenty-eight individuals, aged 9–82 years (68% of those eligible by age), participated. The participants were interviewed, and genealogical, anthropometric, and physiological data were collected.

Data Collection. Age was reported as the Tibetan animal year of birth and was converted to Western years by subtracting the birth year from 1993. Each participant answered questions about general illness symptoms, chronic cough and exertional dyspnea, symptoms of tuberculosis (afternoon sweating, weight loss, and coughing up blood), lifestyle factors, including the frequency and starting age of smoking and tea and alcohol consumption, and low-altitude exposure. One individual born in lowland eastern Tibet had lived in the study village since 1950. The rest of the participants were native life-long residents of Pen-Dri or of nearby villages at altitudes above 3600 m. Adults provided genealogical information about their own households and their extended families. These reports were cross-checked when relatives provided

the same information, and inconsistencies were reconciled during follow-up interviews.

Sa_{O_2} was measured with a Criticare 501+ finger pulse oximeter validated previously (Decker et al. 1989). Values reported are the average of 6 measurements obtained at contiguous 10-s intervals during rest in the seated position after resting for about 15 min. Measures of height and weight were obtained. Remeasurements of a subsample of study participants conducted under the same conditions as the initial measurements found excellent measurement repeatability, as demonstrated by small average differences of 1 ± 0.6% for Sa_{O_2} (n = 21), 0.2 ± 0.3 cm for height (n = 21), and 0.06 ± 0.4 kg for weight (n = 22).

Data on 354 nonpregnant, healthy (by self-report) villagers who were free of symptoms consistent with active tuberculosis infection and who did not have a chronic cough or exertional dyspnea, ranging in age from 9 to 81 years, were analyzed. Volunteers came from 46 pedigrees (groups of individuals with a common ancestor) with from 1 to 146 individuals. Eighty percent of individuals were from pedigrees that contributed five or more study participants. The 158 males averaged 28 ± 1 years of age and had an average body mass index (BMI) of 17.7 ± 0.2 kg/m^2. The 196 females averaged 30 ± 1 years of age and had an average BMI of 17.9 ± 0.2 kg/m^2.

Statistical Analyses. The statistical genetic analyses proceeded from simple classical quantitative genetics models assuming polygenic inheritance to more complex mixed major locus models allowing for the effects of specific major loci in addition to polygenic effects. Quantitative genetic analyses provide estimates of Sa_{O_2} heritability (h^2), the proportion of interindividual phenotypic variation resulting from additive genetic effects, using maximum-likelihood variance decomposition methods (Blangero 1993) in the computer program PAP (Hasstedt 1989; Hopper and Mathews 1982, 1994).

Complex segregation analyses compare the likelihoods of a set of restricted models representing various transmission hypotheses with an unrestricted general model that allows a mixture of up to three normal phenotypic distributions relating to unobservable genotypes. Four classes of restricted models explaining the familial patterning of Sa_{O_2} [sporadic, environmental (or finite mixture), classical polygenic, and mixed Mendelian models] were tested against the unrestricted general model using the unified approach of Lalouel et al. (1983). The sporadic model includes only random environmental effects; all individual trait values are independent of one another. The polygenic model allows only for polygenic inheritance. The environmental (finite mixture) model assumes random environmental factors for major factors and permits residual polygenic inheritance. The mixed Mendelian model incorporates a major genetic factor in addition to transmission probabilities fixed at their Mendelian expectations and allows for residual polygenic inheritance.

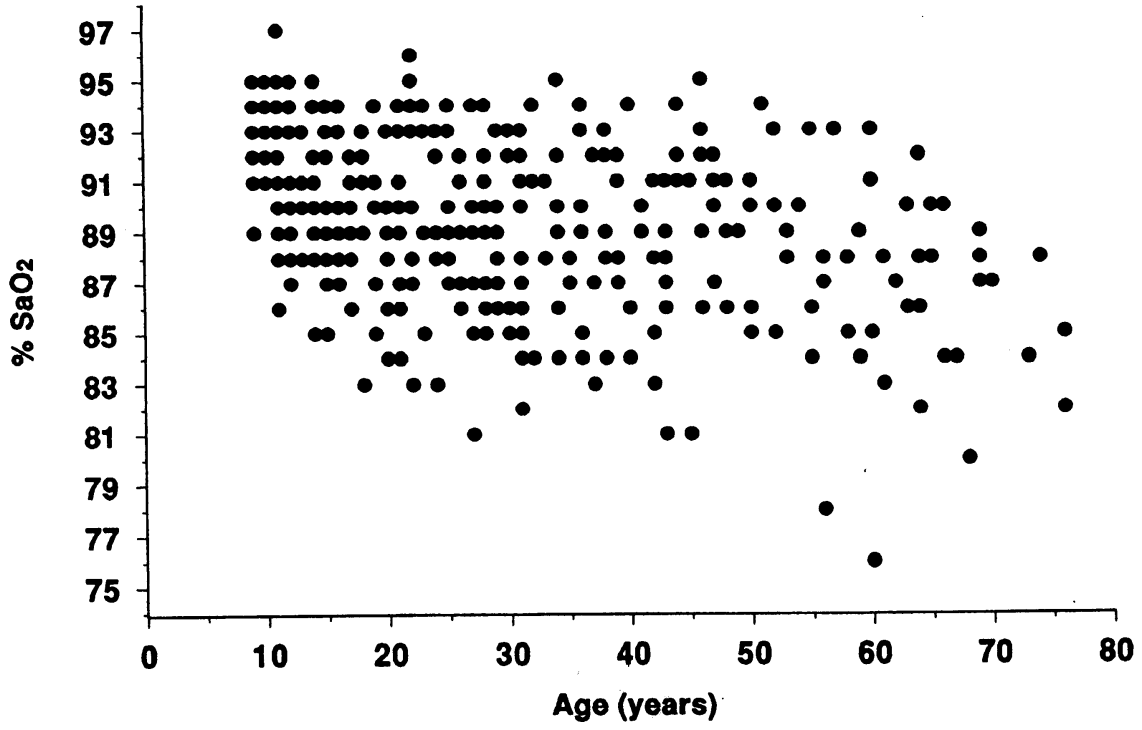

Figure 1. Scatterplot of Sa_{O_2} and age among Pen-Dri villagers (3800–4065 m).

The required parameters for each model were estimated by numerical maximization of the likelihood of the data given the assumed transmission model. Each model was compared with the unrestricted general model using likelihood ratio statistics, obtained as twice the difference between the ln-likelihood of the unrestricted and restricted models. The best model is the one requiring the fewest estimated parameters without being significantly worse than the most general model.

Results

Based on the sample of 354 individuals, the average Sa_{O_2} of the Pen-Dri sample was 89.4 ± 0.2% with a range of 76–97%. There were no significant sex differences (male average = 89.2 ± 0.3%, n = 158; female average = 89.5 ± 0.2%, n = 196). A trend toward lower Sa_{O_2} values at older ages ($r = -0.35, p < 0.05$) accounted for less than 12% of the variation (Figure 1). Six women had an average Sa_{O_2} value of 91.8 ± 0.6% measured once during pregnancy compared with 87.7 ± 0.9% measured 1–3 months after delivery (paired $t = 3.3, p < 0.05$).

The quantitative genetic analyses revealed a significant maximum-likelihood estimate of Sa_{O_2} heritability (h^2) of 0.444 ± 0.096, indicating that 44% of the variation among individuals was attributable to additive genetic effects. This analysis allowed for simultaneous estimates of covariate effects

Table 1. Evaluation of Hypotheses about the Mode of Transmission of the Sa_{O_2} Major Gene: Results of Complex Segregation Analysis of Sa_{O_2} in 354 Tibetans[a]

Parameter	*General Model*	*Environmental (Finite Mixture) Model*	*Mendelian Model*	*Polygenic Model*	*Sporadic Model*
p_A	0.166	0.182	0.258	(1)	(1)
τ_{AA}	1.000	0.182	(1)	(1)	(1)
τ_{Aa}	0.725	0.182	(0.5)	(0.5)	(0.5)
τ_{aa}	0.000	0.182	(0)	(0)	(0)
μ_{AA}[b]	83.11	81.59	82.59	87.77	87.53
μ_{Aa}	87.66	87.77	87.58	87.77	87.53
μ_{aa}	88.47	88.86	88.29	87.77	87.53
σ	2.665	2.790	2.679	3.017	3.054
h^2	0.348	0.451	0.382	0.444	(0)
χ^{2c}	–	4.24	1.40	8.61	37.13
d.f.[d]	–	~1	~1	~3	~4
p	–	0.039	0.236	0.035	<0.001

p_A, frequency of allele A with low phenotypic value.

τ, transmission probability of the subscripted genotype; that is, the probability that a parent of the subscripted genotype produces a gamete with allele A.

μ, phenotypic mean of the subscripted genotype.

σ, within-genotype phenotypic standard deviation.

h^2, heritability.

a. Values in parentheses represent fixed parameters.

b. Means reflect males, age 32 years.

c. Compares a given model with the general model.

d. The number of degrees of freedom is approximate becuase τ_{AA} was estimated on the upper boundary (1) and τ_{aa} was estimated on the lower boundary (0) of the parameter space.

of sex, age, and smoking. Complex segregation analysis tested the hypothesis that a segregating autosomal locus could explain all or part of this genetic variation. Genetic models with three phenotypic distributions attributable to one locus with two alleles were considered because three component distributions of Sa_{O_2} were required to account for the observed variation.

Table 1 summarizes the results of the complex segregation analysis evaluating transmission hypotheses and indicates that the Mendelian model provided the best fit to the general model. All the other models were rejected as being significantly different from the general model. Therefore the presence of a major gene (alleles with large quantitative effects at a segregating autosomal locus) was the most parsimonious explanation for the observed data; that is, it required the fewest estimated parameters without being significantly worse than the most general model. The estimated frequency of the allele A for low Sa_{O_2} was 0.258.

Figure 2 presents a histogram of the observed data and the theoretical genotypic distributions obtained from the parameters of the best fitting model.

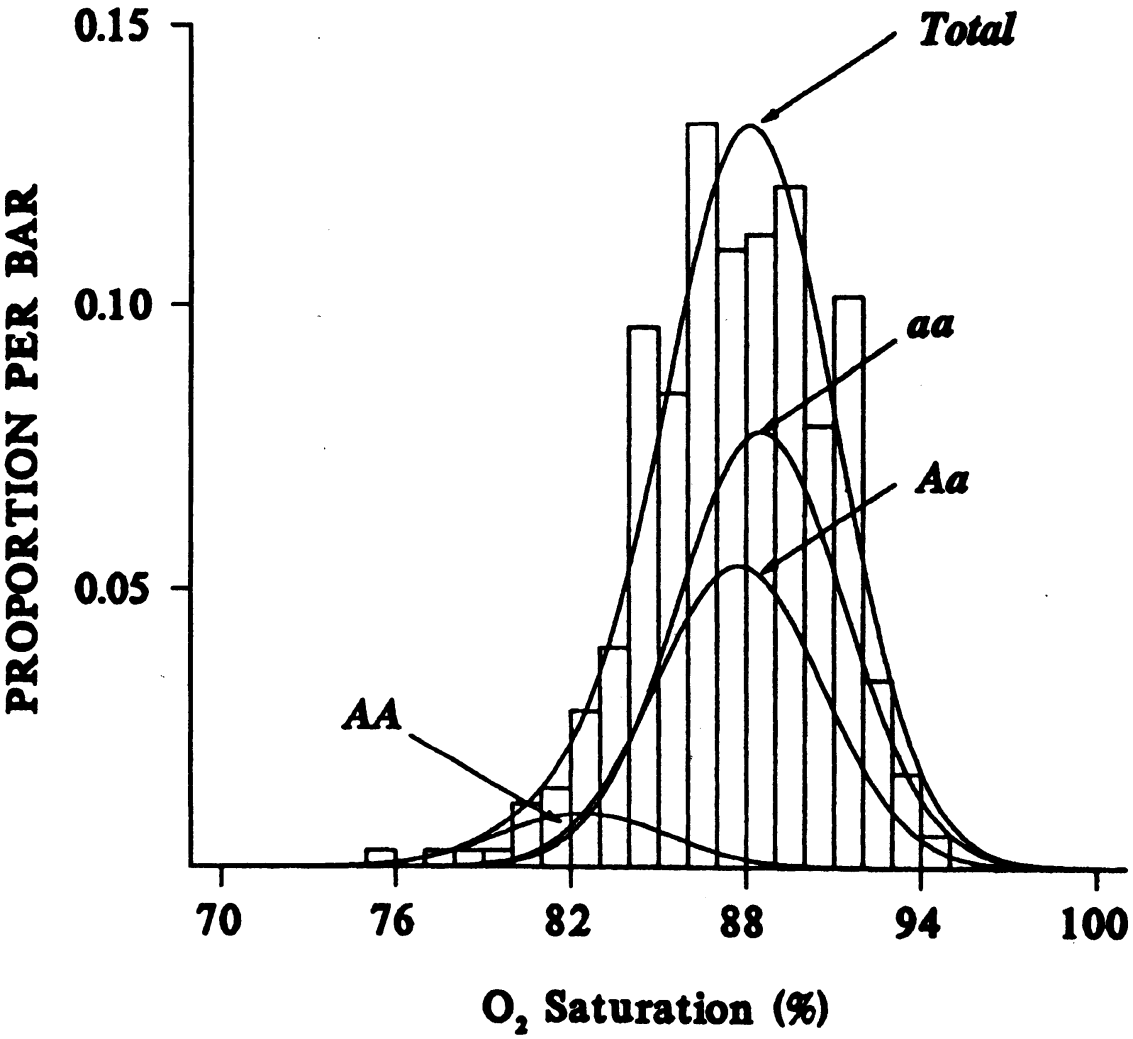

Figure 2. Observed Sa_{O_2} and theoretical genotypic distributions of the Pen-Dri sample (3800–4065 m).

Individuals in the low-Sa_{O_2} distribution have a mean Sa_{O_2} of 82.6%, whereas those in the middle and upper distributions have similar means of 87.6% and 88.3%. The total phenotypic covariance matrix was decomposed into components because of the major locus, residual polygenic effects, and random environmental effects (Blangero and Konigsberg 1991). The results of the variance decomposition indicate that the major locus accounts for 21.2% of the total phenotypic variation in Sa_{O_2} in this population, whereas polygenes account for an additional 30.2% of the variation; that is, 51.3% of the total phenotypic variation is due to genetic effects (additive and dominance).

Discussion

The results from Pen-Dri, Tibet, at 3800–4065 m, confirm and extend the earlier evidence for a major gene for Sa_{O_2} with an autosomal dominant

Table 2. Comparison of Maximum-Likelihood Estimates of the Sa_{O_2} Major Gene in Two Tibetan Samples

	Phala (4850–5450 m)		*Pen-Dri (3800–4065 m)*		*Site Differences*	
Parameter	*Maximum-Likelihood Estimate*	*Standard Error of the Estimate*	*Maximum-Likelihood Estimate*	*Standard Error of the Estimate*	χ^2	*p Value*
p_A	0.4460	0.0952	0.2578	0.1193	1.521	0.2175
μ_{AA}	78.0735	0.9887	82.5852	1.8108		
μ_{Aa}	83.9550	0.4609	87.5821	1.3311		
μ_{aa}	83.9550	0.4609	88.2882	0.7142		
σ	2.9554	0.2435	2.6789	0.1640	0.887	0.3462
h^2	0.4085	0.1358	0.4443	0.0959	0.046	0.8298
$\mu_{Aa} - \mu_{AA}$	5.8815	0.1049	4.9970	1.1835	0.554	0.4566
$\mu_{aa} - \mu_{AA}$	5.8815	0.8570	5.7030	1.9328	0.007	0.9327

p_A, frequency of allele A with low phenotypic value.
μ, phenotypic mean of the subscripted genotype.
σ, within-genotype phenotypic standard deviation.
h^2, heritability.

mode of inheritance in Phala, Tibet, at 4800–5450 m altitude (Beall et al. 1994). Table 2 presents the results of Wald tests of the null hypothesis of no differences in the maximum-likelihood estimates of the model parameters in Phala and Pen-Dri. The Wald test is based on the asymptotic normality of the maximum-likelihood estimates and uses their estimated standard errors. The genotypic means are different at the two sites, as expected, based on the altitude difference. There were no significant differences in the estimated allele frequencies, the phenotypic separation between genotypes, or the phenotypic standard deviation. This strongly suggests that the analyses have detected the same major gene effect in the two samples. There is no known gene flow between the Phala nomad sample and the Pen-Dri villager sample; the two populations live 350–400 mi apart, a 4–5-week trip by yak caravan. The presence in two geographic locations suggests that this polymorphism is characteristic of the Tibetan highlander gene pool. The 4–6% difference between low and high genotypic means in the two samples is roughly similar to the 5% difference between their means. That is, the effect of one or two alleles for high Sa_{O_2} appears physiologically equivalent to living about 1000 m lower.

These findings have implications for understanding the processes of adaptation to chronic high-altitude hypoxia. These results from two high-altitude Tibetan populations suggest that individuals with the dominant allele who are less hypoxic than their counterparts at the same altitude have a selective advantage in that environment.

Acknowledgments This research was supported by the National Science Foundation through award SBR92-21724 awarded to C.M. Beall and by the National Institutes of Health through grant HL43850 awarded to K.P. Strohl. We thank the Tibet Academy of Social Sciences, Lhasa, Tibet Autonomous Region, for hosting this field study. We are grateful to Ben Jiao and Tsering Puntsoh for assistance in data collection and to the officials and residents of Pen-Dri for their hospitality and cooperation.

Received 1 July 1996; revision received 29 January 1997.

Literature Cited

Beall, C.M., J. Blangero, S. Williams-Blangero et al. 1994. A major gene for percent of oxygen saturation of arterial hemoglobin in Tibetan highlanders. *Am. J. Phys. Anthropol.* 95:271–276.

Blangero, J. 1993. Statistical genetic approaches to human adaptability. *Hum. Biol.* 65:941–966.

Blangero, J., and L.W. Konigsberg. 1991. Multivariate segregation analysis using the mixed model. *Genet. Epidemiol.* 8:299–316.

Decker, M.J., P.L. Hoekje, and K.P. Strohl. 1989. Ambulatory monitoring of arterial oxygen saturation. *Chest* 95:717–722.

Hasstedt, S.J. 1989. *Pedigree Analysis Package.* Salt Lake City, UT: Department of Human Genetics, University of Utah.

Hopper, J.L., and J.D. Mathews. 1982. Extensions to multivariate normal models for pedigree analysis. *Ann. Hum. Genet.* 46:373–383.

Hopper, J.L., and J.D. Mathews. 1994. A multivariate normal model for pedigree and longitudinal data and the software "Fisher." *Aust. J. Statist.* 36:153–187.

Lalouel, J.M., D.C. Rao, N.E. Morton et al. 1983. A unified model for complex segregation analysis. *Am. J. Hum. Genet.* 35:816–826.

Niermeyer, S., P. Yang, D. Shanmina et al. 1995. Arterial oxygen saturation in Tibetan and Han infants born in Lhasa, Tibet. *New Engl. J. Med.* 333:1248–1252.

Reeves, J.T., R.E. McCullough, L.G. Moore et al. 1993. Sea level P_{CO_2} relates to ventilatory acclimatization at 4,300 m. *J. Appl. Physiol.* 75:1117–1122.

Samaja, M., A. Veicsteinas, and P. Cerretelli. 1979. Oxygen affinity of blood in altitude Sherpas. *J. Appl. Physiol.* 47:337–341.

Winslow, R.M., C.C. Monge, N.J. Statham et al. 1981. Variability of oxygen affinity of blood: Human subjects native to high altitude. *J. Appl. Physiol.* 51:1411–1416.