J Physiol Sci (2013) 63:183–193
DOI 10.1007/s12576-013-0257-8

ORIGINAL PAPER

VEGFA SNPs and transcriptional factor binding sites associated with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau

Norman E. Buroker · Xue-Han Ning · Zhao-Nian Zhou · Kui Li · Wei-Jun Cen · Xiu-Feng Wu · Wei-Zhong Zhu · C. Ronald Scott · Shi-Han Chen

Received: 5 December 2012 / Accepted: 28 February 2013 / Published online: 4 April 2013
© The Physiological Society of Japan and Springer Japan 2013

Abstract Mountain sickness (MS) occurs among humans visiting or inhabiting high altitude environments. We conducted genetic analyses of seven single nucleotide polymorphisms (SNPs) in the promoter region of VEGFA gene for lowland (Han) and highland (Tibetan) Chinese. The seven SNPs were evaluated in Han and Tibetan patients with acute (A) and chronic (C) MS. We compared 64 patients with AMS with 64 Han unaffected with MS, as well as 48 CMS patients with 32 unaffected Tibetans. The SNPs studied are rs699947, rs34357231, rs79469752, rs13207351, rs28357093, rs1570360, and rs2010963 which are found in the promoter ranging from −2,578 to −634 bp from the transcriptional start site (TSS), respectively. Direct sequencing was used to identify individual genotypes for these SNPs. Arterial oxygen saturation of hemoglobin (SaO_2) was found to be significantly associated with the rs699947, rs34357231, rs13207351, and rs1570360 SNPs in Han patients with AMS, while the rs2010963 SNP was found to approach significance in the AMS study group, but found to be significantly associated in the normal Tibetan study group. The Han and Tibetan control groups were found to diverge significantly for the rs28357093 and rs2010963 SNPs, as measured by genetic distances of 0.073 and 0.054, respectively. All the SNPs are found in transcriptional factor binding sites (TFBS), and their possible role in gene regulation was evaluated with regard to MS. MS was found to be significantly associated with these SNPs compared with their Han and Tibetan control groups, indicating that these nucleotide substitutions result in TFBS changes which apparently have a physiological effect on the development of high altitude sickness.

Keywords High altitude sickness · *VEGFA* promoter · SNPs · Transcriptional factor binding sites

Electronic supplementary material The online version of this article (doi:10.1007/s12576-013-0257-8) contains supplementary material, which is available to authorized users.

N. E. Buroker (✉) · C. R. Scott · S.-H. Chen
Department of Pediatrics, University of Washington, Seattle, WA 98195, USA
e-mail: nburoker@u.washington.edu

X.-H. Ning
Division of Cardiology, Seattle Children's Hospital Institute Foundation, Seattle, WA 98105, USA

Z.-N. Zhou · X.-F. Wu · W.-Z. Zhu
Laboratory of Hypoxia Physiology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

K. Li · W.-J. Cen
People's Hospital of the Tibet Autonomous Region, Lhasa, China

Introduction

The vascular endothelial growth factor (VEGF) is a family of key regulators in critical physiological and pathological angiogenesis [1] including tissue growth, wound healing, rheumatoid arthritis, proliferative retinopathies, cardiovascular disease, and cancer [2], and is a growth factor activator for angiogenesis, vasculogenesis, and endothelial cell growth. In most [3–8] but not all [9] studies, It has been shown that the VEGF is an important component of the pathogenesis of high altitude adaptation and sickness. Presently, 7 VEGF family members and 14 alternative splicing variants have been identified in humans [10–12]. Of the 14 splicing variants, 12 are VEGFA isoforms [12]

with 3 (VEGFA-121, -165 and -189) being differentially expressed in humans visiting or living in high altitude environments and also in chronic mountain sickness (CMS) patients [4, 5]. Among all family members, VEGFA is the most potent and best-known angiogenic protein and exerts its biologic effect through interaction with cell-surface receptors, which triggers a cascade of downstream dimerizations and phosphorylations [13].

High altitude sickness (HAS) arises from two different diseases: acute and chronic mountain sickness. Acute mountain sickness (AMS) is very common in lowlanders who ascend from sea level to altitudes greater than 2,600 m, and is characterized by headache, lightheadedness, breathlessness, fatigue, insomnia, anorexia, and nausea [14, 15]. Symptoms begin 2–3 h after ascent. The condition is generally self-limiting where most symptoms disappear after 2–3 days, although insomnia may persist [16]. AMS must be treated as an emergency and the illness will resolve if no further altitude is gained; however, in some cases, descent to a lower altitude may be necessary in order to reverse the condition. Chronic mountain sickness (CMS) is characterized by polycythemia and severe hypoxemia, which is reversible upon descent from high altitudes [17, 18]. Hematologic, neurologic, cardiac, and respiratory symptoms are manifestations of the disease. The most common symptoms are bone and muscle pain, headaches, dizziness, dyspnea, insomnia, tinnitus, mental fatigue, and a loss of appetite. The severity of the condition increases with advancing age [19]. CMS is a syndrome resulting from the loss of human adaptation to high altitude and can occur in permanent residents residing in this environment [20, 21].

The precise pathogenesis of AMS and CMS is not well understood, but hypoxia is likely to be a major factor [22–26]. This raises the question of why, under the same hypoxic conditions, some individuals are susceptible to AMS and CMS while others are not. Tibetans may be one of the oldest high altitude-adapted ethnic groups in the world with origins from the Neolithic period based on current genetic data [27–30]. Although AMS and CMS are different diseases and are treated differently, they both arise in humans at high altitudes. We hypothesize that Han and Tibetans when contracting these diseases may differ genetically and physiologically from individuals that do not acquire them. In order to examine this theory, we recently completed an investigation of genes from the vascular endothelial growth factor signaling pathway for AMS and CMS patients [31]. Since VEGFA is a key regulator of hypoxia, we decided to expand our investigation to include the promoter region of the gene. Here, we report on seven single nucleotide polymorphisms (SNPs) in the promoter region of the gene and their association with transcriptional factor binding sites (TFBS) among the AMS and CMS patients. In addition, the associations of these SNPs from the VEGFA promoter with the clinical data acquired from the Qinghai-Tibetan Plateau patients provides new insight into the pathogenesis of AMS and CMS in high altitude environments.

Materials and methods

Study groups

The Chinese ethnic groups studied were the Han who are considered upward migrants from low altitudes, and the Tibetans who are high altitude natives. AMS was studied in association with the Han while CMS was studied in association with the Tibetans, resulting in two different HAS groups compared with their respective ethnic controls. All mountain sickness patients in this study had been hospitalized and diagnosed at the Lhasa People Hospital (Tibet, China, at 3,658 m above sea level) from 2002 to 2008. The CMS patients and Tibetan controls normally live at 3,600–4,400 m. AMS was diagnosed by using the current consensus of mountain sickness in Tibet (Diagnosis and Therapeutics for Mountain Sickness, Xizang Autonomous Region), which is in agreement with the Lake Louise scoring system [32]. The Lake Louise consensus on the definition and quantification of altitude illness [32] was the Qinghai diagnostic criteria for measuring CMS. We sampled Han AMS patients from the hospital with symptoms of acute pulmonary edema as diagnosed by a cough accompanied by pink frothy sputum. Moist or bubbling rales in the lungs was suggestive of pulmonary oedema, showing a characteristic shadow on chest X-rays. In addition to the characteristic symptoms of severe acute mountain response, acute cerebral edema, was diagnosed by ataxia, disturbance of consciousness or coma, abnormal plantar reflexes, and papilledema. The AMS Han patients were newcomers from the low land and acquired the illness within 2 days after arriving at the higher altitude of Tibet. We also sampled Tibetan CMS patients as diagnosed by erythropoiesis, pulmonary hypertension and/or high arterial blood pressure, right ventricular hypertrophy, or right and left ventricular hypertrophy. AMS patients had an average age of 34.9 years, while CMS patients had an average age of 53.6 years. Patients with other diseases having similar clinical manifestations were excluded. Healthy Tibetan and Han people from the Lhasa area were randomly selected to serve as control subjects. All patients and controls sampled in the study signed an informed consent approved by the Human Ethics Committee of the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, and the study was approved by the appropriate institutional review board.

Sampling

Buccal brush samples were collected from 85 Hans with AMS and 45 Tibetans with CMS during the high occurrence period (spring and winter) at the Lhasa People Hospital. Samples for controls were also collected via buccal brush from unaffected Han and Tibetan individuals determined to be in good health upon physical examination by doctors at the Lhasa People Hospital. The controls consisted of 79 Han lowlanders who had travelled to the high altitude of Tibet and 31 Tibetan highlander natives. The people from both ethnic control groups had lived at the 3,600 m altitude for at least 6 months prior to being sampled.

Physiological measurements

Heart rate was measured by an electrocardiogram. Blood pressure was measured by an automatic sphygmomanometer (China Jiangsu Diving Medical Equipment). The arterial oxygen saturation levels measured on all participants of the study was conducted with a pulse oximeter device (Model 9500; Nonin Medical, Plymouth, MN, USA). All measurements were made at the Lhasa People Hospital, Tibet (3,658 m altitude).

Genotyping

Genomic DNA was extracted from the buccal brushes using the PureGene DNA method from Gentra Systems (Minneapolis, MN, USA). The DNA yield in this study ranged from 0.5 to 7.6 μg per buccal brush. We found the yield adequate for all PCR reactions conducted in the study. The Vector NTI Advance 11 computer program from Invitrogen (Carlsbad, CA, USA) was used to develop the primers for genotyping each SNP. The upstream VEGFA promoter SNPs [rs699947 (A/C) and rs34357231/rs35569394 (I/D)] were resolved using PCR sense primer 5′-TCCCTGGAGCGTTTTGGT-3′ and anti-sense primer 5′-TAAGTGCTCCCAAAGGCC-3′. The midstream *VEGFA* promoter SNPs [rs79469752 (C/T), rs13207351 (A/G), rs28357093 (A/C), and rs1570360 (A/G)] were resolved using PCR sense primer 5′-TCTGGA CAGAGTTTCCGG-3′ and anti-sense primer 5′-GCTAC CAGCCGACTTTTA-3′. The downstream VEGFA promoter SNP [rs2010963 (C/G)] was resolved using PCR sense primer 5′- AGAAGTCGAGGAAGAGAGA-3′ and anti-sense primer 5′- CGGTGTCTGTCTGTCTGT-3′. The DNA sequencing reaction mixtures were run on an ABI Veriti thermal cycler system at standard conditions and the PCR fragments were detected on 1 % TBE agarose gels. All genotyping was conducted using PCR of genomic DNA and Sanger bi-directional DNA sequencing with a Big Dye® sequencing kit (Applied Biosystems). Direct sequencing was performed for all samples on an automated sequencer ABI Prisms® 3130 Genetic Analyzer (Applied Biosystems). Electropherograms were compared with the *VEGFA* (NCBI RefSeq NM_001171626), wild-type sequences. The Mutation Surveyor DNA variant analysis software from SoftGenetics (State College, PA, USA) was used to confirm sequencing genotypes.

Statistical analysis

TIBCO® Spotfire® v.3.0 statistical software (2012) was used to analyze the clinical data (Table 1) and their association with the VEGFA SNP genotypes (Table 2). Data was evaluated with the analyses of variance within groups and single factor analysis of variance across groups. The criterion for significance was $p < 0.05$ for all comparisons.

Population genetic data analysis was conducted using Arlequin (v.2) software [33]. In our report, we list the allele, genotype frequencies, and the p values from testing Hardy–Weinberg equilibrium (HWE) among the polymorphisms in each study group as well as compute the genetic distances

Table 1 Gender, average ages and the physiological parameters: arterial oxygen saturation of hemoglobin (SaO_2), heart rate (*HR*), systolic arterial blood pressure (*BPs*) and diastolic arterial blood pressure (*BPd*) with standard errors for the AMS Han patients, CMS Tibetans patients, and the corresponding control subjects

Study group	Ethnic group	*n*	Gender	Age average	SaO_2 (%)	BPs (mmHg)	BPd (mmHg)	HR (bpm)
AMS	Han	78	Males	34.2 ± 0.8	64.4 ± 0.9*	116.8 ± 1.4	77.9 ± 1.8	105.1 ± 1.5*
		20	Females	37.0 ± 2.4	53.0 ± 1.8*	108.8 ± 2.8	73.6 ± 2.5	99.1 ± 4.1*
Control	Han	38	Males	36.1 ± 2.4	90.2 ± 0.4	114.0 ± 3.6	76.0 ± 1.8	81.0 ± 2.9
		22	Females	28.0 ± 1.8	91.0 ± 0.5	110.0 ± 3.6	68.0 ± 1.4	79.0 ± 1.6
CMS	Tibetan	37	Males	52.1 ± 1.9	82.0 ± 0.9*	136.5 ± 3.3*	96.1 ± 3.1*	86.3 ± 1.9*
		13	Females	58.1 ± 3.1	80.8 ± 1.1*	148.0 ± 5.3*	112.0 ± 2.9*	91.8 ± 4.3*
Control	Tibetan	17	Males	29.5 ± 1.7	92.8 ± 0.6	115.0 ± 2.8	78.0 ± 2.4	72.1 ± 2.5
		19	Females	30.5 ± 1.4	91.3 ± 0.5	113.0 ± 2.7	76.2 ± 2.5	82.3 ± 1.7

n number of patients or subjects sampled

* $p < 0.05$ vs. ethnic control

Table 2 Genotype, allele frequencies, Hardy–Weinberg equilibrium (HWE), genetic distance and F_{st} statistics for seven VEGFA SNPs of AMS and CMS study groups in a comparison with normal Han (H) and Tibetan (T) Chinese populations, respectively

SNP		Population	*n*	Genotypes			Allele frequencies ± SD		HWE *p* value	Genetic distance pairwise differences			
	SG			**AA**	**AC**	**CC**	**A**	**C**		**SG**	1	2	3
rs699947(−)2578	1	AMS (H)	63	1	27	35	0.230 ± 0.037	0.77 ± 0.037	0.157	1			
	2	Normal (H)	59	2	28	29	0.271 ± 0.041	0.729 ± 0.041	0.187	2	−0.004		
	3	CMS (T)	48	2	16	30	0.208 ± 0.041	0.792 ± 0.041	1	3			
	4	Normal (T)	32		11	21	0.172 ± 0.047	0.828 ± 0.047	0.556	4		0.015	−0.009
				II	**ID**	**DD**	**I**	**D**			1	2	3
rs34357231 or rs35569394 (−)2550-2549ins 18 bp I/D	1	AMS (H)	64	1	28	35	0.234 ± 0.037	0.766 ± 0.037	0.156	1			
	2	Normal (H)	59	2	29	28	0.28 ± 0.041	0.72 ± 0.041	0.193	2	−0.003		
	3	CMS (T)	48	2	16	30	0.208 ± 0.041	0.792 ± 0.041	1	3			
	4	Normal (T)	32		11	21	0.172 ± 0.047	0.828 ± 0.047	0.556	4		0.019	−0.009
				CC	**CT**	**TT**	**C**	**T**			1	2	3
rs79469752 (−)1203	1	AMS (H)	52	44	8	0	0.923 ± 0.026	0.077 ± 0.026	1	1			
	2	Normal (H)	59	55	4	0	0.962 ± 0.018	0.037 ± 0.018	1	2	0.005		
	3	CMS (T)	35	28	7	0	0.90 ± 0.036	0.10 ± 0.036	1	3			
	4	Normal (T)	28	25	3	0	0.946 ± 0.030	0.054 ± 0.030	1	4		−0.01	−0.001
				AA	**AG**	**GG**	**A**	**G**			1	2	3
rs13207351 (−)1190	1	AMS (H)	48	2	26	20	0.313 ± 0.047	0.687 ± 0.047	0.096	1			
	2	Normal (H)	56	5	27	24	0.330 ± 0.045	0.670 ± 0.044	0.762	2	−0.009		
	3	CMS (T)	40	2	16	22	0.25 ± 0.049	0.75 ± 0.049	1	3			
	4	Normal (T)	24	2	11	11	0.313 ± 0.068	0.687 ± 0.068	1	4		−0.014	−0.007
				AA	**AC**	**CC**	**A**	**C**			1	2	3
rs28357093 (−)1179	1	AMS (H)	53	37	15	1	0.839 ± 0.036	0.160 ± 0.036	1	1			
	2	Normal (H)	59	47	11	1	0.889 ± 0.029	0.110 ± 0.029	1	2	0.002		
	3	CMS (T)	46	36	10	0	0.891 ± 0.032	0.109 ± 0.033	1	3			
	4	Normal (T)	29	29	0	0	1			4		0.073*	0.078*
				AA	**AG**	**GG**	**A**	**G**			1	2	3
rs1570360 (−)1154	1	AMS (H)	46	0	18	28	0.196 ± 0.042	0.804 ± 0.042	0.171	1			
	2	Normal (H)	55	0	22	33	0.200 ± 0.038	0.800 ± 0.038	0.099	2	−0.01		
	3	CMS (T)	44	0	11	33	0.125 ± 0.036	0.875 ± 0.036	1	3			
	4	Normal (T)	27	0	5	22	0.093 ± 0.040	0.907 ± 0.040	1	4		0.028	−0.009
				CC	**CG**	**GG**	**C**	**G**			1	2	3
rs2010963 (−)634	1	AMS (H)	61	9	35	17	0.434 ± 0.045	0.566 ± 0.045	0.301	1			
	2	Normal (H)	64	7	35	22	0.383 ± 0.043	0.587 ± 0.048	0.291	2	−0.003		
	3	CMS (T)	47	11	27	9	0.521 ± 0.052	0.479 ± 0.052	0.39	3			
	4	Normal (T)	30	7	20	3	0.567 ± 0.065	0.433 ± 0.065	0.075	4		0.054*	−0.009

*$p < 0.05$ computed from F_{stat} statistics

Table 3 Odds ratio (*OR*), standard error (*SE*) of the OR, 95 % confidence interval (*CI*), Chi square and *p* values for seven VEGFA SNPs in comparison of alleles from the AMS and CMS study groups with their Han (*H*) and Tibetan (*T*) control study groups, respectively

SNP	Study group	2*n*	Alleles		OR[a]	SE	95 % CI	Chi square	*p* value
			A	**C**					
rs699947 (−)2578	AMS (H)	126	29	97	0.803	0.296	0.432–1.494	0.35	0.554
	Normal (H)	118	32	86					
	CMS (T)	96	20	76	1.268	0.416	0.524–3.102	0.135	0.713
	Normal (T)	64	11	53					
			I	**D**					
rs34357231 rs35569394 (−)2550–2549ins 18 bp I/D	AMS (H)	128	30	98	0.788	0.293	0.427–1.455	0.445	0.505
	Normal (H)	118	33	85					
	CMS (T)	96	20	76	1.268	0.416	0.524–3.102	0.135	0.713
	Normal (T)	64	11	53					
rs79469752 (−)1203	AMS (H)	104	96	8	0.842	0.628	0.275–2.580	0.005	0.944
	Normal (H)	118	110	8					
	CMS (T)	70	63	7	0.509	0.715	0.099–2.342	0.392	0.531
	Normal (T)	56	53	3					
			A	**G**					
rs13207351 (−)1190	AMS (H)	96	30	66	0.921	0.298	0.493–1.722	0.016	0.9
	Normal (H)	112	37	75					
	CMS (T)	80	20	60	0.733	0.325	0.309–1.745	0.317	0.573
	Normal (T)	48	15	33					
			A	**C**					
rs28357093 (−)1179	AMS (H)	106	89	17	0.648	0.396	0.279–1.499	0.819	0.366
	Normal (H)	118	105	13					
	CMS (T)	92	82	10					
	Normal (T)	58	58						
			A	**G**					
rs1570360 (−)1154	AMS (H)	92	18	74	0.973	0.355	0.459–2.060	0	1
	Normal (H)	110	22	88					
	CMS (T)	88	11	77	1.4	0.569	0.415-4.962	0.102	0.749
	Normal (T)	54	5	49					
			C	**G**					
rs2010963 (−)634	AMS (H)	122	53	69	1.238	0.258	0.724–2.119	0.492	0.483
	Normal (H)	128	49	79					
	CMS (T)	94	49	45	0.833	0.332	0.412–1.681	0.148	0.7
	Normal (T)	60	34	26					

[a] Odds ratio between alleles of mountain illness and control populations

between the patient and control groups (Table 2). Genetic distances were used to measure the divergence between study groups. The genetic distance can range between 0 and 1, where 1 is complete divergence and 0 indicates that two study groups are genetically identical [34]. Smaller genetic distances indicate a close genetic relationship, whereas larger genetic distances indicate a more distant genetic relationship. F_{stat} is used in association with genetic distance to express the proportion of genetic diversity due to allele frequency differences among study groups [35]. We also computed the odds ratio (OR), confidence intervals, Chi-square and *p* values between the mountain illness and control study groups for alleles of polymorphic loci (Table 3) as a second method of statistically analyzing the genetic data because of the small sample size in this study. Haplotype estimates [36] and linkage disequilibrium (LD) [36, 37] were computed among the *VEGFA* SNPs for each of the four study groups (Fig. 1).

Bonferroni's correction was used for the physiological parameters, HWE, F_{stat}, and OR tests which assumes that a 0.05 significance level obtained for the 8 physiological parameters, 28 HWE, 21 F_{stat}, and 13 OR tests yields an actual significance threshold of 0.0063, 0.0018, 0.0023, and 0.0038, respectively [38].

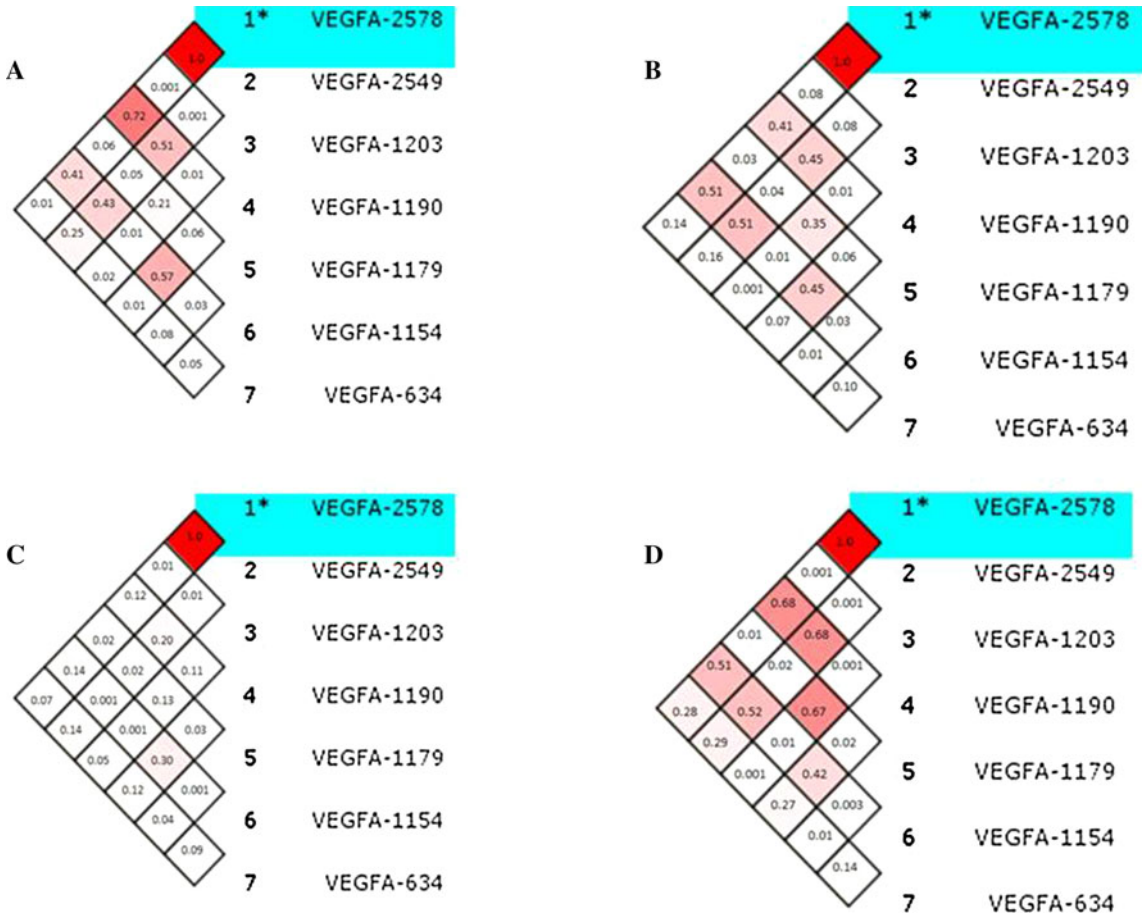

Fig. 1 Different patterns of linkage disequilibrium among seven VEGFA promoter SNPs in the four study groups: **a** control Han Chinese, **b** AMS patients, **c** control Tibetan Chinese, and **d** CMS patients. The degree of genetic linkage between the seven SNPs is estimated as r^2 values where *dark red* color ($r^2 = 1$) in the linkage disequilibrium pattern means that there exists strong pairwise linkage disequilibrium between adjacent SNPs and no color means that LD is weak or non-existent ($r^2 > 0$). The SNPs are labeled by their nucleotide distance in base pairs from the transcription start site (TSS). The tagged (*blue color*) VEGFA-2578 SNP (rs699947) is furthest from the TSS

Transcriptional factor binding sites

The JASPAR CORE database [39, 40] and ConSite [41] were used to identify transcriptional factor binding sites (TFBS) associated with the seven SNPS in the VEGFA promoter. JASPAR is a collection of transcription factor DNA-binding preferences used for scanning genomic sequences where ConSite is a web-based tool for finding cis-regulatory elements in genomic sequences. The SNP locations within TFBS for haplotypes found in each study group are listed in the Supplementary material. The Vector NTI Advance 11 computer program was used to locate the TFBS in the VEGFA promoter.

Results

Physiological study

The gender, average age and physiological parameters for the mountain sickness (AMS and CMS) and normal control study groups are listed in Table 1. The arterial oxygen saturation levels (SaO_2) were significantly lower in both the AMS and CMS study groups while the heart rate (HR) was significantly higher in both mountain sickness groups compared to their ethnic controls. Blood pressure (BPs/d) was found to be significantly higher in only the CMS Tibetan group. It should be noted that the average age of the CMS patients was 53.6 compared to 30.1 for the Tibetan control group; however, this would not be expected to change the outcome of this study because there is little difference in the SNP allele frequencies between the two groups as measured by their genetic distances (Table 2), with the exception of the rs28357093 SNP as discussed below. The measured physiological values of the control study groups fall within the normal ranges of residents at the Lhasa region.

Genetic study

We examined the association of SNPs from the *VEGFA* promoter with the AMS, CMS, and normal Han and

Tibetan study groups (Tables 2, 3). These polymorphisms were found to be in Hardy–Weinberg equilibrium (HWE) for all Chinese study groups (Table 2). During further examination of the genetic data, we found the *VEGFA*-A allele for the rs28357093 SNP (A/C) to be fixed in the Tibetan control resulting in a significant divergence between the CMS and Tibetan control study groups as measured by F_{stat} statistics on the genetic distance (0.078; Table 2). This also resulted in a divergence for the SNP between the Han and Tibetan control groups as measured by F_{stat} statistics on the genetic distance (0.073; Table 2). There were no *VEGFA*-AA genotypes observed for the rs1570360 SNP (A/G) among all four study groups (Table 2). A divergence was also found for the rs2010963 SNP between the Han and Tibetan control groups as measured by F_{stat} statistics on the genetic distance (0.054; Table 2). Odds ratio testing of the seven SNPs between the high altitude sickness groups and their respective control revealed no significant difference between the sick and healthy groups (Table 3). After applying the Bonferroni corrections to the F_{stat} and OR statistical tests, we found no changes in the above significance levels.

Linkage disequilibrium (LD) values were computed between the seven *VEGFA* SNPs for each study group (Fig. 1). Strong LD was found between the *VEGFA* SNPs at −2,578 bp (A/C) and −2,549 bp (I/D) due to their close proximity within 29 bp of each other. The 18 bp (I/D) of the rs34357231/rs35569394 SNPs create a strong linkage block for all the study groups (Table 2; Fig. 1). This LD extends to the midstream SNPs located at −1,190 and −1,154 bp which is about the same between the AMS and Han control study groups but appears much stronger in the CMS study group compared to its Tibetan control group (Table 2; Fig. 1). The LD between SNPs (rs13207351 and rs1570360) at −1,190 and −1,154 bp, respectively and between SNPs (rs79469752 and rs28357093) at −1,203 and −1,179 bp, respectively, among the study groups has previously been reported [42].

VEGFA promoter SNPs and haplotypes

The rs34357231/rs35569394 SNPs are created by an 18-bp insertion/deletion (I/D) polymorphism located in the promoter at −2,549 bp from the TSS [43]. This polymorphism probably causes most of the above LD. The difference between the two SNPs depends on where the deletion comes out of the promoter sequence. Sixteen base pairs of the deletion are the same for both SNPs. The additional two G bp are added at the 3′ end of the 16-bp sequence which is allele D of the rs34357231 SNP while the two G bp are added at the 5′ end of the 16-bp sequence which is allele D′ of the rs35569394 SNP (Fig. 2). There were only four D′ alleles observed in the entire study and those were pooled with the rs34357231 D alleles of the respective study group. Two D′ alleles were found in the CMS group and two were found in the Tibetan control group.

Haplotypes and frequencies for the seven SNPs and allelic associations with TFBS for the four study groups are found in the Supplementary material. The TFBS presented in bold are associated only with that given allele for each SNP while the TFBS with regular type are found associated with both alleles of the SNP. As an example, the TFBS (HIF1α:ARNT) of the hypoxia inducible factor-1α and aryl hydrocarbon receptor nuclear translocator TFs contains the VEGFA-C allele but not the VEGFA-A allele of the rs699947 SNP. The HIF1α and ARNT TFs play an essential role in cellular and systemic responses to hypoxia. The HIF1α:ARNT binding site in the promoter is the only

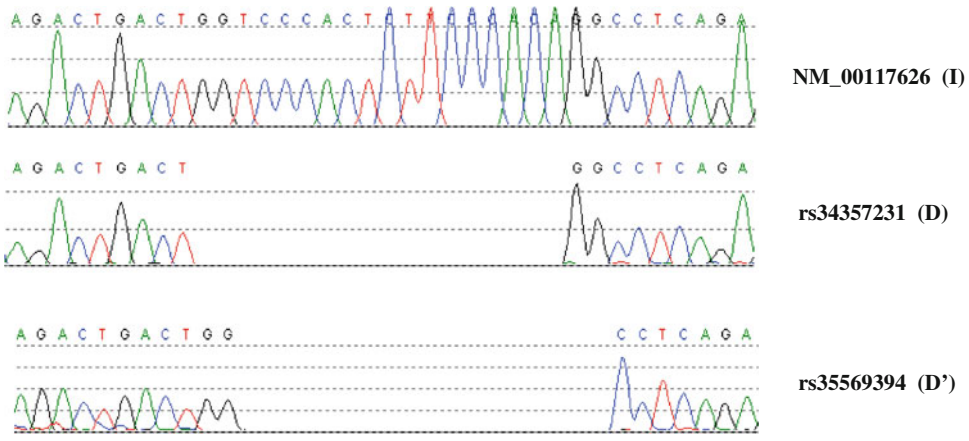

Fig. 2 VEGFA SNPs (rs34357231 and rs35569395) alleles (I, D, and D′) identified within the electropherograms from the direct DNA sequencing data. The reference sequence NM_00117626 contains the 18-bp inserted (I) allele while the rs34357231 SNP illustrates a D allele and the rs35569395 SNP illustrates a D′ allele each having an 18-bp deletion. The difference between the two deletion alleles results from whether the deleted sequence contains two G nucleotides on the 3′ sequence (rs34357231) or on the 5′ sequence (rs35569394)

one found in the VEGFA gene when surveyed with the Vector NTI Advance 11 computer program.

The first haplotype containing the rs699947 VEGFA-C allele and the HIF1α:ARNT TFBS is the most common haplotype among all study groups and more frequent in the Tibetan than in the Han study groups (Supplementary material). In fact, haplotypes 1 and 2 are found over 50 % of the time in Tibetans and over 45 % of the time in Han Chinese, while the third haplotype is found less than 20 % of the cases in the four study groups. All seven SNPs are associated with TFBS specific for a given allele. The total number of haplotypes found in the study groups were 10 (NH), 16 (AMS), 11 (NT), and 10 (CMS).

Clinical and genetic associations

The *VEGFA* SNPs were analyzed for associations with age, red blood cell count, hematocrit, hemoglobin, oxygen saturation level of hemoglobin (SaO_2), heart rate and blood pressure among the mountain sickness groups and their respective controls. A significant association ($F_{stat} = 3.47$, $p < 0.05$) was found between the physiological parameter SaO_2 and the rs699947 and rs34357231 SNPs while a nearly significant association ($F_{stat} = 2.99$, $p < 0.07$) was observed for the rs2010963 SNP from the promoter region of AMS patients (Fig. 3). A significant association ($F_{stat} = 6.18$, $p < 0.01$) was also found between SaO_2 and the rs2010963 SNP within the normal Tibetan study group (Fig. 4) but not the Tibetan CMS patientsn which probably results from the age differential between the two Tibetan study groups (Table 1), suggesting that younger Tibetans maybe more influenced by SaO_2 than older individuals. If the younger Tibetan population having this SNP variation in the regulatory region of the *VEGFA* gene is influenced by SaO_2 during the aging process then the outcome may lead to CMS later in life [44, 45]. We have previously reported that other VEGFA promoter SNPs have been associated with SaO_2 and HR among AMS patients while a similar association was found between the red blood cells (RBCs) parameter and the rs1570360 SNP among CMS patients [31]. No other associations were found between the SNPs and physiological parameters; however, two SNPs (rs699947 and rs34357231) were found to approach a significant correlation with SaO_2 ($F_{stat} = 2.38$, $p < 0.11$) for the CMS study group.

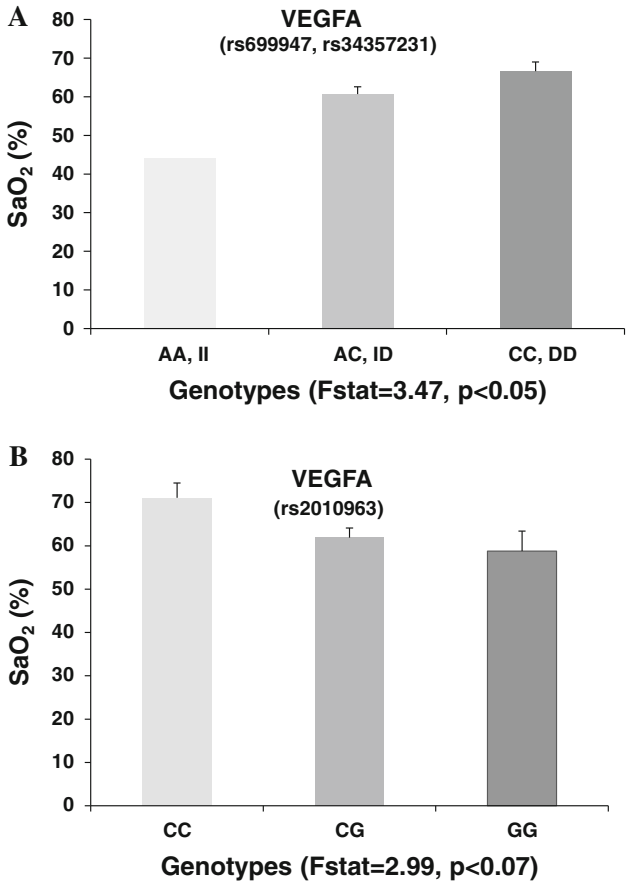

Fig. 3 VEGFA genotypes for SNPs **a** rs699947, rs34357231 and **b** rs2010963 and their significant association with arterial oxygen saturation of hemoglobin (SaO_2) in acute mountain sickness (AMS) patients. The number of AMS patients with the VEGFA SNP genotypes is found in Table 2. The SaO_2 data are expressed as the mean ± SEM

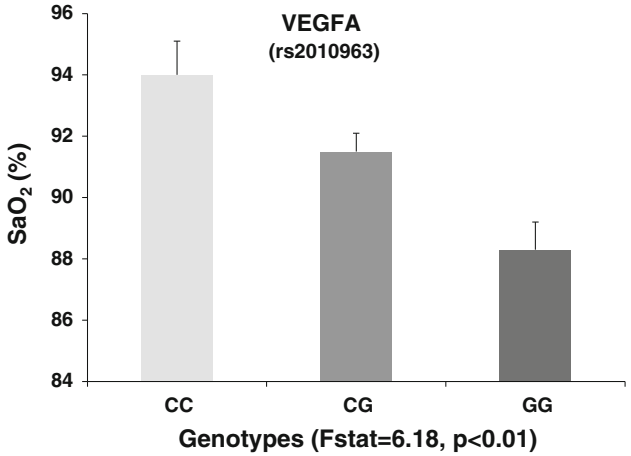

Fig. 4 VEGFA SNP (rs2010963) genotypes and their significant association with arterial oxygen saturation of hemoglobin (SaO_2) for the normal Tibetan study group. The number of normal Tibetan individuals with the VEGFA SNP genotypes is found in Table 2. The SaO_2 data are expressed as the mean ± SEM

Discussion

It has been shown that VEGFA promoter SNPs rs699947 (A/C), rs1570360 (A/G) and rs2010963 (C/G) have been associated with amyotrophic lateral sclerosis (ALS) [46]

where Swedish, Belgium and English individuals with the haplotypes AAG and AGG, respectively, have a significantly greater risk of ALS. These haplotypes lowered circulating VEGFA levels in vivo and reduced gene transcription relative to the CGC haplotype in both normoxic and hypoxic conditions. In a human myoblasts study, the same three SNPs were also found to impact VEGFA gene expression, and maximal oxygen consumption [47]. These SNPs have also been found to be associated with VEGFA protein expression in MCF7 breast cancer cells [48]. In this report, we analyzed seven VEGFA promoter SNPs including the three mentioned in the above investigations in relation to high altitude sickness in Chinese patients. We also identified the VEGFA TFBS in the promoter region which are affected by each of the seven SNPs (Supplementary material). From the Supplementary material, it can be seen that all seven SNPs alter the TFBS in the promoter of VEGFA. The seven SNPs control which TFs bind the promoter region of VEGFA and regulate the gene. Using the TFBS HIF1α:ARNT example as mentioned in "Results", the rs699947 SNP VEGFA-C allele provides a binding location for the HIF1α:ARNT protein dimer to attach the promoter of VEGFA and regulate the gene, where the VEGFA-A allele eliminates this binding site. In high altitude environments like Tibet, the oxygen levels are greatly reduced, which results in hypoxia conditions for individuals inhabiting or visiting such environments. Hypoxia leads to an increase in HIF activity that induces the expression of genes which mediates the adaptive responses through glycolytic enzymes, hemeoxygenase, vascular endothelial growth factor, and erythropoietin [49]. Since the HIF1α:ARNT binding site is under the control of the rs699947 SNP and is the only HIF1α:ARNT site in the VEGFA gene (see "Results"), the presence or absence of this binding site should have a tremendous impact on individuals inhabiting or visiting high altitude environments. However, instead of using the HIF-1α:ARNT TF, a switch could occur to the hypoxia-inducible factor-1α (HIF-1α) TF and its hypoxia-responsive elements (HRE, ACGTC, GCGTG or TACGTGGG [50]) that are also located in the promoter and are unaffected by known SNPs.

Chinese HAS patients with either AMS or CMS display a significant decrease in SaO_2 (Table 1). An examination of the allele frequency data for the seven VEGFA SNPs revealed little difference between the HAS patients and their respective controls (Tables 2, 3) with the exception of the rs28357093 SNP (A/C), where the *VEGFA*-A allele is fixed in the Tibetan control resulting in a significant divergence between the CMS and Tibetan control study groups as measured by F_{stat} statistics on the genetic distance (0.078; Table 2) as previously reported [42]. Also, a significant divergence in VEGFA SNPs (rs28357093 and rs2010963) allele frequencies as measured by F_{stat} statistics ($p < 0.05$) on the genetic distance (0.073 and 0.054, respectively) was observed between the Han and Tibetan control groups, respectively (Table 2). The seven SNPs were found to be in HWE for all study groups (Table 2); however, a visual examination of the table reveals that there are a very low number of homozygous individuals consisting of the allele with the minor frequency for these SNPs. We believe that this results from tight linkage associations between SNPs groups throughout the promoter region as indicated by the heavy LD between the SNPs (Fig. 1). In a previous study, we found the VEGFA SNPs rs13207351 and rs1570360 to be in LD as were the rs28357093 and rs79469752 SNPs among all four study groups [42]. In this study, we found the SNPs rs699947 and rs34357231 to be in heavy LD ($r^2 = 1$) and also in LD with rs13207351 and rs1570360 (Fig. 1), which consists of a four SNP haplotype group. The heavy LD is no doubt due to the 18-bp I/D polymorphism of rs34357231 (Fig. 2), which creates a linkage block with little chance of recombination between closely linked SNPs. The 18-bp inserted sequence contains the AGGCCT palindrome at the 5′ end of the deleted sequence which creates two TFBS (CREB1 and NFIC) whose TFs are known to bind palindromes (cf. Supplementary material). The rs28357093 and rs79469752 SNPs appear to be part of a second linkage group intervening between the SNPs of the first linkage group (Fig. 1). The two linkage groups may be regulated by the chromatin insulator-binding factor (CTCF) which has a binding site (TCACTAGGGGGCGC) at −1,617 bp from the TSS [51] and falls within the domain of the two linkage groups. CTCF is a highly conserved zinc finger TF implicated in diverse genomic regulatory functions, including transcriptional activation/repression, insulation, and imprinting [52], which restricts upstream enhancers from activating VEGFA and thereby restraining angiogenesis [51]. The CTCF protein has been shown to be involved with chromatin loop formation and co-associates with the zinc finger domains [52], such as the Zfx TFBS found to be associated with four of the seven SNPs in the promoter of VEGFA (cf. Supplementary material).

The seven VEGFA promoter SNPs were examined in relation to the clinical physiological parameters of HAS [31, 53, 54] and their respective control groups. The results of four of these SNPs (rs79469752, rs13207351, rs28357093, rs1570360) have previously been reported [31]. The remaining three SNPs were found to have a significant association with SaO_2 (Figs. 3, 4). The tightly linked SNPs (rs699947 and rs34357231) which segregate as a single linkage group displayed a significant ($F_{stat} = 3.47$, $p < 0.05$) association with SaO_2 (Fig. 3a) and the rs2010963 SNP approaches a significant ($F_{stat} = 2.99$, $p < 0.07$) association with SaO_2 (Fig. 3b) in the AMS study group. The Tibetan control group also

displayed a significant ($F_{stat} = 6.18$, $p < 0.01$) association with SaO_2 (Fig. 4), indicating that these VEGFA promoter SNPs appear to be important indicators for people inhabiting or visiting high altitude environments in that certain genotypes are preferentially favored under low oxygen conditions.

The rs699947, rs1570360, and rs2010963 VEGFA SNPs have been studied the most [46–48] in relation to gene regulation where the CGC or CGG haplotypes of these SNPs have strong associations with oxygen consumption. In the present HAS study, these haplotypes have the highest occurrence among all study groups (Supplementary material). These SNPs and the rs1327351 SNP [31] have been significantly associated with SaO_2 in HAS patients indicating a strong association with hypoxia. We believe that the presence of the rs699947 VEGFA-C allele which provides a HIF1α:ARNT binding site plays an important role in an individual's ability to respond to hypoxia, while the VEGFA-A allele which eliminates this TFBS is detrimental to the individual's ability to respond to hypoxia. Although there is little difference in the SNP allele frequencies between the control groups and the HAS groups (Tables 2, 3), the genetic difference lies in the linkage disequilibrium (LD) between these SNPs for the control and HAS groups (Fig. 1), where there are changes in LD levels between the control and HAS groups, especially for CMS. This disequilibrium reflects unfavorable haplotype combinations that are expressed in HAS groups compared to their controls and are emphasized in the Supplementary material table along with transcriptional factor binding site (TFBS) changes.

Acknowledgments We thank Dr. Rhona Jack, Department of Laboratory Medicine, Seattle Children's Hospital Institute Foundation for her constructive criticism in reviewing this manuscript. This study was supported in part by the grants from Children's Hospital and Regional Medical Center (HR5836), National Natural Science Foundation (No. 38970307 and No. 30393130) and National Basic Research Program of China "973" (No. 2006CB504100).

Conflict of interest Authors declare no competing financial interests.

References

1. Nowak DG, Woolard J, Amin EM, Konopatskaya O, Saleem MA, Churchill AJ, Ladomery MR, Harper SJ, Bates DO (2008) Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. J Cell Sci 121:3487–3495
2. Woolard J, Bevan HS, Harper SJ, Bates DO (2009) Molecular diversity of VEGF-A as a regulator of its biological activity. Microcirculation 16:572–592
3. Maloney J, Wang D, Duncan T, Voelkel N, Ruoss S (2000) Plasma vascular endothelial growth factor in acute mountain sickness. Chest 118:47–52
4. Appenzeller O, Minko T, Pozharov V, Bonfichi M, Malcovati L, Gamboa J, Bernardi L (2003) Gene expression in the Andes; relevance to neurology at sea level. J Neurol Sci 207:37–41
5. Appenzeller O, Minko T, Qualls C, Pozharov V, Gamboa J, Gamboa A, Wang Y (2006) Gene expression, autonomic function and chronic hypoxia: lessons from the Andes. Clin Auton Res 16:217–222
6. Patitucci M, Lugrin D, Pages G (2009) Angiogenic/lymphangiogenic factors and adaptation to extreme altitudes during an expedition to Mount Everest. Acta Physiol 196:259–265
7. Walter R, Maggiorini M, Scherrer U, Contesse J, Reinhart WH (2001) Effects of high-altitude exposure on vascular endothelial growth factor levels in man. Eur J Appl Physiol 85:113–117
8. Gao W, Gao Y, Zhang G, Song L, Sun B, Shi J (2005) Hypoxia-induced expression of HIF-1alpha and its target genes in umbilical venous endothelial cells of Tibetans and immigrant Han. Comparative biochemistry and physiology. Toxicol Pharmacol 141:93–100
9. Dorward DA, Thompson AA, Baillie JK, MacDougall M, Hirani N (2007) Change in plasma vascular endothelial growth factor during onset and recovery from acute mountain sickness. Respir Med 101:587–594
10. Dai J, Rabie AB (2007) VEGF: an essential mediator of both angiogenesis and endochondral ossification. J Dent Res 86:937–950
11. Harper SJ, Bates DO (2008) VEGF-A splicing: the key to anti-angiogenic therapeutics? Nat Rev Cancer 8:880–887
12. Xu J, Dou T, Liu C, Fu M, Huang Y, Gu S, Zhou Y, Xie Y (2011) The evolution of alternative splicing exons in vascular endothelial growth factor A. Gene 487:143–150
13. Holmes DI, Zachary I (2005) The vascular endothelial growth factor (VEGF) family: angiogenic factors in health and disease. Genome Biol 6:209
14. Hackett PH, Roach RC (2001) High-altitude illness. N Engl J Med 345:107–114
15. Bartsch P, Bailey DM, Berger MM, Knauth M, Baumgartner RW (2004) Acute mountain sickness: controversies and advances. High Alt Med Biol 5:110–124
16. Ning XH (2006) Health care at high altitude—self-care universal health book. Shanghai Science and Technology, Shanghai, pp 66–68
17. Monge C (1943) Chronic mountain sickness. Physiol Rev 23:166–184
18. Winslow RM, Monge CC (1987) Hypoxia, polycythemia, and chronic mountain sickness. Johns Hopkins University Press, Baltimore
19. Moore LG (2001) Human genetic adaptation to high altitude. High Alt Med Biol 2:257–279
20. Wu TY, Li WS, Wei LY et al (1998) A preliminary studies on the diagonosis of chronic mountain sickness in tibetan populations. Press Committee of the 3rd World Congress on Mountain Medicine and High Altitude Physiology, Matsumoto
21. Leon-Velarde F, McCullough RG, McCullough RE, Reeves JT (2003) Proposal for scoring severity in chronic mountain sickness (CMS). Background and conclusions of the CMS Working Group. Adv Exp Med Biol 543:339–354
22. West JB (2004) The physiologic basis of high-altitude diseases. Ann Intern Med 141:789–800
23. Schoene RB (2008) Illnesses at high altitude. Chest 134:402–416
24. Strohl KP (2008) Lessons in hypoxic adaptation from high-altitude populations. Sleep Breathing 12:115–121
25. Wilson MH, Newman S, Imray CH (2009) The cerebral effects of ascent to high altitudes. Lancet Neurol 8:175–191
26. Martin D, Windsor J (2008) From mountain to bedside: understanding the clinical relevance of human acclimatisation to high-altitude hypoxia. Postgrad Med J 84:622–627 quiz 626

27. Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J et al (1999) Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. Am J Hum Genet 65:1718–1724
28. Su B, Xiao C, Deka R, Seielstad MT, Kangwanpong D, Xiao J, Lu D, Underhill P, Cavalli-Sforza L, Chakraborty R et al (2000) Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. Hum Genet 107:582–590
29. Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC (1994) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93:189–199
30. Du R, Xiao C, Cavalli-Sforza LL (1997) Genetic distances between Chinese populations calculated on gene frequencies of 38 loci. Sci China C Life Sci 40:613–621
31. Buroker NE, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH (2012) AKT3, ANGPTL4, eNOS3, and VEGFA associations with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. Int J Hematol 96:200–213
32. Hackett PH (1992) The diagnoses accord with the Lake Louise scoring system. In: Sutton JR, Houston CS (eds) Hypoxia and mountainsickness. Pergamon, New York, pp 327–330
33. Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: a software for population genetics data analysis, 2.000 Edition. Geneva
34. Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390
35. Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nat Rev Genet 10:639–650
36. Yoo J, Seo B, Kim Y (2005) SNPAnalyzer: a web-based integrated workbench for single-nucleotide polymorphism analysis. Nucleic Acids Res 33:W483–W488
37. Ding K, Zhou K, He F, Shen Y (2003) LDA—a java-based linkage disequilibrium analyzer. Bioinformatics 19:2147–2148
38. Weir BS (1990) Genetic data analysis: methods for discrete population genetic data. Sinauer, Sunderland
39. Bryne JC, Valen E, Tang MH, Marstrand T, Winther O, da Piedade I, Krogh A, Lenhard B, Sandelin A (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic Acids Res 36:D102–D106
40. Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res 32:D91–D94
41. Sandelin A, Wasserman WW, Lenhard B (2004) ConSite: web-based prediction of regulatory elements using cross-species comparison. Nucleic Acids Res 32:W249–W252
42. Buroker NE, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH (2012) AKT3, ANGPTL4, eNOS3, and VEGFA associations with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. Int J Hematol 96:200–213
43. Brogan IJ, Khan N, Isaac K, Hutchinson JA, Pravica V, Hutchinson IV (1999) Novel polymorphisms in the promoter and 5′ UTR regions of the human vascular endothelial growth factor gene. Hum Immunol 60:1245–1249
44. Monge CC, Whittembury J (1976) Chronic mountain sickness. Johns Hopkins Med J 139(Suppl):87–89
45. Sime F, Monge C, Whittembury J (1975) Age as a cause of chronic mountain sickness (Monge's disease). Int J Biometeorol 19:93–98
46. Lambrechts D, Storkebaum E, Morimoto M, Del-Favero J, Desmet F, Marklund SL, Wyns S, Thijs V, Andersson J, van Marion I et al (2003) VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. Nat Genet 34:383–394
47. Prior SJ, Hagberg JM, Paton CM, Douglass LW, Brown MD, McLenithan JC, Roth SM (2006) DNA sequence variation in the promoter region of the VEGF gene impacts VEGF gene expression and maximal oxygen consumption. Am J Physiol Heart Circ Physiol 290:H1848–H1855
48. Stevens A, Soden J, Brenchley PE, Ralph S, Ray DW (2003) Haplotype analysis of the polymorphic human vascular endothelial growth factor gene promoter. Cancer Res 63:812–816
49. Aggarwal S, Negi S, Jha P, Singh PK, Stobdan T, Pasha MA, Ghosh S, Agrawal A, Prasher B, Mukerji M (2010) EGLN1 involvement in high-altitude adaptation revealed through genetic analysis of extreme constitution types defined in Ayurveda. Proc Natl Acad Sci USA 107:18961–18966
50. Forsythe JA, Jiang BH, Iyer NV, Agani F, Leung SW, Koos RD, Semenza GL (1996) Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. Mol Cell Biol 16:4604–4613
51. Tang M, Chen B, Lin T, Li Z, Pardo C, Pampo C, Chen J, Lien CL, Wu L, Ai L et al (2011) Restraint of angiogenesis by zinc finger transcription factor CTCF-dependent chromatin insulation. Proc Natl Acad Sci USA 108:15231–15236
52. Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. Cell 137:1194–1211
53. Buroker NE, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Ge M, Fan LP, Zhu WZ, Portman MA et al (2010) Genetic associations with mountain sickness in Han and Tibetan residents at the Qinghai-Tibetan Plateau. Clin Chim Acta 411:1466–1473
54. Buroker NE, Ning XH, Zhou ZN, Li K, Cen WJ, Wu XF, Zhu WZ, Scott CR, Chen SH (2012) EPAS1 and EGLN1 associations with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. Blood Cells Mol Dis 49:67–73