EGLN1 involvement in high-altitude adaptation revealed through genetic analysis of extreme constitution types defined in Ayurveda

Shilpi Aggarwal[a], Sapna Negi[a], Pankaj Jha[a], Prashant K. Singh[a], Tsering Stobdan[a], M. A. Qadar Pasha[a], Saurabh Ghosh[b], Anurag Agrawal[a], Indian Genome Variation Consortium[a], Bhavana Prasher[c,1], and Mitali Mukerji[a,1]

[a]Genomics and Molecular Medicine, Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research (CSIR), New Delhi 110007, India; [b]Human Genetics Unit, Indian Statistical Institute, Kolkata 700108, India; and [c]Planning and Performance Division, Council of Scientific and Industrial Research (CSIR), New Delhi 110001, India

Edited* by Charles R. Cantor, Sequenom, San Diego, CA, and approved September 20, 2010 (received for review May 6, 2010)

It is being realized that identification of subgroups within normal controls corresponding to contrasting disease susceptibility is likely to lead to more effective predictive marker discovery. We have previously used the Ayurvedic concept of *Prakriti*, which relates to phenotypic differences in normal individuals, including response to external environment as well as susceptibility to diseases, to explore molecular differences between three contrasting *Prakriti* types: *Vata, Pitta, and Kapha. EGLN1* was one among 251 differentially expressed genes between the *Prakriti* types. In the present study, we report a link between high-altitude adaptation and common variations rs479200 (C/T) and rs480902 (T/C) in the *EGLN1* gene. Furthermore, the TT genotype of rs479200, which was more frequent in *Kapha* types and correlated with higher expression of *EGLN1*, was associated with patients suffering from high-altitude pulmonary edema, whereas it was present at a significantly lower frequency in *Pitta* and nearly absent in natives of high altitude. Analysis of Human Genome Diversity Panel-Centre d'Etude du Polymorphisme Humain (HGDP-CEPH) and Indian Genome Variation Consortium panels showed that disparate genetic lineages at high altitudes share the same ancestral allele (T) of rs480902 that is overrepresented in *Pitta* and positively correlated with altitude globally ($P < 0.001$), including in India. Thus, *EGLN1* polymorphisms are associated with high-altitude adaptation, and a genotype rare in highlanders but overrepresented in a subgroup of normal lowlanders discernable by Ayurveda may confer increased risk for high-altitude pulmonary edema.

high-altitude pulmonary edema | *Prakriti* | Indian Genome Variation Database | phenotype | hypoxia

Ayurveda, an ancient system of Indian medicine documented and practiced since 1500 B.C., deals with interindividual variability for personalized and predictive medicine (1). This system of medicine phenotypically classifies individuals into seven broad constitution types termed *Prakriti*, among which *Vata* (V), *Pitta* (P), and *Kapha* (K), the most contrasting constitutions, exhibit readily recognizable phenotypes, respond differently to diet, drugs, and external environment as well as vary in predisposition to specific diseases (*SI Materials and Methods*). We have earlier shown differences between the three most contrasting *Prakriti* types of Indo-European origin in biochemical profiles and genome-wide expression and observed significant overrepresentation of hub and housekeeping genes within the differentially expressed genes (2).

We postulate that the genetic variations that underlie differential expression correlating with *Prakriti* phenotypes could provide leads for understanding adaptation to external environment and susceptibility to diseases. In this study, we observed significant genetic differences in five of the differentially expressed genes among the *Prakriti* types. We further studied *EGLN1*, a key oxygen sensor gene that negatively regulates the activity of hypoxia-inducible factor (HIF-1A). In physiological normoxic conditions, *EGLN1* hydroxylates the constitutively expressed HIF at two proline residues, leading to its polyubiquitination by the Von Hippel Lindau (VHL) E-3 ligase complex and subsequent degradation by the proteosomal machinery (3). Hypoxia leads to the inactivation of *EGLN1*, thereby increasing HIF that induces the expression of genes, which mediates adaptive responses through glycolytic enzymes, hemeoxygenase (cellular level), vascular endothelial growth factor (local), and erythropoietin (systemic level). Because oxygen homeostasis plays a key role in a large number of cellular, physiological, and systemic processes, we hypothesized that interindividual variations in *EGLN1* could contribute to differences in hypoxia responsiveness such as in high-altitude conditions. We analyzed the allele frequencies of two common variations (rs479200 and rs480902) in the *EGLN1* gene in populations from different altitudes represented in the Indian Genome Variation Consortium (IGVC) and HGDP-CEPH Human Genome Diversity panels (4, 5). We observed these variations not only to be linked to high-altitude adaptation but also to be associated with increased risk of developing high-altitude pulmonary edema (HAPE) in Indo-European sojourners. Thus, our study could establish a link between variations in *EGLN1* and high-altitude adaptation as well as susceptibility to HAPE, taking lead from expression and genetic differences in normal individuals identified from three contrasting constitution types described in Ayurveda.

Results

Distribution of Common Variations in Extreme Constitution Types. We studied the distribution of 141 tag SNPs encompassing 30 genes (Dataset S1) selected from the 251 differentially expressed genes between the V, P, and K from our earlier study in the same cohort (2). The details of recruitment and assessment of *Prakriti* types are provided in *SI Materials and Methods*. Ninety-two individuals who were not phenotyped for their constitution types but were from the same ethno-genetic background, namely Indo-European (IE), and large populations (IE-LP) were used as

Author contributions: M.M. designed research; S.A., S.N., P.J., P.K.S., and B.P. performed research; S.A., S.N., T.S., M.A.Q.P., I.G.V.C., B.P., and M.M. contributed new reagents/analytic tools; S.A., P.J., S.G., A.A., B.P., and M.M. analyzed data; and S.A., P.J., A.A., B.P., and M.M. wrote the paper.

Conflict of interest statement: S.A., M.A.Q.P., B.P., and M.M. are the inventors and have filed patent application no. 1336DEL2010 in India. There are no implications of this patent application on the publication of the manuscript, because the provisional patent application has already been filed. S.N., P.J., P.K.S., S.G., A.A., T.S., and The Indian Genome Variation Consortium have been acknowledged for contributing to the invention but do not fulfill the criteria of inventorship.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: mitali@igib.res.in or bhavana@csir.res.in.

The full list of authors participating in the Indian Genome Variation Consortium can be found at http://igvbrowser.igib.res.in/gbrowse/igvc.html.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1006108107/-/DCSupplemental.

heterogeneous phenotype controls (IE pool).The details of the populations are provided in *Materials and Methods* below. We observed that 14 SNPs from five genes (*AKT3*, *EGLN1*, *FAS*, *FBN2*, and *RAD51*) have significant allele frequency differences between the constitution types, even after correction for multiple testing with false-discovery rate (FDR) threshold set at 5% (Table 1). Although we had selected tag SNPs, majority of the SNPs in *AKT3* were in linkage disequilibrium (LD) and were different between P and K. Allele frequencies of rs480902 and rs479200 in *EGLN1* were significantly different between P and K. At the *FBN2* locus, rs1435514 showed significant allele frequency difference between P and K, and at the *RAD51* locus, K differed significantly from V at rs11858338, rs3092982, rs11855560, and rs12593359. At the *FAS* locus (rs2296603), P differed significantly from V. These differences (with the exception of *RAD51*) were very striking, because the alleles flip from being less frequent in one group to being more frequent in the other group. The observed genotypic differences also corroborated (except *FBN2*) with expression differences between the same *Prakriti* groups. Most importantly, after the constitution types were pooled, these contrasting allele frequencies were averaged out, and the pooled frequencies did not differ significantly from the IE background (Fig. 1). Further comparison of each constitution group with the IE pool revealed significant difference (FDR correction at 5%) between P and IE with respect to *AKT3* (rs2220276 and rs2291409) and between V and IE with respect to *RAD51* (rs11858338, rs12593359, rs11855560, and rs3092982) and *INSR* (rs8110533) (Table S1).

Genetic Variations in *EGLN1* Correlate with Altitude in IGVC and HGDP-CEPH Populations. Because *EGLN1* is a key oxygen sensor gene, we reasoned that variations in this gene, if meaningful, might also exhibit differences in allele frequencies across populations of different geographical locations, including those residing at high altitudes. An earlier study by the IGVC had analyzed the extent of genetic relatedness and homogeneity in 55 Indian populations from diverse linguistic and ethnic lineages of different geographical regions using various population genetic measures such as population differentiation by F_{ST}, genetic distance by Nei's D_A distance, system structure, and principal component analysis (PCA), and it identified five genetically close, near homogeneous clusters (4, 6). We studied patterns of distribution of *EGLN1* variations (rs480902 and rs479200) across a representative set of 24 Indian populations from the clusters described above. There was a significant difference ($P = 4.01 \times 10^{-7}$) in the allele frequencies of rs480902 and rs479200 between Tibeto-Burman (TB) populations (TB-N-IP1 and TB-N-SP1) residing at an altitude 3,500 m above sea level and other members (TB-NE-LP1, IE-N-IP2, and IE-NE-IP1) of the same genetic cluster but residing at low altitude (Fig. 2*A* and Table 2). At each of these SNPs, the alleles that were more frequent in the high-altitude populations (rs480902, T = 0.71; rs479200, C = 0.71) were also overrepresented in the P group (0.64 for both). The IE populations, which reside in Jammu and Kashmir, also had an overrepresentation of T allele of rs480902 (0.56) (Fig. 2*B*) and C allele of rs479200 (0.52) that was present in P constitution types. Given the diversity of Indian populations, these observations could also be a consequence of population stratification. We carried out a principal-component analysis of the V, P, and K cohort (Ayur) and the 24 Indian populations on a panel of 2,060 unlinked SNPs to investigate potential inflation of the odds ratio (OR) because of population stratification. ANOVA did not reveal any differences in the Ayur samples with IE populations of North India, and no V, P, or K individuals clustered with high-altitude populations (Fig. S1). The TB populations at high altitude also did not differ from members of the same linguistic background or genetic cluster as revealed by principal component analysis, but they differed significantly at the *EGLN1* locus with respect to the altitude (Table 2 and Fig. S1). The T allele of rs480902 associated with high altitude is found at high frequency (0.71) in the outgroup African population (OG-W-IP) residing in India. This population, called Siddi, are descendants from Bantu-speaking parts of East Africa (7). We also observed a conservation of LD between rs479200 and rs480902 across all of the populations (Fig. S2).

We further analyzed the *EGLN1* gene in the HGDP-CEPH Human Genome Diversity Panel that has sampled populations from various geographical locations all around the world (5). We determined the altitudes from Google earth based on the reported latitude and longitude of each population in the HGDP-CEPH panel (Tables S2 and S3). From a genome-wide study on the Illumina 650K array platform (8), the genotype for SNPs of *EGLN1* were retrieved for all of the HGDP-CEPH populations (http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/). We observed a significant correlation ($P < 0.01$) of allele frequencies of four SNPs (rs973252, rs480902, rs2808611, and rs2808614) with altitude, irrespective of genetic relatedness between the populations (Figs. 2*C* and 3*A* and Table 3). However, there were two populations, Burusho and Kalash, that, although they reside at very high altitudes, had underrepresentation of the *EGLN1* alleles that were associated with other high-altitude populations. These SNPs that were positively correlated with altitude span a region of ~29 kb in the first intron of the *EGLN1* gene and map to the

Table 1. SNPs that show significant difference between the constitution types after FDR correction for multiple testing set at a threshold of significance ($P < 0.05$)

Gene	SNP	Variation	Expression differences	Comparison allele frequency (1 vs. 2)	Allele	Allele frequency 1	Allele frequency 2	*P* value
AKT3	rs2345994	C/T	K+	PvsK	C	0.28	0.58	9.21E-04
AKT3	rs6672195	C/T	K+	PvsK	C	0.28	0.58	9.21E-04
AKT3	rs4590656	T/C	K+	PvsK	T	0.28	0.58	9.21E-04
AKT3	rs1973284	G/A	K+	PvsK	G	0.28	0.61	2.45E-04
AKT3	rs2220276	T/A	K+	PvsK	T	0.26	0.61	1.10E-04
AKT3	rs2291409	A/G	K+	PvsK	A	0.26	0.61	1.10E-04
EGLN1	rs480902	T/C	P-	PvsK	T	0.64	0.31	4.55E-04
EGLN1	rs479200	C/T	P-	PvsK	T	0.36	0.71	2.17E-04
FAS	rs2296603	T/C	P+	VvsP	T	0.35	0.64	1.41E-03
FBN2	rs1435514	T/C	V+	PvsK	C	0.76	0.45	7.66E-04
RAD51	rs11855560	T/C	V+	KvsV	T	0.32	0.08	7.77E-04
RAD51	rs11858338	G/A	V+	KvsV	A	0.61	0.87	5.68E-04
RAD51	rs12593359	T/G	V+	KvsV	T	0.39	0.13	5.68E-04
RAD51	rs3092982	C/G	V+	KvsV	C	0.39	0.11	2.24E-04

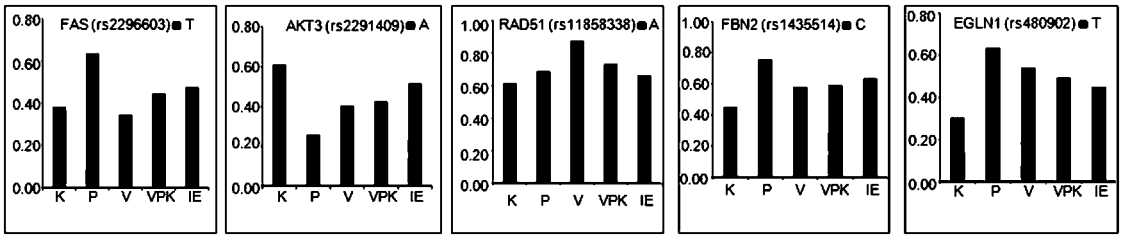

Fig. 1. Representation of allele frequencies of common variations among extreme constitution types. A representative set of SNPs that shows significant difference between the constitution types *Kapha* (K), *Pitta* (P), *Vata* (V), and differences from the V, P, and K/IE pool are depicted. The gene and SNP with the alleles are given in each panel. IE represents individuals with heterogeneous phenotypes from Indo-European populations, and V, P, and K represent individuals of different constitution types pooled into a single group.

same region that revealed differences in the Indian population with respect to altitude, and they also differed between the constitution types (Fig. S3). The T allele of the SNP rs480902 (also included in the 650K Illumina array) was highly correlated with altitude (Kendall's rank correlation: $P < 0.001$; $\tau = 0.2903123$). This allele, although associated with high altitude, is found in all HGDP-CEPH populations, and it had higher frequencies predominantly in sub-Saharan Africa (Fig. 3*B*). In majority of the European populations, the C allele of rs480902 was more frequent (Fig. 3*B*).

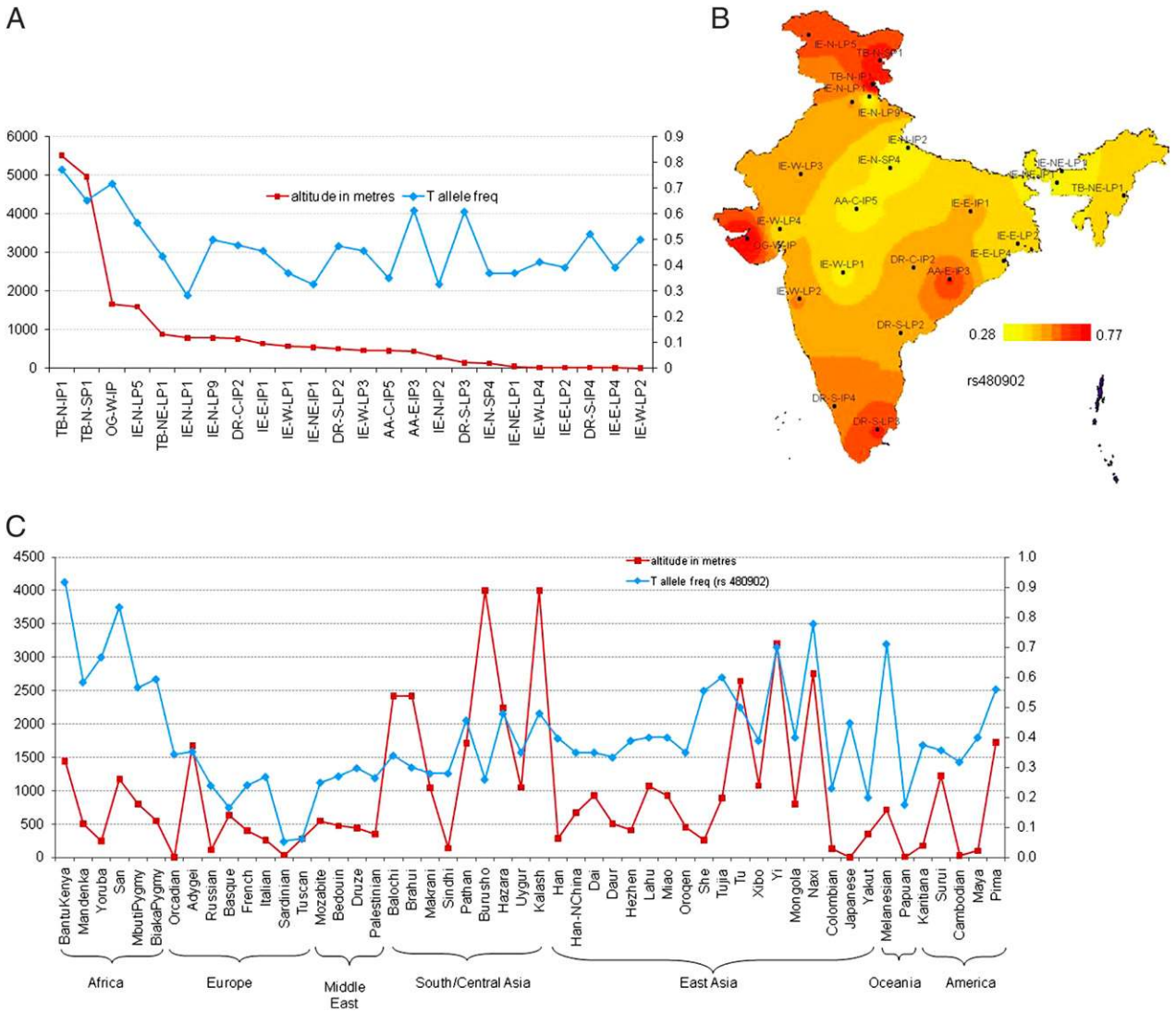

Fig. 2. Distribution of T allele frequency of rs480902 in diverse IGVC and HGDP-CEPH populations from different altitudes. (*A*) Frequency in the 24 IGV populations and their altitude. (*B*) Spatial frequency map of rs480902 in IGV populations. The color gradient below the map depicts the range of observed frequency of the T allele from minimum to maximum. (*C*) Frequency distribution in the HGDP-CEPH panel of 52 populations along with their altitudes. Diverse continental populations residing at high altitudes selectively retain the ancestral T allele.

Table 2. Comparison of allele frequency of *EGLN1* SNPs between Indian populations of the same genetic cluster residing at high and low altitudes

SNP	Allele	Allele frequency		*P* value: high vs. low
		High altitude	Low altitude	
rs480902	T	0.71	0.36	4.01E-07
rs479200	C	0.71	0.36	4.01E-07

Association of Common Variations in the *EGLN1* Gene with HAPE. At the genetic level, K differed significantly from P and V with respect to two SNPs, rs480902 and rs479200, that span an ~12-kb region in the first intron of the *EGLN1* gene. Compared with the TC and CC genotypes at rs479200, the TT genotype that was overrepresented in K also had significantly higher expression of *EGLN1* (one-tailed *t* test; *P* value = 0.017) (Fig. 4*A*). Because higher expression of *EGLN1* is inversely correlated to HIF activity, we hypothesized that individuals with genotypes associated with high *EGLN1* expression may not be able to perform well under hypoxic conditions. To test this hypothesis, we studied a cohort of IE sojourners to high altitudes who suffered from HAPE as well as natives of that high-altitude region. Interestingly, the TT genotype of rs479200 that was associated with higher expression of *EGLN1* in the K constitution had a significantly higher frequency (0.44) in HAPE patients compared with natives (0.05) of the high altitude (Fig. 4*B*). In addition, the frequency of C allele of rs480902 and the T allele of rs479200 (0.63 and 0.64, respectively) in HAPE patients was similar to K type (0.69 and 0.71). The alleles associated with the K constitution were significantly underrepresented in P constitution (0.36 and 0.36) as well as the natives (0.28 and 0.21) of high altitude (Fig. 4*C* and Table S4). After V, P, and K were pooled, both the SNPs assumed an allele frequency similar to IE population, and their frequency difference from HAPE patients was also not apparent.

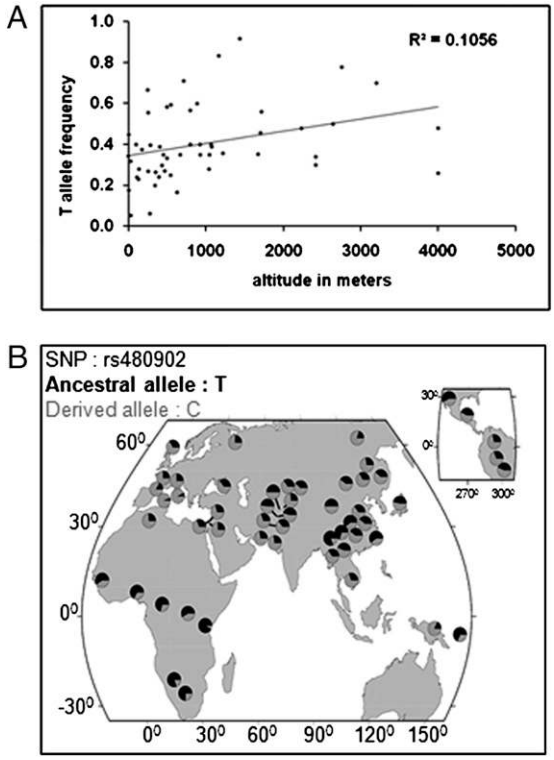

Fig. 3. Allele frequency distribution of rs480902 in HGDP-CEPH populations. (*A*) Correlation of T allele frequency of rs480902 with increasing altitude (R^2 = 0.1056) in HGDP-CEPH populations. (*B*) Spatial frequency map of rs480902 in the HGDP-CEPH populations retrieved from the HGDP selection browser. Frequencies of the ancestral T allele that is predominantly present in populations residing at high altitude and the derived C allele are represented by dark and light shades, respectively.

Discussion

Interindividual differences in susceptibility to diseases and response to environment and drugs are, to a large extent determined by genomic variations. A large fraction of these variations could be a consequence of population history, drift, or adaptation to spatially varying selective pressures such as diet and climate. However, given the large amount of variations in an individual's genome, linking these to a phenotype is an extremely challenging task. According to Ayurveda, response to external environment (diet, weather, lifestyle, stress, and drugs), susceptibility, and progression of disease are largely determined by an individual's basic constitution (*Prakriti*), which can be phenotypically analyzed (1, 2). Therefore, classifying normal individuals based on Ayurveda constitution types may also allow us to identify meaningful phenotype to genotype links. Our earlier observations revealed significant differences in biochemical parameters and gene expression between the three contrasting *Prakriti* types of IE origin identified using phenotyping methods of Ayurveda (2). Here, we show significant differences in allele frequencies of common variations in five genes (*FAS*, *AKT3*, *FBN2, EGLN1*, and *RAD51)* between the *Prakriti* groups in the same study population described earlier. After the V, P, and K samples were pooled, these SNPs assumed a frequency similar to the background population. We hypothesized that these variations, which are linked to *Prakriti* groups that are differently predisposed to diseases, may lead us to identification of predictive markers for differential responsiveness to disease and environment.

As a proof of concept, we studied *EGLN1* because it plays a key role in oxygen homeostasis and is also believed to be a target of many pharmacological interventions that aim to stabilize HIF or lower HIF activity (9–11). Although variations in genes of pathways related to hypoxia, such as HIF-1, endothelial function, and vascular remodeling, have been studied (12–17), none of the studies so far have reported polymorphisms linked to *EGLN1* in high-altitude adaptation. We analyzed the allele frequencies of *EGLN1* polymorphisms that differed between the constitution types in populations residing at different altitudes as well as in

Table 3. Correlation of frequency of *EGLN1* SNPs with altitude in the HGDP-CEPH Human Genome Diversity Panel

SNP	Allele	Correlation value (τ)	*P* value
rs973252	A	0.2305814	0.008664
rs480902	T	0.2903123	0.001337
rs519504	A	−0.05634704	0.2853
rs545937	T	−0.05466884	0.2908
rs2808611	G	0.2271038	0.009633
rs2808614	G	0.2184982	0.01191
rs7542797	C	−0.1412878	0.07377
rs1622146	C	−0.06183828	0.2631

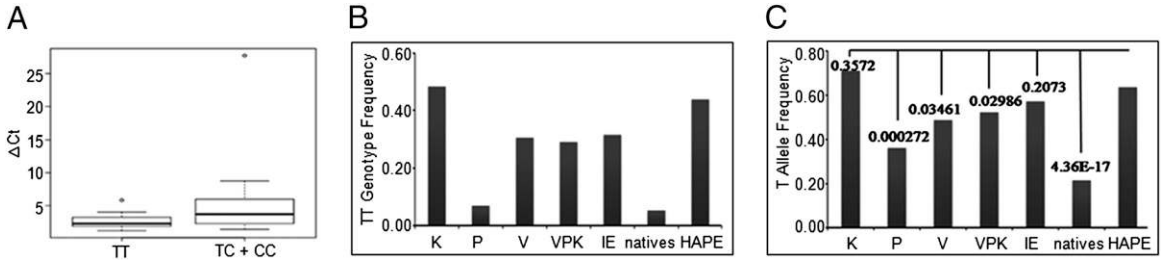

Fig. 4. Correlation between *EGLN1* genotypes of rs479200 and expression and the association of TT genotype and the T allele with HAPE. (*A*) Box plot representing ΔCT values of gene expression of *EGLN1* by RT-PCR in TT, TC, and CC genotypes of rs479200 in Ayurveda samples. (*B*) Frequency of TT genotype of rs479200 in different constitution types (K, P, and V), VPK, IE, natives of high altitude, and patients of HAPE. (*C*) Frequency of T allele of rs479200 in different constitution types (K, P, and V), VPK, IE, natives of high altitude, and patients of HAPE. Fisher's exact test was performed for association analysis of *EGLN1* SNP rs479200 between different controls and HAPE. The numbers over each of the bars represent the *P* values of comparison of each group with HAPE.

subjects who develop HAPE, a condition that normally occurs in unacclimated sojourners at altitudes above 2,500 m and accounts for most of the deaths caused by altitude sickness (18). Analysis of 24 diverse Indian populations revealed that TB populations residing at high altitudes had a significantly higher frequency of T allele of rs480902 and C allele of rs479200 (overrepresented in P constitution types) compared with populations that resided at low altitude but were from the same genetic cluster. Our observation in Indian population was further corroborated with the analysis of the HGDP-CEPH panel, which showed that disparate genetic lineages at high altitude share the same ancestral allele (T) of rs480902. This indicates a selection for retention of an ancestral physiological adaptation at high altitudes, except for in populations Kalash and Burusho, which seem to have acquired adaptation to high altitude through a different mechanism. These populations have inhabited the high altitude much more recently compared with the Tibetan and Andean Highlanders (19). It would be interesting to further explore this finding.

The role of *EGLN1* in high-altitude adaptation is further substantiated by the presence of higher frequencies of T allele and TT genotypes of rs479200 in IE sojourners who develop HAPE. The TT genotype, corresponding to higher gene expression of *EGLN1*, is overrepresented in K and rare in natives and P, which raises the possibility that K may have a higher risk of HAPE and P *Prakriti* could be more protected. The comparison of sojourners who develop HAPE with healthy individuals of same genetic background (IE pool) did not reveal significant differences. This could be because IE pool is comprised of heterogeneous constitution types, and in the absence of phenotypic stratification, the effect of these variations are masked. Although a cohort of companions of IE sojourners that did not develop HAPE on multiple ascents would have been of much interest, such a cohort was not accessible because of highly sensitive military areas.

Interestingly, Ayurveda assigns *Prakriti* not only to humans but also to environment and food, and it makes specific mention of adaptation as well as dietary and lifestyle recommendations based on one's *Prakriti* for achieving healthy balance. An interpretation of our results that P constitution is more protected at high altitudes would be consistent with the Ayurvedic school of thought that considers mountains mainly as K and V dominant regions (*SI Materials and Methods*), and therefore, there would be higher prevalence of K and V diseases.

EGLN1 gene, owing to its important function as an oxygen sensor, is relevant to the human hypoxic response, both at high altitude in hypoxic conditions or in cellular hypoxia. Furthermore, EGLN1 is being considered as an important pharmacological target. Therefore, it is important to study *EGLN1* variations in diseases and drug response, where cellular hypoxia is implicated in pathogenesis. The SNPs that are associated with high-altitude adaptation, both in the Indian study as well as in the global populations, encompass the first intron of the EGLN1 gene. This region is highly conserved and harbors a segmental duplication as well as conserved DNA regulatory elements (Fig. S3). Functional characterization of this region would provide insights into mechanisms of regulation of the *EGLN1* gene.

Conclusion

Our study shows that expression and genetic analysis of healthy individuals phenotyped using the principles of Ayurveda could uncover genetic variations that are associated with adaptation to external environment and susceptibility to diseases. We show, through genetic analysis, that two contrasting constitutions within nondiseased normals derived from the same genetic background differ both at the expression and genetic level with respect to the *EGLN1* gene, and these differences are linked to high-altitude adaptation and susceptibility to HAPE. Our work further suggests that variations in the hypoxia response pathway are common in most of the world population and could attain different allele frequencies as a consequence of positive selection.

The involvement of the *EGLN1* gene in high-altitude adaptation in Tibetan highlanders has been shown by two independent groups using whole-genome approaches while our manuscript was under review (20–22).

Materials and Methods

Study Subjects. The study was carried out in four different cohorts of samples that are described in detail in *SI Materials and Methods*. Briefly, the samples comprised of (*i*) 96 individuals of extreme constitution types V (39), P (29), and K (28) identified from an initial phenotyping of 850 volunteers on the basis of Ayurveda methods and recruited in our earlier study (2) and (*ii*) 552 samples from 24 diverse Indian populations from the existing panel of IGVC (6). These include 92 heterogeneous phenotype controls (IE pool) from IE North Indian large populations (size > 10 million). Our earlier study on Indian Genome Variation that had sampled diverse populations from different linguistic, geographical, and ethnic background revealed five near homogeneous genetic clusters, where IE large population from Northern India was one of the clusters (4, 6). (*iii*) Additionally, the samples included 96 unrelated HAPE patients from IE background and (*iv*) 96 samples of unrelated natives of Leh recruited through Sonam Norboo Memorial (SNM) Hospital, Leh (altitude, ~3,500 m), Jammu, and Kashmir, India (17).

Genetic and Expression Analysis. A total of 158 SNPs from 30 genes that exhibited expression differences in our earlier study (*SI Materials and Methods* and Dataset S1) and 2,060 SNPS that were used for population stratification were genotyped in V, P, and K samples as well as the IGVC panel (http://igvbrowser.igib.res.in) using Illumina Bead Array platform (*SI Materials and Methods*). Genotyping of rs480902 and rs479200 on HAPE samples and natives was carried out using single base primer extension assay (SNaPSHOT ddNTP Primer extension kit; Applied Biosystems) on an ABI Prism 3100 Genetic Analyzer. The genotype data on *EGLN1* SNPs from 52 populations were retrieved from HGDP selection browser (http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/). Relative expression of *EGLN1* between the constitution types was measured by real-time quantitative (TaqMan) PCR using two genes, *ASAH1* and *MAN1A1*, as internal control (details in *SI Materials and Methods*).

Statistical Analysis. We used Fisher's exact test for estimating allelic frequency differences and testing genotypic and allelic associations. Correction for multiple testing was done using the FDR method. EIGENSTRAT (23) was used for analysis of population stratification in the IGVC panel and Ayurveda samples. Gene expression normalization factor for each sample based on the geometric mean of two internal controls was performed, and differences in expression of *EGLN1* with respect to rs479200 genotypes were compared using a one-tailed *t* test. We used Kendall's rank correlation to study the relation of altitude with different SNPs of *EGLN1* in HGDP-CEPH populations.

Details of samples, SNP selection, experiments, and data analysis are described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Prof. Samir K. Brahmachari for conceiving the idea of integration of the ancient Indian system of personalized and predictive medicine with modern genomics and critical review, Drs. Vani Brahmachari, Ram Niwas Prasher, and Arijit Mukhopadhyay for critical review; Ankita and Amit Mandal for analysis support; and Drs. Ghulam Mohammad and Tashi Thinlas, Sonam Norboo Memorial (SNM) Hospital (Leh, Jammu, and Kashmir, India) for the collection of HAPE samples. Financial support from the Council of Scientific and Industrial Research Senior Research Fellowship (SRF) (to S.A.), Department of Science and Technology (DST) Grant B6.25 (to M.M.), and Council of Scientific and Industrial Research Grants CMM0016 and MLP3601 (to M.M.) is acknowledged.

1. Sharma PV (2000) *Caraka Samhita* (Chaukhamba Orientalia, Varanasi, India).
2. Prasher B, et al. (2008) Whole genome expression and biochemical correlates of extreme constitutional types defined in Ayurveda. *J Transl Med* 6:48.
3. Semenza GL (2009) Regulation of oxygen homeostasis by hypoxia-inducible factor 1. *Physiology (Bethesda)* 24:97–106.
4. Indian Genome Variation Consortium (2008) Genetic landscape of the people of India: A canvas for disease gene exploration. *J Genet* 87:3–20.
5. Cann HM, et al. (2002) A human genome diversity cell line panel. *Science* 296:261–262.
6. Indian Genome Variation Consortium (2005) The Indian Genome Variation database (IGVdb): A project overview. *Hum Genet* 118:1–11.
7. Singh KS (2003) *People of India* (Gujarat Manohar, Prakashan, India), pp 1295–1298.
8. Li JZ, et al. (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319:1100–1104.
9. Ikeda E (2005) Cellular response to tissue hypoxia and its involvement in disease progression. *Pathol Int* 55:603–610.
10. Loinard C, et al. (2009) Inhibition of prolyl hydroxylase domain proteins promotes therapeutic revascularization. *Circulation* 120:50–59.
11. Luukkaa M, et al. (2009) Expression of the cellular oxygen sensor PHD2 (EGLN-1) predicts radiation sensitivity in squamous cell cancer of the head and neck. *Int J Radiat Biol* 85:900–908.
12. Mortimer H, Patel S, Peacock AJ (2004) The genetic basis of high-altitude pulmonary oedema. *Pharmacol Ther* 101:183–192.
13. Qadar Pasha MA, et al. (2001) Angiotensin converting enzyme insertion allele in relation to high altitude adaptation. *Ann Hum Genet* 65:531–536.
14. Morrell NW, Sarybaev AS, Alikhan A, Mirrakhimov MM, Aldashev AA (1999) ACE genotype and risk of high altitude pulmonary hypertension in Kyrghyz highlanders. *Lancet* 353:814.
15. Smith TG, Robbins PA, Ratcliffe PJ (2008) The human side of hypoxia-inducible factor. *Br J Haematol* 141:325–334.
16. Liu K, Sun X, Wang S, Hu B (2007) Association between polymorphisms of HIF-1alpha C1772T and G1790A and hypoxic acclimation in high altitude in Tibetans. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* 24:654–658.
17. Ahsan A, et al. (2004) eNOS allelic variants at the same locus associate with HAPE and adaptation. *Thorax* 59:1000–1002.
18. Peacock AJ (1995) High altitude pulmonary oedema: Who gets it and why? *Eur Respir J* 8:1819–1821.
19. Mehdi SQ, et al. (1999) *Genomic Diversity: Application in Human Population Genetics*, eds Papiha SS, Deka R, Chakraborty R (Plenum, New York), pp 83–90.
20. Yi X, et al. (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329:75–78.
21. Simonson TS, et al. (2010) Genetic evidence for high-altitude adaptation in Tibet. *Science* 329:72–75.
22. Storz JF (2010) Evolution. Genes for high altitudes. *Science* 329:40–41.
23. Price AL, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909.