Archaic introgression contributed to shape the adaptive modulation of angiogenesis and cardiovascular traits in human high-altitude populations from the Himalayas

Giulia Ferraretti[1*], Paolo Abondio[2*], Marta Alberti[1], Agnese Dezi[3], Phurba T. Sherpa[4], Paolo Cocco[5], Massimiliano Tiriticco[5], Marco di Marcello[5], Guido Alberto Gnecchi-Ruscone[6], Luca Natali[5,7], Angela Corcelli[8], Giorgio Marinelli[5], Davide Peluzzi[5], Stefania Sarno[1], Marco Sazzini[1,9#]

[1]Laboratory of Molecular Anthropology & Centre for Genome Biology, Dept. of Biological, Geological and Environmental Sciences, University of Bologna, Italy

[2]Dept. of Cultural Heritage, Ravenna Campus, University of Bologna, Italy

[3]Dept. of Emergency and Organ Transplantation, University of Bari Aldo Moro, Italy

[4]Mount Everest Summiters Club, Rolwaling, Dolakha, Nepal

[5]Explora Nunaat International, Montorio al Vomano, Teramo, Italy

[6]Dept. of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[7]Italian Institute of Human Paleontology, Rome, Italy

[8]Dept. of Basic Medical Science, Neuroscience and Sense Organs, University of Bari Aldo Moro, Italy

[9]Interdepartmental Centre Alma Mater Research Institute on Global Changes and Climate Change, University of Bologna, Italy

*These authors equally contributed

[#]For correspondence: marco.sazzini2@unibo.it (MS)

Abstract It is well established that several *Homo Sapiens* populations experienced admixture with extinct human species during their evolutionary history. Sometimes, such a gene flow could have played a role in modulating their capability to cope with a variety of selective pressures, thus resulting in archaic adaptive introgression events. A paradigmatic example of this evolutionary mechanism is offered by the *EPAS1* gene, whose most frequent haplotype in Himalayan highlanders was proved to reduce their susceptibility to chronic mountain sickness and to be introduced in the gene pool of their ancestors by admixture with Denisovans. In this study, we aimed at further expanding the investigation of the impact of archaic introgression on more complex adaptive responses to hypobaric hypoxia evolved by populations of Tibetan/Sherpa ancestry, which have been plausibly mediated by soft selective sweeps and/or polygenic adaptations rather than by hard selective sweeps. For this purpose, we used a combination of composite-likelihood and gene network-based methods to detect adaptive loci in introgressed chromosomal segments from Tibetan whole genome sequence data and to shortlist those presenting Denisovan-like derived alleles

that participate to the same functional pathways and are absent in populations of African ancestry, which are supposed to do not have experienced Denisovan admixture. According to this approach, we identified multiple genes putatively involved in archaic introgression events and that, especially as regards *TBC1D1*, *RASGRF2*, *PRKAG2*, and *KRAS*, have plausibly contributed to shape the adaptive modulation of angiogenesis and of certain cardiovascular traits in high-altitude Himalayan peoples. These findings provided unprecedented evidence about the complexity of the adaptive phenotype evolved by these human groups to cope with challenges imposed by hypobaric hypoxia, offering new insights into the tangled interplay of genetic determinants that mediates the physiological adjustments crucial for human adaptation to the high-altitude environment.

Introduction

The scientific community currently agrees that the *Homo sapiens* species experienced admixture with extinct Hominins since traces of such inbreeding events are still detectable in the genomes of modern humans (Gouy et al., 2020). In fact, people belonging to non-African populations show 1–2% of Neanderthal ancestry (Green et al., 2010; Prüfer et al., 2014), while Melanesians and East Asians present 3% and 0.2% of Denisovan ancestry, respectively (Reich et al., 2010; Meyer et al., 2013; Prüfer et al., 2014; Racimo et al., 2017). Despite evidence supporting selection against introgressed alleles has been collected (Simonti et al., 2016; Racimo et al., 2017; McArthur et al., 2021), some of the genomic segments showing signatures ascribable to archaic introgression were also proved to have been targeted by natural selection in modern human populations, thus providing examples for the occurrence of Adaptive Introgression (AI) events (Racimo et al., 2017).

So far, several studies have indeed identified introgressed archaic alleles at high frequency in human genes involved in metabolism or in the response to environmental conditions, such as temperature, sunlight, and altitude (Prüfer et al., 2014; Benjamin et al., 2014; Sankararaman et al., 2014; Huerta-Sánchez et al., 2014; Gittelman et al., 2016; Racimo et al., 2017; Enard et al., 2018; Danneman et al., 2018). Moreover, some genes that play a role in immune responses to pathogens are found to be characterized by a similar pattern of variability (Laurent et al., 2011; Enard et al., 2018) and certain Neanderthal alleles have been shown to be associated with down-regulation of gene expression in brain and testes (McCoy et al., 2017; Racimo et al., 2017; Dannemann et al., 2018). These works collectively attest how genetic variants introduced in the human gene pool by admixture with archaic species can significantly impact our biology by possibly comporting modifications in the modulation of several functional pathways. In particular, the high frequency of some archaic alleles in protein-coding and/or regulatory genomic regions suggests a possible adaptive role for Neanderthal and/or Denisovan variants, pointing to a further evolutionary mechanism having potentially contributed to the processes of human biological adaptation to different environmental and cultural settings. By introducing new alleles in the gene pool of a given population, admixture in fact provides a very rapid opportunity for natural selection to act on it (Huerta-Sánchez et al., 2014; Jeong et al., 2014; Racimo et al.,

2015; Hamid et al., 2021) and this is supposed to have likely occurred during the evolutionary history of *H. Sapiens*, particularly after the last Out of Africa migration in the late Pleistocene (Israel et al., 2018; Vahdati et al., 2022). According to this view, gene flow from extinct Hominin species could have facilitated the adaptation of *H. Sapiens* populations to peculiar Eurasian environments.

For instance, a Denisovan origin of the adaptive *EPAS1* haplotype, which confers reduced susceptibility to chronic mountain sickness to Tibetan and Sherpa highlanders (Beall, 2007; Bigham et al., 2010; Yi et al., 2010; Peng et al., 2011; Xu et al., 2011) is well established (Huerta-Sánchez et al., 2014; Zhang et al., 2021). However, the hard selective sweep experienced in high-altitude Himalayan populations by the *EPAS1* introgressed haplotype has been demonstrated to account only for an indirect aspect of their adaptive phenotype, which does not explain most of the physiological adjustments they evolved to cope with hypobaric hypoxia (Gnecchi-Ruscone et al., 2018). Therefore, how far gene flow between Denisovans and the ancestors of Tibetan/Sherpa peoples facilitated the evolution of other key adaptive traits of these populations remains to be elucidated.

To fill this gap, and to overcome the main limitation of most approaches currently used to test for AI (i.e., inferring archaic introgression and the action of natural selection separately by means of different algorithms, which increases the risk of obtaining biased results due to confounding variables), we assembled a dataset of whole genome sequences (WGS) from 27 individuals of Tibetan ancestry living at high altitude (Cho et al., 2017; Jeong et al., 2018) and we analysed it using a composite-likelihood method specifically developed to detect AI events at once (Setter et al., 2020). Notably, this method was designed to recognize AI mediated by subtle selective events (as those involved in polygenic adaptation) and/or soft selective sweeps, which represent the evolutionary mechanisms that are supposed to have played a more relevant role than hard selective sweeps during the adaptive history of human groups characterized by particularly small effective population size, such as Tibetans and Sherpa (Gnecchi-Ruscone et al., 2018). Coupled with validation of the identified putative adaptive introgressed loci through i) the assessment of the composition of gene networks made up of functionally related DNA segments presenting archaic derived alleles that are absent in human groups which are supposed to do not have experienced Denisovan admixture, such as African ones, ii) the confirmation that natural selection targeted these genomic regions in populations of Tibetan ancestry, as well as iii) the quantification of genetic distance between modern and archaic haplotypes, such an approach provided new evidence about the biological functions that have mediated high-altitude adaptation in Himalayan populations and that have been favourably shaped by admixture of their ancestors with Denisovans.

Results

Spatial distribution of genomic variation and ancestry components of Tibetan samples

After quality control (QC) filtering of the available WGS data, we obtained a dataset made up of 27 individuals of Tibetan ancestry characterized for 6,921,628 single nucleotide variants (SNVs). To assess whether this dataset represents a reliable proxy for the genomic variation observable in high-altitude

Himalayan populations, we merged it with genome-wide genotyping data for 1,086 individuals of East-Asian ancestry belonging to both low- and high-altitude groups (Gnecchi-Ruscone et al., 2017; Landini et al., 2021). We thus obtained an extended dataset including 231,947 SNVs (Supplement table 1), which was used to perform population structure analyses.

Results from ADMIXTURE and Principal Components Analysis (PCA) were found to be concordant with those described in previous studies (Jeong et al., 2014; Gnecchi-Ruscone et al., 2017; Gnecchi-Ruscone et al., 2018; Yang et al., 2021). According to the ADMIXTURE model showing the best predictive accuracy (K = 7) (Figure 1 – figure supplement 1), the examined WGS exhibited a predominant genetic component that was appreciably represented also in other populations speaking Tibeto-Burman languages, such as Tu, Yizu, Naxi, Lahu, and Sherpa (Figure 1A and Figure 1 – figure supplement 2). Such a component reached an average proportion of around 78% in individuals of Tibetan ancestry from Nepal included in the extended dataset, as well as of more than 80% in the subjects under investigation, who live in the Nepalese regions of Mustang and Ghorka (Figure 1A and 1B). This suggests that after their relatively recent migration in Nepalese high-altitude valleys, these communities might have experienced a higher degree of isolation and genetic drift with respect to populations that are still settled on the Tibetan Plateau, in which the same ancestry fraction did not exceed 64% (Figure 1A and 1B). Nevertheless, the overall ADMIXTURE profile of the considered WGS appeared to be quite comparable to those inferred according to genome-wide genotyping data for other Tibetan populations (Figure 1A and 1B, Figure 1 - figure supplement 2). Similarly, PCA pointed to the expected divergence of Tibetan and Sherpa high-altitude groups from the cline of genomic variation of East-Asian lowland populations (Abdulla et al., 2009; Jeong et al., 2014; Gnecchi-Ruscone et al., 2017; Zhang et al., 2017; Wang et al., 2022) (Figure 1C). Remarkably, the WGS under investigation clustered within the bulk of genome-wide data generated for other groups from the Tibetan Plateau, thus supporting their representativeness as concerns the overall genetic background of high-altitude Himalayan populations.

Detecting putative adaptive introgression signatures in Tibetan genomes

To identify genomic regions showing signatures putatively ascribable to AI events, we scanned Tibetan WGS with the *VolcanoFinder* algorithm and we computed the composite Likelihood Ratio (LR) and -logα statistics for each polymorphic site (Setter et al., 2020). We then considered the most significant results by focusing on loci showing LR values falling in the positive tail (i.e., top 5%) of the obtained distribution (see Materials and methods, Supplement table 2).

According to such an approach, we were first able to recapitulate the AI event previously described for the *EPAS1* gene (Figure 2 – figure supplement 1A) (Huerta-Sánchez et al., 2014; Hu et al., 2017; Zhang et al., 2021). In fact, this chromosomal interval was found to be characterized by a remarkable number of variants (N=19) showing significant LR scores, as well as by high overall values of -logα (Figure 2 - figure supplement 1A), suggesting respectively the plausible archaic origin of many alleles at this gene and an appreciable action of natural selection on it. Five of these significant SNVs have been already described as

Denisovan-like derived alleles at outstanding frequency (i.e., ranging between 0.96 and 1) in Tibetans but not in other modern human populations (Supplement table 3) (Zhang et al., 2021). On the contrary, at the genomic region encompassing *EGLN1* (which we have considered as a negative control for AI, see Materials and methods) we detected high -logα values coupled with a low number of SNVs (N=3) showing significant LR scores, with only one being remarkably above the adopted significance threshold (Figure 2 – figure supplement 1B and Supplement table 4). Overall, these findings are concordant with evidence from literature that suggest adaptive evolution of both *EPAS1* and *EGLN1* loci in high-altitude Himalayan populations (Yang et al., 2017; Liu et al., 2022), although only the former was proved to have been impacted by archaic introgression (Huerta-Sanchez et al., 2014; Hu et al., 2017; X. Zhang et al., 2021).

Moreover, we were able to confirm other introgression signatures previously inferred from WGS data for populations of Tibetan ancestry, such as those involving the *PRKCE* gene and the *MIRLET7BHG* long non-coding region, which are located in the overlapping upstream chromosomal intervals respectively of *EPAS1* and *PPARA* (Figure 2 – figure supplement 2A and 2B). In line with what reported for *EPAS1*, also *PPARA* has been already proposed to play a role in the modulation of high-altitude adaptation of Himalayan human groups (Simonson et al., 2010; Horscroft et al., 2017; Zhang et al., 2021). Interestingly, AI signatures identified by *VolcanoFinder* in the *MIRLET7BHG* locus extended also in the *PPARA* gene, as well in its downstream region (Figure 2 – figure supplement 2A), supporting the findings described by Hu et al. (2017). Finally, we observed patterns comparable to those at *EPAS1* and *PPARA* for 10 additional genomic regions that were differentially pointed out by previous studies as Tibetan and/or Han Chinese DNA segments potentially carrying introgressed Denisovan alleles (Hu et al. 2017, Browning et al. 2018, Zhang et al., 2021) (Figure 2A and 2B, Figure 2 – figure supplement 3A, Supplement table 5).

Validating genomic regions affected by archaic introgression

To validate signatures of archaic introgression at the candidate AI loci identified with *VolcanoFinder*, we relied on the approach described by Gouy & Excoffier (2020). In detail, we used the *Signet* algorithm to identify networks of genes participating to the same functional pathway and presenting archaic-like (i.e., Denisovan) derived alleles observable in the putative admixed group (i.e., Tibetans) but not in an outgroup of African ancestry (i.e., Yoruba, YRI), by assuming that only Eurasian *H. sapiens* populations experienced Denisovan admixture (see Materials and methods).

After having crosschecked results from the *VolcanoFinder* and *Signet* analyses, we identified six gene networks that turned out to be consistently significant across all the *Signet* runs performed and that included a total of 15 genes pointed by *VolcanoFinder* as candidate AI loci (Supplement table 6). Four of these loci composed the gene network overall ascribable to the *Pathways in Cancer* biological functions, which included also the *EPAS1* positive control for AI (Figure 3A). Most of the other genes supported by both the analyses were instead observed in significant networks belonging to the *Ras signalling* and *AMPK signalling* pathways (Supplement table 6).

Interestingly, gene networks belonging to the *Pathways in Cancer* and *Ras signalling pathway* appeared to be tightly related from a functional perspective because included oncogenes that promote the initiation and progression of tumour growth by stimulating cell proliferation and angiogenesis (Kranenburg et al., 2004). In particular, according to the KEGG database *KRAS* and *RASGRF2* genes from the *Ras signalling* network were found to contribute also to the *Pathways in Cancer* functions (Figure 3A), especially by interacting with the identified candidate introgressed loci *PLCB1*, *RASGRP2*, *DAPK1*, *MAPK1*, *FOS*, and *VEGFA* to modulate the *VEGF signalling pathway*, which is activated in hypoxic conditions and induces the transcription of genes that promote angiogenesis (Figure 3A) (Maxwell & Ratcliffe, 2002; Kranenburg et al., 2004).

Also, the *AMPK signalling pathway* is known to be activated in different cell types by stresses such as deprivation of oxygen and/or glucose, leading to the inhibition of energy-consuming biosynthetic pathways (e.g., protein and glycogen synthesis) and to the activation of ATP-producing catabolic pathways, such as fatty acid oxidation and glycolysis (Kanehisa & Goto, 2000; Chen et al., 2018; Dengler, 2020).

No significant gene networks involving the *EGLN1* genomic region considered as a negative control for AI were instead reconstructed with the *Signet* approach (Supplement table 6), suggesting that the very few variants at this locus that showed *VolcanoFinder* LR scores above the adopted significant threshold might represent false positive results (Figure 2 – figure supplement 1B and Supplement table 4).

Shortlisting introgressed genomic regions characterized by adaptive evolution

To further shortlist the most robust candidate genes involved in AI events, we applied the *LASSI* algorithm to phased Tibetan WGS data with the aim of searching for genomic signatures ascribable to the action of natural selection (Harris & DeGiorgio, 2020) (see Materials and methods).

This enabled us to confirm the strong selective events occurred at the *EPAS1* and *EGLN1* genes, as previously reported by multiple studies conducted on high-altitude Himalayan populations (Beall et al., 2010; Yi et al., 2010;Simonson et al., 2010; Horscroft et al., 2017; Zhang et al., 2021), as well as to corroborate adaptive evolution of some of the genes pointed out by both *VolcanoFinder* and *Signet* analyses. In fact, several chromosomal intervals associated to these loci presented values of the computed *T* statistic that fall within the top 5% of the related distribution (Figure 4C and 4D, Figure 4 – figure supplement 1, 2C, 2D, 3C and 3D).

More in detail, in addition to *EPAS1*, genomic windows associated to the *DAPK1*, *GNG7*, *AK5*, *TBC1D1*, *PLCB1*, *RASGRF2*, and *PRKAG2* introgressed loci supported by both *VolcanoFinder* and *Signet* approaches were found to present scores within the top 5% of the *T* distribution, suggesting that their haplotype diversity was appreciably shaped by positive selection. Interestingly, adaptive evolution of the *TBC1D1* and *RASGRF2* genes has been previously proposed by studies conducted on different populations of Tibetan ancestry (Peng et al., 2011; Zheng et al., 2023).

Estimating genetic distance between modern and archaic sequences

As a final step for prioritizing the most convincing AI genes supported by *VolcanoFinder*, *Signet* and *LASSI* approaches, as well as to explicitly test whether the Denisovan human species represented a plausible source of archaic alleles for them, we merged Tibetan WGS data with those from low-altitude Han Chinese (CHB) and YRI populations sequenced by the 1000 Genomes Project (Auton et al., 2015), and with the Denisovan genome. We then used the *Haplostrip* algorithm (Marnetto & Huerta-Sánchez, 2017) to estimate genetic distance between modern and archaic haplotypes at the candidate AI genes reported in the previous paragraph. We especially considered genomic windows that included Denisovan-like derived alleles and that presented values of the likelihood *T* statistic supporting an adaptive evolution (see Materials and methods).

Among the tested putative AI loci, *TBC1D1*, *RASGRF2*, *PRKAG2*, and *KRAS* were found to present substantial proportions of Tibetan haplotypes that cluster close to the Denisovan sequence, thus showing the lowest numbers of pairwise differences with respect to it as compared with CHB or YRI haplotypes (Figure 4A and 4B, Figure 4 – figure supplement 2A and 2B). More in detail, 61% of the *TBC1D1* Tibetan haplotypes turned out to be the nearest ones to the archaic sequence by entailing only two pairwise differences with respect to it (Figure 4A), while 29% of Tibetan haplotypes inferred for the *RASGRF2* gene were even identical to the Denisovan DNA (Figure 4B). Interestingly, both these chromosomal intervals were classified by the *LASSI* method as regions whose variation pattern was conformed with the soft selective sweep model, presenting three potential adaptive haplotypes (i.e., those haplotypes that plausibly carry putative advantageous alleles and thus increased in frequency due to positive selection). At each gene, one of these haplotypes was found to contain the Denisovan-like derived alleles that are completely absent in YRI (Figure 4A and 4D). A similar pattern was observed also for the considered *PRKAG2* and *KRAS* genomic windows (Figure 4 – figure supplement 2A and 2B).

Overall, the distribution of similarities between modern and archaic haplotypes described for the four identified AI candidate loci appears to be comparable to that obtained for *EPAS1*, with the sole relevant distinction being represented by an even more pronounced differentiation between Tibetan and CHB patterns plausibly ascribable to the occurrence of a hard (rather than soft) selective sweep at *EPAS1* (Figure 4 – figure supplement 3A), as previously proposed (Simonson et al. 2010; Huerta-Sánchez et al., 2014). In fact, at *EPAS1* the haplotype carrying the lowest number of pairwise differences (N=3) with respect to the archaic one belongs to the Tibetan population, in which it reached 74% of frequency, being instead absent in all the others modern groups considered (Figure 4 – figure supplement 3A). Conversely, the *EGLN1* genomic window showing the highest *T* score according to *LASSI* analysis was characterized by an opposite pattern, with CHB haplotypes being overrepresented among those with the lowest number of pairwise differences (N=3) with respect to the Denisovan genome (Figure 4 – figure supplement 3B). Moreover, the sole *EGLN1* putative adaptive haplotype inferred by *LASSI* for Tibetans was among those presenting the highest number of differences as compared with the archaic sequence, being characterized exclusively by alleles that are not observed in the Denisovan genome and thus suggesting that natural selection targeted modern rather than archaic *EGLN1* variation (Figure 4 – figure supplement 3B).

Discussion

In the present study, we aimed at investigating how far gene flow between the Denisovan archaic human species and the ancestors of modern populations settled in high-altitude regions of the Himalayas contributed to the evolution of key adaptive traits of these human groups, in addition to having conferred them reduced susceptibility to chronic mountain sickness (Huerta-Sánchez et al., 2014). For this purpose, we used WGS data from individuals of Tibetan ancestry to search for genomic signatures ascribable to AI mediated by weak selective events rather than by hard selective sweeps, under the assumption that soft sweeps and/or processes of polygenic adaptation are more likely to have occurred in such remarkably isolated and small-effective population size groups (Gnecchi-Ruscone et al., 2018).

By assembling a large genome-wide dataset including both low- and high-altitude populations, we first framed the available WGS data into the landscape of East Asian genomic variation. This confirmed that the genomes under investigation are well representative of the overall profiles of ancestry components observable in high-altitude Himalayan peoples (Figure 1A and 1B). In fact, the considered individuals were found to show close genetic similarity to other populations of Tibetan ancestry (Jeong et al., 2014) and to Sherpa people from Nepal (Gnecchi-Ruscone et al., 2017), as well as to appreciably diverge from the cline of variation of lowland East-Asians (Abdulla et al., 2009) (Figure 1C).

Based on this evidence, we submitted Tibetan WGS to a pipeline of analyses that implemented multiple independent approaches aimed at identifying genomic regions characterized by signatures putatively ascribable to AI events and by tight functional correlations with each other. According to such a rationale, we shortlisted the candidate introgressed loci that most likely contribute to the same adaptive trait by searching for chromosomal intervals including loci simultaneously showing: i) significant LR scores computed by the *VolcanoFinder* algorithm (Figure 2A and 2B, Figure 2 – figure supplement 3 and 4), ii) Denisovan-like derived alleles belonging to significant networks of functionally related genes reconstructed with the *Signet* method and completely absent in populations of African ancestry (Figure 3A and 3B, Supplement table 6), iii) signatures ascribable to the action of natural selection as pointed out by *LASSI* analysis (Figure 4C and 4D, Figure 4 – figure supplement 2C and 2D), and iv) haplotypes more similar to the Denisovan ones rather than to those observed in other modern human populations, as depicted by the *Haplostrips* approach (Figure 4A and 4B; Figure 4 – figure supplement 2A and 2B).

Overall, in addition to *EPAS1*, which we considered as a positive control for AI, a total of 18 genes encompassed within the putative AI chromosomal intervals identified by the *VolcanoFinder* method (Supplement table 5) were found to have been previously proposed as genomic regions impacted by introgression of Denisovan alleles in Tibetan and/or Han Chinese populations (Huerta-Sánchez et al., 2014; Hu et al., 2017; Browning et al., 2018; Yang et al., 2017; X. Zhang et al., 2021). Among them, *PPARA*, *PRKCE* and *TBC1D1* (Figure 2A, Figure 2 – figure supplement 2A and 2B) were also specifically suggested to have played an adaptive role in high-altitude groups from the Tibetan Plateau (Simonson et al., 2010; Peng et al., 2011; Horscroft et al., 2017; Arciero et al., 2018; Deng et al., 2019; X. Zhang et al., 2021; Zheng et al., 2023). Interestingly, *PPARA* encodes for a nuclear transcription factor whose decreased expression in the

myocardium of rats exposed to hypoxia seems to contribute to the maintenance of the correct heart contractile function despite such a stressful condition (Cole et al., 2016). Similarly, the PRKCE protein kinase C has been demonstrated to exert a cardio-protective role against ischemic injury (Scruggs et al., 2016). Moreover, *TBC1D1* encodes for a protein whose serine phosphorylation sites are targeted by AMP-activated protein kinases (AMPK) after the activation of the *AMPK signalling pathway* as a result of the increase cellular AMP/ATP ratio caused by stresses that induce a lower ATP production (e.g., deprivation of oxygen and/or glucose) or that accelerate ATP consumption (e.g., intense muscle contraction) (Kanehisa & Goto, 2000; Vichaiwong et al., 2010). In addition, another member of the *AMPK signalling pathway*, *PRKAG2*, was suggested by both *VolcanoFinder* analysis and literature data to present putative introgressed Denisovan alleles in Tibetan populations (Figure 2B) (Zhang et al., 2021). Mutations at this locus are known to cause the *PRKAG2 cardiac syndrome*, an inherited disease characterized by ventricular pre-excitation, supraventricular arrhythmias, and cardiac hypertrophy (Zhang et al., 2013; Porto et al., 2016). Dysregulation of AMPK activity mediated by reduction in *PRKAG2* expression and leading to the impairment of glycogen metabolism in cell cultures has been proposed as a possible cause for the development of this pathological condition (Zhang et al., 2013). Conversely, enhanced activation of the *AMPK signalling pathway* during pregnancy coupled with *PRKAG2* overexpression was observed in the placenta of women living at high-altitudes when compared with women living in low-altitude regions (Lorca et al., 2021). Finally, high-altitude individuals of Tibetan ancestry were found to exhibit reduced incidence of major adverse cardiovascular events with respect to low-altitude controls possibly indicating the involvement of protective cardiac mechanisms in the modulation of high-altitude adaptations as previously proposed (Kolář & Ošťádal, 2004; Mallet et al., 2018; Lei et al., 2024). We can thus speculate that adaptive evolution at the *PPARA*, *PRKCE*, *TBC1D1* and *PRKAG2* genomic regions in Tibetans might have contributed to the development of protective mechanisms that reduce cardiovascular risk associated to the hypoxic stress.

In addition to these genes, other two of the identified candidate AI loci (i.e., *RASGRF2* and *KRAS*) have been previously proved to have been targeted by natural selection in Tibetan populations (Peng et al., 2011) or to present putative introgressed archaic alleles (Hu et al., 2017; Browning et al., 2018) (Figure 2 – figure supplement 3 and 4). The proteins encoded by these loci are strictly related from a functional perspective, with the Ras protein specific guanine nucleotide releasing factor 2 representing a calcium-regulated nucleotide exchange factor that activates the RAS protein codified by the proto-oncogene *KRAS* (Kanehisa & Goto, 2000; Sayers et al., 2022).

Despite evidence reported in literature seem to corroborate *VolcanoFinder* results that indicate *PPARA* and *PRKCE* as putatively implicated in AI events experienced by Tibetan ancestors, only *EPAS1*, *TBC1D1*, *RASGRF2*, *PRKAG2* and *KRAS* loci were finally retained after the adopted filtering procedure to represent the most reliable candidate AI loci.

Archaic introgression at these genomic regions was first confirmed by *Signet* analysis, with *TBC1D1* and *PRKAG2* being included in a significant gene network belonging to the *AMPK signalling pathway*, while *RASGRF2* and *KRAS* participate to that related to the *Ras signalling pathway* (Supplement table 6). The

same analysis pointed to *EPAS1* as a member of a significant network belonging to the *Pathways in cancer*, a complex group of biological functions such as those involved in Ras, MAPK, VEGF, and HIF-1 signalling cascades (Kanehisa & Goto, 2000) (Figure 3A and Supplement table 6). These findings emphasize a link between biological mechanisms activated within the context of hypoxic tumour microenvironments in different types of cancers and those involved in high-altitude adaptations, especially as concerns functions that might underlie the improved angiogenesis observed in Tibetan and Sherpa individuals (Gnecchi-Ruscone et al., 2018; Calderón - Gerstein & Torres - Samaniego, 2021). For instance, accumulation of the HIF-1α transcriptional factor in the nucleus of cells close to hypoxic tumour masses comports the activation of diverse biological responses such as the formation of dense capillary structures that permit oxygen and nutrients supplies to cancer cells, thus determining tumour progression and/or treatment failure (Brahimi-Horn et al., 2007; Calderón - Gerstein & Torres - Samaniego, 2021). In line with this evidence, the *Ras* and *MAPK/ERK signalling pathways* have been proposed to play a significant role in promoting angiogenesis by triggering *VEGF* expression, being possibly implicated in the adaptive response to hypoxia evolved by high-altitude populations (Figure 3A) (Kanehisa & Goto, 2000; Kranenburg et al., 2004). Moreover, in the study by Lorenzo et al. (2014) a gain-of-function mutation in the *EGLN1* gene was observed in Tibetans and was proved to enhance the catalytic activity of the HIF prolyl 4-hydroxylase 2 (*PHD2*) under hypoxic conditions. This alters the binding of HIF-2α (the isoform 2 of the inducible hypoxia transcriptional factor encoded by *EPAS1*) and negatively regulates the activation of the *HIF-1 signalling pathway* during hypoxia, eventually offering protection against the detrimental effects of prolonged polycythaemia. When HIF-2α exerts its transcriptional activities along with p300 protein and HIF-1β, it enhances *VEGF* expression and permits the activation of the *VEGF signalling pathway* (Kanehisa & Goto, 2000; Rashid et al., 2021). Accordingly, down regulation of the *HIF-1 signalling pathway* comports the reduction of VEGF mRNA expression (Greenberger et al., 2008; Zhang et al., 2018). Coupled with these observations, results from the *Signet* analysis suggest that in individuals of Tibetan ancestry, when the *HIF-1 signalling pathway* is likely down-regulated in chronic hypoxic conditions (Lorenzo et al., 2014), adaptive changes at the *Ras/MAPK signalling pathways* could represent alternative biological mechanisms that in its place enable to sustain improved angiogenesis and thus permit adequate tissue oxygenation (Figure 3A).

The same five candidate genes, in addition to the *EGLN1* locus considered as a negative control for AI, were also confirmed by the *LASSI* method to have adaptively evolved in the studied populations (Figure 4C and 4D, Figure 4 – figure supplement 1A and 1B, Figure 4 - figure supplement 2, 3C and 3D). In fact, several genomic windows in their associated chromosomal intervals and/or in their flanking regions presented positive values of the computed likelihood *T* statistic, many of which falling in the top 5% of the related distribution, which indicate their non-neutral evolution. Interestingly, *TBC1D1*, *RASGRF2*, *PRKAG2*, *KRAS* and *EPAS1* significant genomic windows pointed out by *LASSI* were also found to include Denisovan-like derived alleles that are completely absent in the YRI African population, suggesting that positive selection acted in Tibetans on haplotypes containing archaic introgressed variation. Such a scenario was further supported by the *Haplostrips* analysis, which revealed for all the loci mentioned above patterns of similarity

between Tibetan and Denisovan haplotypes that are comparable to that observed for *EPAS1* (Figure 4A and 4B, Figure 4 - figure supplement 2A and 2B), with the sole exception being represented by *EGLN1* (Figure 4 – figure supplement 1B, Figure 4 – figure supplement 3B and 3D, Figure 4 – figure supplement 3A).

Overall, we collected multiple evidence supporting both the archaic origin and the adaptive role of variation at *TBC1D1*, *RASGRF2*, *PRKAG2* and *KRAS* genes in populations of Tibetan ancestry. Genetic signatures at such loci are especially consistent with the hypothesis of adaptive events mediated by soft selective sweeps and/or polygenic mechanisms that involved haplotypes including both modern and archaic introgressed alleles. Therefore, the results obtained have succeeded in expanding the knowledge about AI events having involved the ancestors of modern high-altitude Himalayan populations and Denisovans and emphasized once more the complexity of the adaptive phenotype evolved by these human groups to cope with challenges imposed by hypobaric hypoxia. Accordingly, they offer new insights for future studies aimed at elucidating the molecular mechanisms by which several genes along with *TBC1D1*, *PRKAG2*, *RASGRF2* and *KRAS* interact with each other and contribute to mediate physiological adjustments that are crucial for human adaptation to the high-altitude environment.

Materials and methods

Dataset composition and curation

The dataset used in the present study included WGS data for 27 individuals of Tibetan ancestry recruited from the high-altitude Nepalese districts of Mustang and Ghorka (Jeong et al., 2018). Although these subjects reside in Nepal, they have been previously proved to speak Tibetan dialects and to live in communities showing religious and social organizations proper of populations settled on the Tibetan Plateau, being also biologically representative of high-altitude Tibetan people (Cho et al., 2017). To filter for high-quality genotypes, the dataset was subjected to QC procedures using the software PLINK v1.9 (Purcell et al., 2007). In detail, we retained autosomal SNVs characterized by no significant deviations from the Hardy-Weinberg Equilibrium ($P > 5.3 \times 10^{-9}$ after Bonferroni correction for multiple testing), as well as loci/samples showing less than 5% of missing data. Moreover, we removed SNVs with ambiguous A/T or G/C alleles and multiallelic variants, obtaining a dataset composed by 27 individuals and 6,921,628 SNVs. WGS data were finally phased with SHAPEIT2 v2.r904 (Delaneau et al., 2013) by applying default parameters and using the 1000 Genomes Project dataset as a reference panel (Auton et al., 2015) and HapMap phase 3 recombination maps.

Population structure analyses

To assess representativeness, genetic homogeneity, and ancestry composition of Tibetan WGS included in the dataset, we performed genotype-based population structure analyses. For this purpose, we merged the unphased WGS dataset with genome-wide genotyping data for 34 East-Asian populations (Gnecchi-Ruscone et al., 2017;Landini et al., 2021) and we applied the same QC described above. The obtained extended dataset included 231,947 SNVs and was used to assess the degree of recent shared ancestry (i.e., identity-by-

descent, IBD) for each pair of individuals. Identity-by-state (IBS) estimates were thus used to calculate an IBD kinship coefficient and a threshold of PI_HAT = 0.270 was considered to remove closely related subjects to the second degree (Ojeda-Granados et al., 2022). To discard variants in linkage disequilibrium (LD) with each other we then removed one SNV for each pair showing $r^2 > 0.2$ within sliding windows of 50 SNVs and advancing by five SNVs at the time. The obtained LD-pruned dataset was finally filtered for variants with minor allele frequency (MAF) < 0.01 and used to compute PCA utilizing the *smartpca* method implemented in the EIGENSOFT package (Patterson et al., 2006), as well as to run the ADMIXTURE algorithm version 1.3.0 (Alexander et al., 2009) by testing K = 2 to K = 12 population clusters. In detail, 25 replicates with different random seeds were run for each K and we retained only those presenting the highest log-likelihood values. In addition, cross-validation errors were calculated for each K to identify the value that best fit the data. Both PCA and ADMIXTURE results were visualized with the R software version 4.0.5. ADMIXTURE results were visualized with the R software version 4.0.5.

Detecting signatures of adaptive introgression

To identify chromosomal regions showing signatures putatively ascribable to AI events, we submitted the phased WGS dataset to the *VolcanoFinder* pipeline, which relies on the analysis of the allele frequency spectrum of the population that is supposed to have experienced archaic introgression (Setter et al., 2020). The *VolcanoFinder* algorithm was chosen among several approaches developed to detect AI signatures according to the following reasons. First of all, it can test jointly both archaic introgression and adaptive evolution according to a model that differs from those considered by other statistics that are aimed at identifying chromosomal segments showing low divergence with respect to a specific archaic sequence and/or enriched in alleles uniquely shared between the admixed group and the archaic source and characterized by a frequency above a certain threshold in the population under study (Racimo et al., 2017). In fact, these methods are especially useful to test an evolutionary scenario conformed to that expected in the case that adaptation was mediated by hard selective sweeps. On the contrary, *VocanoFinder* was proved to have an elevated power in the identification of AI events mediated by more than one predominant haplotype (Setter et al., 2020), as expected when soft sweeps/polygenic adaptations occurred. Moreover, *VolcanoFinder* relies on less demanding computational efforts with respect to algorithms that require to be trained on genomic simulations built specifically to reflect the evolutionary history of the population under study (Gower et al., 2021; Zhang et al., 2023), but possibly introducing bias in the obtained results if the information that guides simulations is not accurate.

We thus scanned Tibetan WGS data using the *VolcanoFinder* method and calculating two statistics for each polymorphic site: α (subsequently converted in -logα) and LR which are informative respectively of: i) the strength of natural selection and ii) the conformity to the evolutionary model of adaptive introgression (see Supplement Information for further details about the evolutionary model tested). Since elevated LR scores are assumed to support AI signatures (Setter et al., 2020), we filtered the most significant results obtained by focusing on SNVs showing LR values falling in the positive tail (i.e., top 5%) of the distribution built for

such a statistic. We then visualized the distribution of both α and LR parameters across the genomic regions including the *EPAS1* and *EGLN1* genes, which we considered respectively as positive and negative control loci for AI and by investigating chromosomal intervals spanning 50kb upstream and downstream to these genes. We then defined the new candidate AI genomic regions based on their conformity with the pattern observed for the positive AI control gene (i.e., according to the detection of multiple peaks of LR scores consistently distributed in the entire genomic region considered and coupled with elevated values of the -logα parameter). Moreover, we relied on evidence advanced by previous studies aimed at detecting archaic introgression signatures from WGS data for individuals with Tibetan and/or Han Chinese ancestries (Hu et al., 2017; Browning et al., 2018; X. Zhang et al., 2021) to filter out loci potentially targeted by natural selection in Tibetans, but with questionable archaic origins.

Identifying gene networks including Denisovan-like derived alleles

To validate archaic introgression signatures inferred with *VolcanoFinder* by using an independent method, we followed the approach described by Gouy & Excoffier (2020), with the aim of identifying biological functions whose underlying genomic regions might have been significantly shaped by Denisovan introgression. For this purpose, we used the *Signet* algorithm to reconstruct network of genes that participate to the same biological pathway and that also included Denisovan-like derived alleles observable in the Tibetan population but not in an outgroup population of African ancestry, by assuming that only Eurasian *H. sapiens* populations experienced Denisovan admixture. To do so, we first compared the Tibetan and Denisovan genomes to assess which SNVs were present in both modern and archaic sequences. These loci were further compared with the ancestral reconstructed refence human genome sequence to discard those presenting an ancestral state (i.e., that we have in common with several primate species). Moreover, we further filtered the considered variants by retaining only those alleles that were completely absent (i.e., with frequency equal to zero) in the YRI population sequenced by the 1000 Genomes Project (Auton et al., 2015). We then calculated the average frequency in the Tibetan population of the Denisovan-like derived alleles observable in each gene and we used the obtained genomic distribution to inform *Signet* analysis (see Supplement information). We performed five independent runs of the *Signet* algorithm to check for consistency of the significant gene networks and functional pathways identified and we finally depicted the confirmed results using Cytoscape v3.9.1 (Shannon et al., 2003).

Testing adaptive evolution of candidate introgressed loci

To confirm signatures ascribable to the action of natural selection at the putative introgressed loci supported by both *VolcanoFinder* and *Signet* analyses, we applied the *LASSI* likelihood method (Harris & DeGiorgio, 2020) on the available Tibetan WGS data. Such an approach detects and distinguishes genomic regions that have experienced different types of selective events (i.e., strong and weak ones) and has been demonstrated to be more powerful in doing so than other haplotype-based approaches (Harris & DeGiorgio, 2020).

Specifically, we calculated for each genomic window the likelihood *T* statistic, which measures the conformity of variability patterns of the analysed region to those expected according to a haplotype frequency spectrum under adaptive rather than neutral evolution. In addition, the *LASSI* algorithm calculated the parameter *m* (i.e., the number of sweeping haplotypes) for each genomic region, thus classifying them as affected by hard sweeps (when $m = 1$) or soft sweeps (when $m > 1$). In detail, *T* scores significantly different from zero indicate the conformity with a non-neutral evolutionary scenario, with ever higher likelihood scores being indicative of increasingly robust evidence for a selective event (Harris & DeGiorgio, 2020).

The method requires to fix a custom value for the length of the considered genomic windows, which are measured in terms of the number of SNVs included in them, and to move windows by 1 SNV after each computation. Therefore, we selected this fixed-length value (i.e., 13 SNVs) by estimating the average number of SNVs included into a haplotype block as defined for the population under study by using the *--blocks* function implemented in PLINK v1.9 (Purcell et al., 2007). Moreover, by following the indications by Harris & DeGiorgio (2020) of choosing values for the fixed number of haplotypes in the spectrum (i.e., K values) <10 for increasing the power of the *T* statistic, we set it at seven. Finally, we choose the likelihood model 3 to calculate the *T* statistic and we applied the *LASSI* algorithm to the phased WGS dataset.

We then focused on the genomic windows showing *T* scores falling in the positive tail (i.e., top 5%) of the obtained distribution and we crosschecked these results with those significant ones pointed out by *VolcanoFinder* and *Signet* analyses to shortlist genomic regions having plausibly experienced both archaic introgression and adaptive evolution.

Haplostrip **analysis**

To explicitly test whether the putative adaptive introgressed loci pointed out by *VolcanoFinder*, *Signet*, and *LASSI* analyses present variation patterns compatible with a scenario of introgression from the Denisovan archaic human species, we estimated genetic distance between modern and archaic haplotypes inferred for those genomic windows supported by all the methods mentioned above, as well as for *EPAS1* and *EGLN1* for the sake of comparison with established positive and negative control genes that have been previously proved to be involved or not in AI events. For this purpose, we used the *Haplostrip* pipeline, as described in previous studies (Huerta-Sánchez et al., 2014;Marnetto & Huerta-Sánchez, 2017). Moreover, since the *EGLN1* gene did not include any Denisovan-like variant as identified according to the criteria described in the previous paragraphs, we choose to build the *Haplostrips* heatmap by considering the *EGLN1* genomic window associated with the highest value of the *LASSI* statistic.

In detail, we merged the 27 Tibetan whole genomes under study (Jeong et al., 2018) with 27 CHB, 27 YRI WGSs (Auton et al., 2015) and with the Denisovan genome (Meyer et al., 2013) (downloadable at: http://cdna.eva.mpg.de/neandertal/altai/Denisovan/). The CHB population, which is known to share a relatively recent common ancestry with Tibetans, was used as a “negative low-altitude control” (i.e., as a group whose ancestors experienced Denisovan introgression, but did not evolve high-altitude adaptation). YRI individuals were instead used as the outgroup population (i.e., a population that presumably did not

experience Denisovan admixture), as previously proposed (Zhang et al., 2021). We then phased the assembled dataset with SHAPEIT2 v2.r904 (Delaneau et al., 2013), as described in the *Dataset composition and curation* section and we run the *Haplostrip* algorithm.

Acknowledgements

We acknowledge support from the Fondazione Cassa di Risparmio in Bologna through the project "Genetic adaptation and acclimatization to high altitude as experimental models to investigate the biological mechanisms that regulate physiological responses to hypoxia", which was granted to MS (n. 2019.0552).

Additional information

Authors contributions

Giulia Ferraretti and Paolo Abondio, Data curation, Software, Formal analysis, Investigation, Writing - original draft; Marta Alberti, Software, Formal analysis; Agnese Dezi, Phurba T. Sherpa, Paolo Cocco, Massimiliano Tiriticco, Marco di Marcello, Guido Alberto Gnecchi-Ruscone, Luca Natali, Angela Corcelli, Giorgio Marinelli, Davide Peluzzi, Data curation, Writing - review and editing; Stefania Sarno, Formal analysis, Writing - original draft; Marco Sazzini, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing - review and editing.

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Arciero, E., Kraaijenbrink, T., Asan, Haber, M., Mezzavilla, M., Ayub, Q., Wang, W., Pingcuo, Z., Yang, H., Wang, J., Jobling, M. A., Van Driem, G., Xue, Y., De Knijff, P., & Tyler-Smith, C. (2018). Demographic history and genetic adaptation in the himalayan region inferred from genome-wide SNP genotypes of 49 populations. *Molecular Biology and Evolution*, *35*(8), 1916–1933. https://doi.org/10.1093/molbev/msy094

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., … Schloss, J. A. (2015). A global reference for human genetic variation. In *Nature* (Vol. 526, Issue 7571, pp. 68–74). Nature Publishing Group. https://doi.org/10.1038/nature15393

Beall, C. M. (2007). Two routes to functional adaptation: Tibetan and andean high-altitude natives. *In the Light of Evolution*, *1*, 239–255. https://doi.org/10.17226/11790

Beall, C. M., Cavalleri, G. L., Deng, L., Elston, R. C., Gao, Y., Knight, J., Li, C., Li, J. C., Liang, Y., McCormack, M., Montgomery, H. E., Pan, H., Robbins, P. A., Shianna, K. V., Tam, S. C., Tsering, N., Veeramah, K. R., Wang, W., Wangdui, P., … Zheng, Y. T. (2010). Natural selection on EPAS1 (HIF2α) associated with low hemoglobin concentration in Tibetan highlanders. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11459–11464. https://doi.org/10.1073/pnas.1002443107

Benjamin, V., & M, A. J. (2014). Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. *Science*, *343*(6174), 1017–1021. https://doi.org/10.1126/science.1245938

Bigham, A., Bauchet, M., Pinto, D., Mao, X., Akey, J. M., Mei, R., Scherer, S. W., Julian, C. G., Wilson, M. J., López Herráez, D., Brutsaert, T., Parra, E. J., Moore, L. G., & Shriver, M. D. (2010). Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data. *PLOS Genetics*, *6*(9), e1001116. https://doi.org/10.1371/journal.pgen.1001116

Brahimi-Horn, M. C., Chiche, J., & Pouysségur, J. (2007). Hypoxia and cancer. In *Journal of Molecular Medicine* (Vol. 85, Issue 12, pp. 1301–1307). https://doi.org/10.1007/s00109-007-0281-3

Browning, S. R., Browning, B. L., Zhou, Y., Tucci, S., & Akey, J. M. (2018). Analysis of Human Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture. *Cell*, *173*(1), 53-61.e9. https://doi.org/10.1016/j.cell.2018.02.031

Calderón - Gerstein, W. S., & Torres - Samaniego, G. (2021). High altitude and cancer: An old controversy. In *Respiratory Physiology and Neurobiology* (Vol. 289). Elsevier B.V. https://doi.org/10.1016/j.resp.2021.103655

Catar, R., Witowski, J., Wagner, P., Annett Schramm, I., Kawka, E., Philippe, A., Dragun, D., & Jörres, A. (2013). The proto-oncogene C-Fos transcriptionally regulates VEGF production during peritoneal inflammation. *Kidney International*, *84*(6), 1119–1128. https://doi.org/10.1038/ki.2013.217

Chen, X., Li, X., Zhang, W., He, J., Xu, B., Lei, B., Wang, Z., Cates, C., Rousselle, T., & Li, J. (2018). Activation of AMPK inhibits inflammatory response during hypoxia and reoxygenation through modulating JNK-mediated NF-κB pathway. *Metabolism: Clinical and Experimental*, *83*, 256–270. https://doi.org/10.1016/j.metabol.2018.03.004

Cho, J. I., Basnyat, B., Jeong, C., di Rienzo, A., Childs, G., Craig, S. R., Sun, J., & Beall, C. M. (2017). Ethnically Tibetan women in Nepal with low hemoglobin concentration have better reproductive outcomes. *Evolution, Medicine, and Public Health*, *2017*(1), 82–96. https://doi.org/10.1093/emph/eox008

Cole, M. A., Jamil, A. H. A., Heather, L. C., Murray, A. J., Sutton, E. R., Slingo, M., Sebag-Montefiore, L., Tan, S. C., Aksentijevíc, D., Gildea, O. S., Stuckey, D. J., Yeoh, K. K., Carr, C. A., Evans, R. D., Aasum, E., Schofield, C. J., Ratcliffe, P. J., Neubauer, S., Robbins, P. A., & Clarke, K. (2016). On the pivotal role of PPARa in adaptation of the heart to hypoxia and why fat in the diet increases hypoxic injury. *FASEB Journal*, *30*(8), 2684–2697. https://doi.org/10.1096/fj.201500094R

Dannemann, M., & Racimo, F. (2018). Something old, something borrowed: admixture and adaptation in human evolution. In *Current Opinion in Genetics and Development* (Vol. 53, pp. 1–8). Elsevier Ltd. https://doi.org/10.1016/j.gde.2018.05.009

Delaneau, O., Zagury, J. F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, *10*(1), 5–6. https://doi.org/10.1038/nmeth.2307

Deng, L., Zhang, C., Yuan, K., Gao, Y., Pan, Y., Ge, X., He, Y., Yuan, Y., Lu, Y., Zhang, X., Chen, H., Lou, H., Wang, X., Lu, D., Liu, J., Tian, L., Feng, Q., Khan, A., Yang, Y., … Xu, S. (2019). Prioritizing natural-selection signals from the deep-sequencing genomic data suggests multi-variant adaptation in Tibetan highlanders. *National Science Review*, *6*(6), 1201–1222. https://doi.org/10.1093/nsr/nwz108

Dengler, F. (2020). Activation of ampk under hypoxia: Many roads leading to rome. In *International Journal of Molecular Sciences* (Vol. 21, Issue 7). MDPI AG. https://doi.org/10.3390/ijms21072428

Enard, D., & Petrov, D. A. (2018). Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. *Cell*, *175*(2), 360-371.e13. https://doi.org/10.1016/j.cell.2018.08.034

Gittelman, R. M., Schraiber, J. G., Vernot, B., Mikacenic, C., Wurfel, M. M., & Akey, J. M. (2016). Archaic Hominin Admixture Facilitated Adaptation to Out-of-Africa Environments. *Current Biology*, *26*(24), 3375–3382. https://doi.org/https://doi.org/10.1016/j.cub.2016.10.041

Gnecchi-Ruscone, G. A., Abondio, P., De Fanti, S., Sarno, S., Sherpa, M. G., Sherpa, P. T., Marinelli, G., Natali, L., Di Marcello, M., Peluzzi, D., Luiselli, D., Pettener, D., & Sazzini, M. (2018). Evidence of polygenic adaptation to high altitude from Tibetan and Sherpa genomes. *Genome Biology and Evolution*, *10*(11), 2919–2930. https://doi.org/10.1093/gbe/evy233

Gnecchi-Ruscone, G. A., Abondio, P., de Fanti, S., Sarno, S., Sherpa, M. G., Sherpa, P. T., Marinelli, G., Natali, L., di Marcello, M., Peluzzi, D., Luiselli, D., Pettener, D., & Sazzini, M. (2018). Evidence of polygenic adaptation to high altitude from Tibetan and Sherpa genomes. *Genome Biology and Evolution*, *10*(11), 2919–2930. https://doi.org/10.1093/gbe/evy233

Gnecchi-ruscone, G. A., Jeong, C., Fanti, S. De, Sarno, S., Gentilini, D., Blasio, A. M. Di, Sherpa, M. G., Sherpa, P. T., Marinelli, G., Marcello, M. Di, Natali, L., Peluzzi, D., Pettener, D., Rienzo, A. Di, Luiselli, D., & Sazzini, M. (2017). The genomic landscape of Nepalese Tibeto-Burmans reveals new insights into the recent peopling of Southern Himalayas. *Scientific Reports*, *January*, 1–12. https://doi.org/10.1038/s41598-017-15862-z

Gouy, A., & Excoffier, L. (2020). Polygenic Patterns of Adaptive Introgression in Modern Humans Are Mainly Shaped by Response to Pathogens. *Molecular Biology and Evolution*, *37*(5), 1420–1433. https://doi.org/10.1093/molbev/msz306

Gouy, A., Excoffier, L., & Nielsen, R. (2020). Polygenic Patterns of Adaptive Introgression in Modern Humans Are Mainly Shaped by Response to Pathogens. *Molecular Biology and Evolution*, *37*(5), 1420–1433. https://doi.org/10.1093/molbev/msz306

Gower, G., Picazo, P. I., Fumagalli, M., & Racimo, F. (2021). Detecting adaptive introgression in human evolution using convolutional neural networks. *ELife*, *10*. https://doi.org/10.7554/eLife.64669

Green, R., Krause, J., Briggs, A., Rasilla Vives, M. de la, & Fortea Pérez, F. J. (2010). A draft sequence of the neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021PubMedID:20448178

Greenberger, L. M., Horak, I. D., Filpula, D., Sapra, P., Westergaard, M., Frydenlund, H. F., Albæk, C., Schrøder, H., & Ørum, H. (2008). A RNA antagonist of hypoxia-inducible factor-1α, EZN-2968, inhibits tumor cell growth. *Molecular Cancer Therapeutics*, *7*(11), 3598–3608. https://doi.org/10.1158/1535-7163.MCT-08-0510

Hamid, I., Korunes, K. L., Beleza, S., & Goldberg, A. (2021). Rapid adaptation to malaria facilitated by admixture in the human population of cabo verde. *ELife*, *10*, 1–24. https://doi.org/10.7554/eLife.63177

Harris, A. M., & DeGiorgio, M. (2020). A likelihood approach for uncovering selective sweep signatures from haplotype data. *Molecular Biology and Evolution*, *37*(10), 3023–3046. https://doi.org/10.1093/molbev/msaa115

Horscroft, J. A., Kotwica, A. O., Laner, V., West, J. A., Hennis, P. J., Levett, D. Z. H., Howard, D. J., Fernandez, B. O., Burgess, S. L., Ament, Z., Gilbert-Kawai, E. T., Vercueil, A., Landis, B. D., Mitchell, K., Mythen, M. G., Branco, C., Johnson, R. S., Feelisch, M., Montgomery, H. E., … Murray, A. J. (2017). Metabolic basis to sherpa altitude adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(24), 6382–6387. https://doi.org/10.1073/pnas.1700527114

Hu, H., Petousi, N., Glusman, G., Yu, Y., Bohlender, R., Tashi, T., Downie, J. M., Roach, J. C., Cole, A. M., Lorenzo, F. R., Rogers, A. R., Brunkow, M. E., Cavalleri, G., Hood, L., Alpatty, S. M., Prchal, J. T., Jorde, L. B., Robbins, P. A., Simonson, T. S., & Huff, C. D. (2017). Evolutionary history of Tibetans inferred from whole-genome sequencing. *PLoS Genetics*, *13*(4). https://doi.org/10.1371/journal.pgen.1006675

Huerta-Sánchez, E., Jin, X., Asan, Bianba, Z., Peter, B. M., Vinckenbosch, N., Liang, Y., Yi, X., He, M., Somel, M., Ni, P., Wang, B., Ou, X., Huasang, Luosang, J., Cuo, Z. X. P., Li, K., Gao, G., Yin, Y., … Nielsen, R. (2014). Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, *512*(7513), 194–197. https://doi.org/10.1038/nature13408

Israel, H., W, W. G., Rolf, Q., Mathieu, D., Rainer, G., Leslie, K., Avner, A., Miryam, B.-M., Helene, V., Norbert, M., Luis, A. J., María, M.-T., María, B. de C. J., Cinzia, F., Laura, M.-F., Rachel, S., Hila, M., A, K. V., Viviane, S., … Mina, W.-E. (2018). The earliest modern humans outside Africa. *Science*, *359*(6374), 456–459. https://doi.org/10.1126/science.aap8369

Jeong, C., Alkorta-Aranburu, G., Basnyat, B., Neupane, M., Witonsky, D. B., Pritchard, J. K., Beall, C. M., & Di Rienzo, A. (2014). Admixture facilitates genetic adaptations to high altitude in Tibet. *Nature Communications*, *5*, 1–7. https://doi.org/10.1038/ncomms4281

Jeong, C., Witonsky, D. B., Basnyat, B., Neupane, M., Beall, C. M., Childs, G., Craig, S. R., Novembre, J., & Di Rienzo, A. (2018). Detecting past and ongoing natural selection among ethnically Tibetan women at high altitude in Nepal. *PLOS Genetics*, *14*(9), e1007650. https://doi.org/10.1371/journal.pgen.1007650

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. In *Nucleic Acids Research* (Vol. 28, Issue 1). http://www.genome.ad.jp/kegg/

Kolář, F., & Ošťádal, B. (2004). Molecular Mechanisms of Cardiac Protection by Adaptation to Chronic Hypoxia. *Physiol. Res*, *53*, 3–13. http://www.biomed.cas.cz/physiolres

Kranenburg, O., Gebbink, M. F. B. G., & Voest, E. E. (2004). Stimulation of angiogenesis by Ras proteins. In *Biochimica et Biophysica Acta - Reviews on Cancer* (Vol. 1654, Issue 1, pp. 23–37). Elsevier. https://doi.org/10.1016/j.bbcan.2003.09.004

Landini, A., Yu, S., Gnecchi-Ruscone, G. A., Abondio, P., Ojeda-Granados, C., Sarno, S., de Fanti, S., Gentilini, D., di Blasio, A. M., Jin, H., Nguyen, T. T., Romeo, G., Prata, C., Bortolini, E., Luiselli, D., Pettener, D., & Sazzini, M. (2021). Genomic adaptations to cereal-based diets contribute to mitigate metabolic risk in some human populations of East Asian ancestry. *Evolutionary Applications*, *14*(2), 297–313. https://doi.org/10.1111/eva.13090

Laurent, A.-R., J., J. M., Subhash, K., Alasdair, M., Klara, D., Loren, G., Farbod, B., Baback, G., Ma, L., A., P. F., Joshua, K., Mary, C., Derek, M., Raja, R., Meral, B., E., M. S. G., Martin,

M., A., G. L., Sofia, T., … Peter, P. (2011). The Shaping of Modern Human Immune Systems by Multiregional Admixture with Archaic Humans. *Science*, *334*(6052), 89–94. https://doi.org/10.1126/science.1209202

Lei, L., Liu, M., Ma, D., Lei, X., Zeng, S., Li, P., Huang, K., Lyu, J., & Lei, Q. (2024). Cardioprotective effects of high-altitude adaptation in cardiac surgical patients: a retrospective cohort study with propensity score matching. *Frontiers in Cardiovascular Medicine*, *11*. https://doi.org/10.3389/fcvm.2024.1347552

Liu, C. C., Witonsky, D., Gosling, A., Lee, J. H., Ringbauer, H., Hagan, R., Patel, N., Stahl, R., Novembre, J., Aldenderfer, M., Warinner, C., di Rienzo, A., & Jeong, C. (2022). Ancient genomes from the Himalayas illuminate the genetic history of Tibetans and their Tibeto-Burman speaking neighbors. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28827-2

Lorca, R. A., Houck, J. A., Laurent, L. C., Matarazzo, C. J., Baker, K., Horii, M., Nelson, K. K., Bales, E. S., Euser, A. G., Parast, M. M., Moore, L. G., & Julian, C. G. (2021). High altitude regulates the expression of AMPK pathways in human placenta. *Placenta*, *104*, 267–276. https://doi.org/10.1016/j.placenta.2021.01.010

Lorenzo, F. R., Huff, C., Myllymäki, M., Olenchock, B., Swierczek, S., Tashi, T., Gordeuk, V., Wuren, T., Ge, R. L., McClain, D. A., Khan, T. M., Koul, P. A., Guchhait, P., Salama, M. E., Xing, J., Semenza, G. L., Liberzon, E., Wilson, A., Simonson, T. S., … Prchal, J. T. (2014). A genetic mechanism for Tibetan high-altitude adaptation. *Nature Genetics*, *46*(9), 951–956. https://doi.org/10.1038/ng.3067

Mallet, R. T., Manukhina, E. B., Ruelas, S. S., Caffrey, J. L., & Downey, H. F. (2018). Cardioprotection by intermittent hypoxia conditioning: evidence, mechanisms, and therapeutic potential. *Am J Physiol Heart Circ Physiol*, *315*, 216–232. https://doi.org/10.1152/ajpheart.00060.2018.-The

Marnetto, D., & Huerta-Sánchez, E. (2017). Haplostrips: revealing population structure through haplotype visualization. *Methods in Ecology and Evolution*, *8*(10), 1389–1392. https://doi.org/10.1111/2041-210X.12747

Maxwell, P. H., & Ratcliffe, P. J. (2002). Oxygen sensors and angiogenesis. *Seminars in Cell and Developmental Biology*, *13*(1), 29–37. https://doi.org/10.1006/scdb.2001.0287

McArthur, E., Rinker, D. C., & Capra, J. A. (2021). Quantifying the contribution of Neanderthal introgression to the heritability of complex traits. *Nature Communications*, *12*(1), 1–14. https://doi.org/10.1038/s41467-021-24582-y

McCoy, R. C., Wakefield, J., & Akey, J. M. (2017). Impacts of Neanderthal-Introgressed Sequences on the Landscape of Human Gene Expression. *Cell*, *168*(5), 916-927.e12. https://doi.org/https://doi.org/10.1016/j.cell.2017.01.038

Meyer, M., Kircher, M., Gansauge, M., Li, H., Mallick, S., Schraiber, J. G., Jay, F., Prüfer, K., De, C., Sudmant, P. H., Alkan, C., Fu, Q., Do, R., Rohland, N., Siebauer, M., Green, R. E., Bryc, K., Briggs, A. W., Dabney, J., … Eichler, E. E. (2013). *A High Coverage Genome Sequence From an Archaic*. *338*(6104), 222–226. https://doi.org/10.1126/science.1224344.A

Monje, P., Hernández-Losa, J., Lyons, R. J., Castellone, M. D., & Gutkind, J. S. (2005). Regulation of the transcriptional activity of c-Fos by ERK: A novel role for the prolyl isomerase Pin1. *Journal of Biological Chemistry*, *280*(42), 35081–35084. https://doi.org/10.1074/jbc.C500353200

Ojeda-Granados, C., Abondio, P., Setti, A., Sarno, S., Gnecchi-Ruscone, G. A., González-Orozco, E., De Fanti, S., Jiménez-Kaufmann, A., Rangel-Villalobos, H., Moreno-Estrada, A., & Sazzini, M. (2022). Dietary, Cultural, and Pathogens-Related Selective Pressures Shaped Differential Adaptive Evolution among Native Mexican Populations. *Molecular Biology and Evolution*, *39*(1). https://doi.org/10.1093/molbev/msab290

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, *2*(12), 2074–2093. https://doi.org/10.1371/journal.pgen.0020190

Peng, Y., Yang, Z., Zhang, H., Cui, C., Qi, X., Luo, X., Tao, X., Wu, T., Ouzhuluobu, Basang, Ciwangsangbu, Danzengduojie, Chen, H., Shi, H., & Su, B. (2011). Genetic variations in tibetan populations and high-altitude adaptation at the Himalayas. *Molecular Biology and Evolution*, *28*(2), 1075–1081. https://doi.org/10.1093/molbev/msq290

Porto, A. G., Brun, F., Severini, G. M., Losurdo, P., Fabris, E., Taylor, M. R. G., Mestroni, L., & Sinagra, G. (2016). Clinical Spectrum of PRKAG2 Syndrome. *Circulation: Arrhythmia and Electrophysiology*, *9*(1), e003121. https://doi.org/10.1161/CIRCEP.115.003121

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P. H., de Filippo, C., Li, H., Mallick, S., Dannemann, M., Fu, Q., Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., … Pääbo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, *505*(7481), 43–49. https://doi.org/10.1038/nature12886

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Racimo, F., Marnetto, D., & Huerta-Sánchez, E. (2017). Signatures of archaic adaptive introgression in present-day human populations. *Molecular Biology and Evolution*, *34*(2), 296–317. https://doi.org/10.1093/molbev/msw216

Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, *16*(6), 359–371. https://doi.org/10.1038/nrg3936

Rashid, M., Zadeh, L. R., Baradaran, B., Molavi, O., Ghesmati, Z., Sabzichi, M., & Ramezani, F. (2021). Up-down regulation of HIF-1α in cancer progression. In *Gene* (Vol. 798). Elsevier B.V. https://doi.org/10.1016/j.gene.2021.145796

Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., Viola, B., Briggs, A. W., Stenzel, U., Johnson, P. L. F., Maricic, T., Good, J. M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E. E., … Pääbo, S. (2010). Genetic history of an archaic hominin group from Denisova cave in Siberia. *Nature*, *468*(7327), 1053–1060. https://doi.org/10.1038/nature09710

Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., Patterson, N., & Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*, *507*(7492), 354–357. https://doi.org/10.1038/nature12961

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., … Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, *50*(D1), D20–D26. https://doi.org/10.1093/nar/gkab1112

Scruggs, S. B., Wang, D., & Ping, P. (2016). PRKCE gene encoding protein kinase C-epsilon—Dual roles at sarcomeres and mitochondria in cardiomyocytes. In *Gene* (Vol. 590, Issue 1, pp. 90–96). Elsevier B.V. https://doi.org/10.1016/j.gene.2016.06.016

Setter, D., Mousset, S., Cheng, X., Nielsen, R., DeGiorgio, M., & Hermisson, J. (2020). VolcanoFinder: Genomic scans for adaptive introgression. In *PLoS Genetics* (Vol. 16, Issue 6). https://doi.org/10.1371/journal.pgen.1008867

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Simonson, T. S., Yang, Y., Huff, C. D., Yun, H., Qin, G., Witherspoon, D. J., Bai, Z., Lorenzo, F. R., Xing, J., Jorde, L. B., Prchal, J. T., & Ge, R. (2010). Genetic evidence for high-altitude adaptation in Tibet. *Science (New York, N.Y.)*, *329*(5987), 72–75. https://doi.org/10.1126/science.1189406

Simonti, C. N., Vernot, B., Bastarache, L., Bottinger, E., Carrell, D. S., Chisholm, R. L., Crosslin, D. R., Hebbring, S. J., Jarvik, G. P., Kullo, I. J., Li, R., Pathak, J., Ritchie, M. D., Roden, D. M., Verma, S. S., Tromp, G., Prato, J. D., Bush, W. S., Akey, J. M., … Capra, J. A. (2016). The phenotypic legacy of admixture between modern humans and Neandertals. *Science*, *351*(6274), 737–741. https://doi.org/10.1126/science.aad2149

Stevens, C., Lin, Y., Sanchez, M., Amin, E., Copson, E., White, H., Durston, V., Eccles, D. M., & Hupp, T. (2007). A germ line mutation in the death domain of DAPK-1 inactivates ERK-induced apoptosis. *Journal of Biological Chemistry*, *282*(18), 13791–13803. https://doi.org/10.1074/jbc.M605649200

Vichaiwong, K., Purohit, S., An, D., Toyoda, T., Jessen, N., Hirshman, M. F., & Goodyear, L. J. (2010). Contraction regulates site-specific phosphorylation of TBC1D1 in skeletal muscle. *Biochemical Journal*, *431*(2), 311–320. https://doi.org/10.1042/BJ20101100

Wang, Y., Zou, X., Wang, M., Yuan, D., Yang, L., Zeng, Y., Cheng, F., Tang, R., & He, G. (2022). The genomic history of southwestern Chinese populations demonstrated massive population migration and admixture among proto-Hmong–Mien speakers and incoming migrants. *Molecular Genetics and Genomics*, *297*(1), 241–262. https://doi.org/10.1007/s00438-021-01837-3

Xu, S., Li, S., Yang, Y., Tan, J., Lou, H., Jin, W., Yang, L., Pan, X., Wang, J., Shen, Y., Wu, B., Wang, H., & Jin, L. (2011). A genome-wide search for signals of high-altitude adaptation in tibetans. *Molecular Biology and Evolution*, *28*(2), 1003–1011. https://doi.org/10.1093/molbev/msq277

Yang, J., Jin, Z. B., Chen, J., Huang, X. F., Li, X. M., Liang, Y. B., Mao, J. Y., Chen, X., Zheng, Z., Bakshi, A., Zheng, D. D., Zheng, M. Q., Wray, N. R., Visscher, P. M., Lu, F., & Qu, J. (2017). Genetic signatures of high-altitude adaptation in Tibetans. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(16), 4189–4194. https://doi.org/10.1073/pnas.1617042114

Yang, X. Y., Rakha, A., Chen, W., Hou, J., Qi, X. bin, Shen, Q. K., Dai, S. S., Sulaiman, X., Abdulloevich, N. T., Afanasevna, M. E., Ibrohimovich, K. B., Chen, X., Yang, W. K., Adnan, A., Zhao, R. H., Yao, Y. G., Su, B., Peng, M. S., & Zhang, Y. P. (2021). Tracing the Genetic Legacy of the Tibetan Empire in the Balti. *Molecular Biology and Evolution*, *38*(4), 1529–1536. https://doi.org/10.1093/molbev/msaa313

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Xi Ping Cuo, Z., Pool, J. E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T. S., Zheng, H., Liu, T., He, W., Li, K., Luo, R., Nie, X., Wu, H., Zhao, M., Cao, H., … Wang, J. (2010). Sequencing of Fifty Human Exomes Reveals Adaptations to High Altitude. *Science*, *329*(5987), 75–78. https://doi.org/10.1126/science.1190371.Sequencing

Zhang, B. li, Xu, R. liang, Zhang, J., Zhao, X. xian, Wu, H., Ma, L. ping, Hu, J. qiang, Zhang, J. liang, Ye, Z., Zheng, X., & Qin, Y. wen. (2013). Identification and functional analysis of a novel PRKAG2 mutation responsible for Chinese PRKAG2 cardiac syndrome reveal an important role of non-CBS domains in regulating the AMPK pathway. *Journal of Cardiology*, *62*(4), 241–248. https://doi.org/10.1016/j.jjcc.2013.04.010

Zhang, C., Lu, Y., Feng, Q., Wang, X., Lou, H., Liu, J., Ning, Z., Yuan, K., Wang, Y., Zhou, Y., Deng, L., Liu, L., Yang, Y., Li, S., Ma, L., Zhang, Z., Jin, L., Su, B., Kang, L., & Xu, S. (2017). Differentiated demographic histories and local adaptations between Sherpas and Tibetans. *Genome Biology*, *18*(1). https://doi.org/10.1186/s13059-017-1242-y

Zhang, J., Xu, J., Dong, Y., & Huang, B. (2018). Down-regulation of HIF-1α inhibits the proliferation, migration, and invasion of gastric cancer by inhibiting PI3K/AKT pathway and VEGF expression. *Bioscience Reports*, *38*(6). https://doi.org/10.1042/BSR20180741

Zhang, X., Kim, B., Singh, A., Sankararaman, S., Durvasula, A., & Lohmueller, K. E. (2023). MaLAdapt Reveals Novel Targets of Adaptive Introgression From Neanderthals and Denisovans in Worldwide Human Populations. *Molecular Biology and Evolution*, *40*(1). https://doi.org/10.1093/molbev/msad001

Zhang, X., Witt, K. E., Bañuelos, M. M., Ko, A., Yuan, K., Xu, S., Nielsen, R., & Huerta-Sanchez, E. (2021). The history and evolution of the Denisovan-EPAS1 haplotype in Tibetans. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(22), 1–9. https://doi.org/10.1073/pnas.2020803118

Zheng, W., He, Y., Guo, Y., Yue, T., Zhang, H., Li, J., Zhou, B., Zeng, X., Li, L., Wang, B., Cao, J., Chen, L., Li, C., Li, H., Cui, C., Bai, C., Baimakangzhuo, Qi, X., Ouzhuluobu, & Su, B. (2023). Large-scale genome sequencing redefines the genetic footprints of high-altitude adaptation in Tibetans. *Genome Biology*, *24*(1). https://doi.org/10.1186/s13059-023-02912-1

Figures

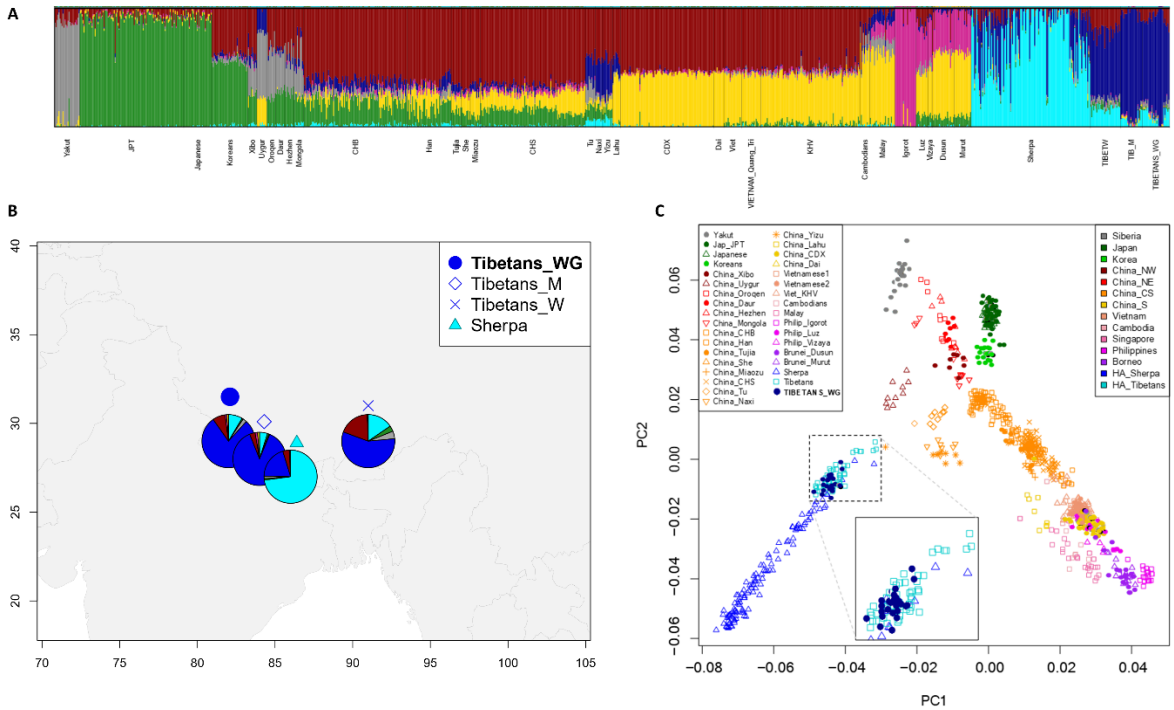

Figure 1. Population structure analyses performed on the extended dataset including Tibetan, Sherpa, and lowland East-Asian individuals. (A) Admixture analysis showed the best predictive accuracy when seven (K=7) population clusters were tested. Populations included in the dataset are labelled according to population names and acronyms reported in Supplement table 1. **(B)** Map showing geographic location and admixture proportions at K=7 of the high-altitude groups included in the extended dataset. The label *Tibetans_WG* indicates whole genome sequence data for individuals of Tibetan ancestry analysed in the present study. Additional information about the considered samples (e.g., number of individuals per group, reference study, and used abbreviations) are reported in Supplement table 1. **(C)** PCA plot considering PC1 vs PC2 and summarizing genomic divergence between high-altitude Tibetan/Sherpa people and the cline of variation observable for lowland East-Asian populations. The enlarged square displays clustering between Tibetan samples sequenced for the whole genome (i.e., blue dots) and Tibetan samples characterized by genome-wide data (i.e., light-blue squares).

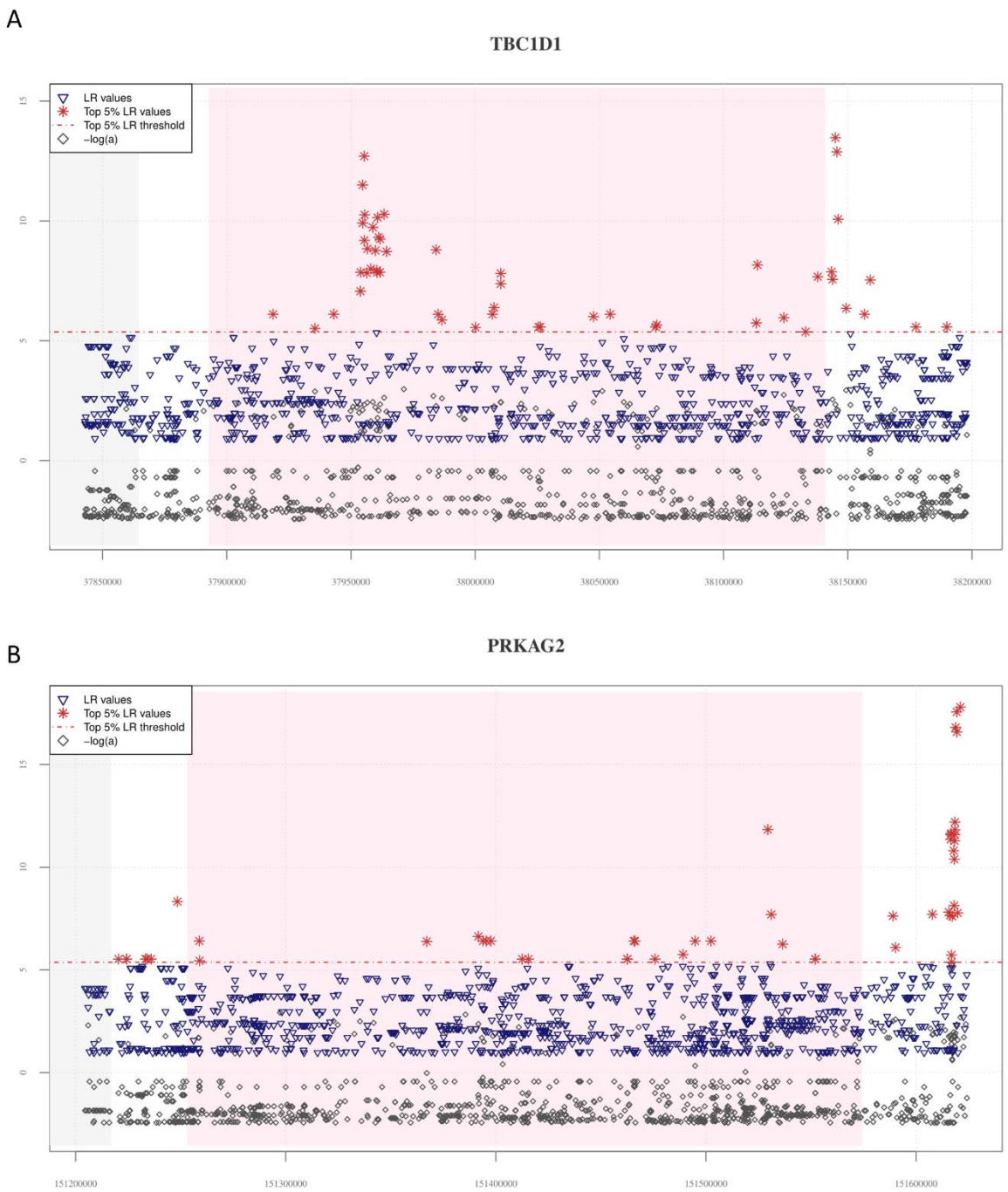

Figure 2. Distribution of *VolcanoFinder* statistics suggestive of putative adaptive introgressed loci across the *TBC1D1* and *PRKAG2* genomic regions. On the x-axis are reported genomic positions of each SNV, while on the y axis are displayed the related statistics obtained. Pink background indicates the chromosomal interval occupied by the considered genes, while the grey background identifies those genes (i.e., *PGM2* in the *TBC1D1* downstream genomic region and the *RHEB* gene in the upstream *PRKAG2* region) possibly involved in regulatory transcription mechanisms. The dashed red line identifies the threshold set to filter for significant LR values (i.e., top 5% of LR values). For both these genomic regions the distribution of LR and -logα are concordant with those observed at the *EPAS1*

positive control for AI. **(A)** A total of 50 significant LR values (red stars) and -logα (grey diamonds) values resulted collectively elevated in both the *TBC1D1* gene and its downstream genomic regions. A remarkable concentration of significant LR values characterizing 19 SNVs was especially observable in the first portion of the gene. **(B)** The entire *PRKAG2* genomic region was found to comprise 46 SNVs showing significant LR values, with the greatest peaks being located in the downstream region associated to such gene. Peaks detected for the LR statistic are accompanied by peaks of -logα values.

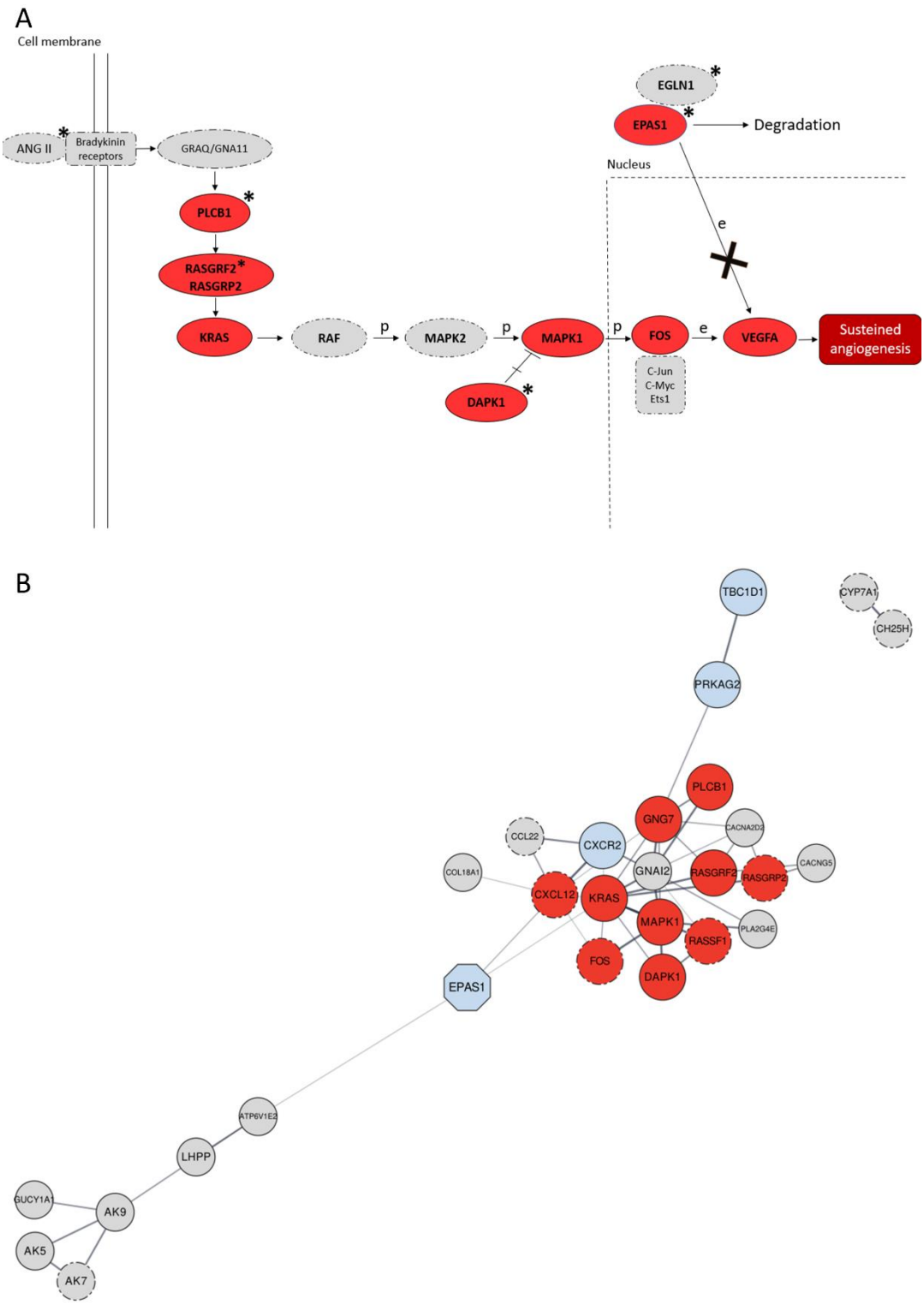

Figure 3. Significant gene networks including Denisovan-like derived alleles according to the *Signet* analysis. (A) Schematic representation of the activation of the *RAS*/*MAPK*(*ERK*) axis after interaction of the bradykinin receptors with their ligands (e.g., ANG II) within the framework of the *Pathways in Cancer* network. Genes supported by both *Signet* (i.e. belonging to the significant network associated to *Pathways in cancer*) and *VolcanoFinder* (i.e., including at least a SNV showing LR value within top 5% of the obtained results) analyses as potentially introgressed loci, are highlighted in red and present solid outline. Grey circles with dotted-dashed contour instead indicate genes supported only

by *Signet*, while loci marked with stars are those including genomic windows showing *LASSI* T statistic within top 5% of the related distribution.

After the interaction between ANGII (active enzyme angiotensin II) and bradykinin receptors, activation of the Ras protein encoded by *KRAS* mediated by RAS-GTPases (e.g., *RASGRF2*) comports a series of phosphorylation reactions that eventually promotes angiogenesis (Kranenburg et al., 2004). In detail, phosphorylation of the MAPK1 protein and prevention of MAPK1-DAPK-1-dependent apoptosis leads to increased MAPK1 activity (Kanehisa & Goto, 2000; Stevens et al., 2007) that causes improved *FOS* mRNA expression (Monje et al., 2005). *FOS* together with other proteins (e.g., Jun) forms the AP-1 transcription factor, which bounds to the *VEGF* promoter region upregulating its expression in endothelial cells (Catar et al., 2013) and sustaining angiogenesis when the HIF-1 signalling cascade is inhibited (Lorenzo et al., 2014). **(B)** Gene network built by setting co-expression as force function and by displaying the entire set of genes identified by the *Signet* algorithm as belonging to significant pathways including Denisovan-like derived alleles. Genes whose variation pattern was supported by both *VolcanoFinder* and *Signet* analyses (e.g., *TBC1D1*) as shaped by archaic introgression are displayed with a solid black outline. The *EPAS1* positive control locus that has been previously proved to have mediated adaptive introgression in Tibetan populations was represented as light-blue octagonal. Genes included in pathways involved in angiogenesis (e.g., *RASGRF2*) and/or activated in hypoxic conditions (e.g., *PRKAG2*) are reported as dark red and light-blue circles respectively, while the remining fraction of significant genes are represented as light-grey circles. The closeness or the distance between all nodes reflects the tendency to be co-expressed with each other and all the connections inferred are characterized by a confident score ≥ 0.7.

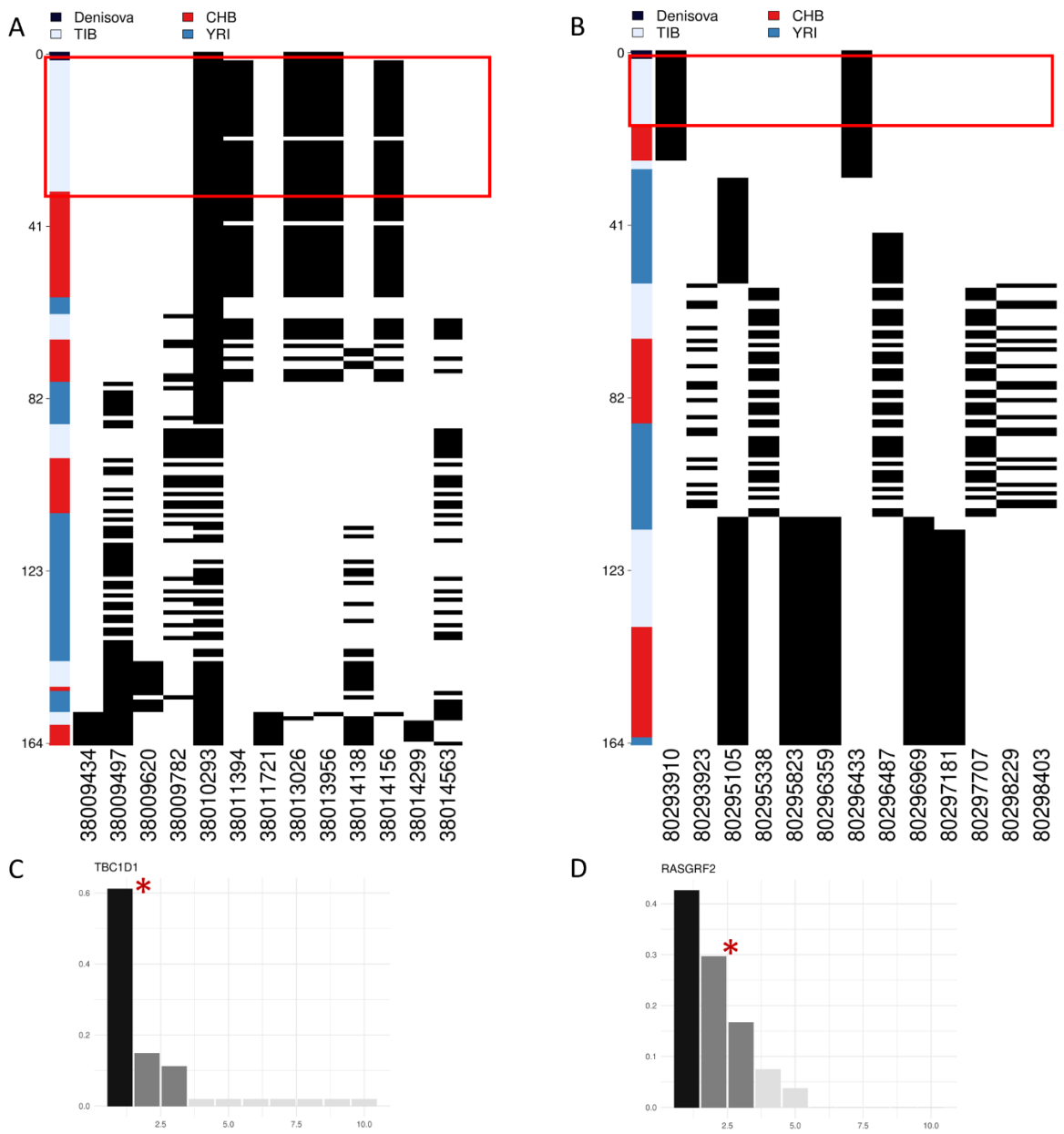

Figure 4 – Representation of genetic distances between modern and archaic haplotypes and barplots showing haplotype frequency spectra for *TBC1D1* and *RASGRF2* candidate AI genes. Haplotypes are reported in rows, while derived (i.e., black square) and ancestral (i.e., white square) alleles are displayed in columns. Haplotypes are ranked from top to bottom according to their number of pairwise differences with respect to the Denisovan sequence. **(A)** Heatmap displaying divergence between Tibetan, CHB and YRI *TBC1D1* haplotypes with respect to the Denisovan genome. A total of 33 *TBC1D1* haplotypes (i.e., 61% of the overall haplotypes inferred for such a region) belonging to individuals with Tibetan ancestry are plotted in the upper part of the heatmap thus presenting the smallest number of pairwise differences with respect to the Denisovan sequence. **(B)** Heatmap displaying divergence between Tibetan, CHB and YRI *RASGRF2* haplotypes with respect to the

Denisovan genome. A total of 16 Tibetan haplotypes in the *RASGRF2* genomic region present no differences with respect to the Denisovan sequence.

As regards barplots, on the x axis are reported the haplotypes detected in the considered genomic windows, while on the y axis is indicated the frequency for each haplotype. The black and dark-grey bars indicate the more frequent haplotypes (i.e., the putative adaptive haplotypes inferred by the *LASSI* method), while red stars mark those haplotypes carrying Denisovan-like derived alleles. **(C)** *TBC1D1* haplotype frequency spectrum. The *TBC1D1* gene presents a haplotype pattern qualitatively comparable to that observed at *EPAS1* (Figure 4 – figure supplement 3A), with a predominant haplotype carrying archaic derived alleles and reaching elevated frequencies in Tibetan populations. In line with this observation, such a pattern was inferred by *LASSI* as conformed with a non-neutral evolutionary scenario, even if it seems to be characterized by a soft rather than a hard selective sweep due to the occurrence of three sweeping haplotypes. **(D)** *RASGRF2* haplotype frequency spectrum. A soft selective sweep was inferred also for the considered *RASGRF2* genomic window, although frequencies reached by the sweeping haplotypes turned out to be more similar with each other. The second most represented haplotype was that carrying the archaic derived alleles and, reached a frequency of 29% in the Tibetan group.