

# Multiple selective events at the *PRDM16* functional pathway shaped adaptation of western European populations to different climate conditions

Andrea Quagliariello<sup>1,2,3</sup>, Sara De Fanti<sup>1,2</sup>, Cristina Giuliani<sup>1,2</sup>, Paolo Abondio<sup>1</sup>, Patrizia Serventi<sup>1</sup>, Stefania Sarno<sup>1,2</sup>, Marco Sazzini<sup>1,2\*</sup> & Donata Luiselli<sup>1,2\*</sup>

1) Laboratory of Molecular Anthropology, Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, 40126, Italy

e-mail: andrea.quagliariello@opbg.net

2) Centre for Genome Biology, Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, 40126, Italy

3) Human Microbiome Unit, Bambino Gesù Children's Hospital and Research Institute, IRCCS, Rome, Italy

**Summary** - Several studies highlighted the role of climate in shaping many human evolutionary processes. This occurred even in relatively recent times, having affected various human phenotypic traits, among which metabolic processes that orchestrate absorption and accumulation of substances to maintain energy homeostasis, that is critical for the survival of individuals in high energy-expenditure environments. To date, most researches have focalized on detection of climatic influence on SNPs frequency in populations exposed to extreme environmental conditions or by comparing variation patterns between populations from different continents. In this study, we instead explored the genetic background of distinct western European human groups at loci involved in nutritional and thermoregulation processes, to test whether patterns of differential local adaptation to environmental conditions could be appreciated also at a lower geographical scale. Taking advantage from the 1000 Genomes Project data, genetic information for 21 genes involved in nutritional and thermoregulation processes was analysed for three western European populations. The applied Anthropological Genetics methods pointed to appreciable differentiation between the examined groups especially for the *PRDM16* gene. Moreover, several neutrality tests suggested that balancing selection has acted on different regions of the gene in people from Great Britain, as well as that more recent positive selection could have also targeted some *PRDM16* SNPs in Finn and Italian populations. These series of adaptive footprints are plausibly related to climate variability in both ancient and relatively recent times. Since this locus is involved in thermoregulation mechanisms and adipogenesis, local adaptations mediated by a pathway related to the brown adipose tissue activity could have evolved in response to changing cold temperature exposures of such populations.

**Keywords** - Natural selection, Thermoregulation, Brown adipose tissue, *PRDM16*.

## Introduction

As consequence of the Out of Africa migration, anatomically modern humans developed the ability to face considerably different environmental conditions, characterized by a wide range

of nutritional and climatic landscapes, by evolving new physiological traits and/or by adopting innovative cultural strategies.

In particular, following the end of the last ice age, approximately 14,000 years ago, a major warming event raised global temperatures

\* Contributed equally to this work

roughly to their current values (Wright, 1993), while less intense dramatic climatic changes and meteorological events have been responsible for temperatures fluctuations in the subsequent history of human societies (Zhang *et al.*, 2007).

Accordingly, recent studies have investigated the most relevant phenotypes, as well as the underlying genetic differences between continental groups of populations, with the aim of highlighting the influence of climate on broad processes of human adaptation (Hancock *et al.*, 2011b; Turchin *et al.*, 2012; Raj *et al.*, 2013). The achieved results demonstrated that humans are extremely vulnerable to long-term temperature changes and that climate has actually impacted on several human traits, such as body size, skin pigmentation, energy expenditure, nutrients metabolism, etc. (Sturm, 2009; Jablonski & Chaplin, 2010; Sazzini *et al.*, 2014). Traces of these adaptive events have been detected for instance by the analysis of single nucleotide polymorphisms (SNPs) distribution, pointing to several SNPs at genes involved in these biological processes whose frequency patterns reflect a latitudinal gradient among various human populations (Hancock *et al.*, 2010a; Hancock *et al.*, 2011b). These genetic differences due to environmental adaptations influenced physiological processes especially linked to metabolism and thermogenesis, as highlighted by studies showing that adaptation to local temperature conditions deeply influenced energy expenditure, metabolism rate and susceptibility to several metabolic disorders (Hancock *et al.*, 2008; Yamane *et al.*, 2010; Hancock *et al.*, 2011a; Brown, 2012; Sazzini *et al.*, 2016).

Furthermore, the evolutionary processes that shaped the development of new metabolic phenotypes in response to specific environments may be driven not only by the necessity to accurately regulate the body heat production (i.e. thermogenesis), but also by cultural (e.g. dietary) selective pressures. In fact humans are able to considerably alter their own ecosystem to improve their subsistence by directly shaping their ecological niche and influencing their own evolutionary history through cultural innovations (Cordain *et al.*,

2005; Laland & Brown, 2006; Ruddiman, 2013; Lewis & Maslin, 2015). For instance, during human evolutionary history inside and outside Africa, several shifts in dietary habits have occurred accordingly to changing local ecosystems and resources availability. In particular, transition from a hunter-gatherer lifestyle to an agricultural one represented one of the most challenging dietary shifts in recent human evolution, leading to high relevance of cereals and tubers and poor iron and protein intake in our diets (Cordain *et al.*, 2002; Mathieson *et al.*, 2015).

Up to date, most part of the studies on human adaptation to diet, subsistence techniques and climate conditions have focused on broad divergence between populations adapted to extremely different environments (Voight *et al.*, 2006; Lopez Herraez *et al.*, 2009; Hancock *et al.*, 2010a; Hancock *et al.*, 2010b). On the contrary, the present case study aims at investigating patterns of genetic variation at a lower geographical scale (i.e. within the same continent) to test whether moderate differences in climate and dietary habits could have appreciably shaped the evolution of relatively close human groups. For this purpose, western European populations constitute interesting case studies according to their relatively shared ancestry and to their concomitant exposure to different bio-climatic zones, ranging from Mediterranean to Boreal environments, and thus characterized by considerable divergences in temperatures and nutritional resources (Rivas-Martinez *et al.*, 2006; Skoglund *et al.*, 2012; Lazaridis *et al.*, 2014; Allentoft *et al.*, 2015).

That being so, this approach enabled us to dissect the potential legacy of natural selection on these populations, shedding light on some of the evolutionary mechanisms that could have favoured their dispersal into the European continent.

## Methods

### *Data retrieval*

To perform population genetics analyses, genotype data for 21 genes pertaining to several metabolic and thermogenic pathways were

retrieved from the 1,000 Genomes Project phase I database (Altshuler *et al.*, 2012) (<ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/phase1/>). More in detail, the obtained dataset was composed of genotypes of 280 individuals belonging to three European populations characterized by limited genetic admixture and representative of the western European genetic variability observable at different latitudes (Tuscans from Italy, TSI; British from England and Scotland, GBR; Finnish, FIN).

#### *Population genetics analyses*

To explore patterns of genetic variation among the selected populations, the PLINK package v.1.07 (Purcell *et al.*, 2007) was used to prune SNPs in Linkage Disequilibrium (LD) and to prevent potential bias on the applied multivariate analyses. By proceeding with a 50 SNPs-sliding windows approach and by shifting windows of 10 SNPs every time, LD values have been calculated between each pair of SNPs and one variant per pair with a pairwise genotypic correlation ( $r^2$ ) higher than 0.1 was removed.

The obtained pruned dataset was then used to perform the Discriminant analysis of principal components (DAPC) (Jombart *et al.*, 2010), which is particularly informative to explore genetic relationships among pre-defined groups of individuals, using the R *adegenet* package and in order to explore patterns of population structure.

#### *Descriptive and site frequency spectrum-based neutrality statistics*

Nucleotide diversity, estimated as the average number of pairwise difference ( $\pi$ ), the number of polymorphic site ( $S$ ), as well as Tajima's  $D$  ( $D$ ), Fu and Li's  $D$  and  $F$  ( $D$ ,  $F$ ) statistics, were calculated using the DnaSP package v.5.10 (Librado & Rozas, 2009) on the whole *PRDM16* gene that turned out to be the best candidate locus among those investigated after population structure analyses. These statistics were computed for each population by applying a sliding windows approach (i.e. each window consisted of 10 kb and was progressively shifted of 10 kb). To test the significance of the obtained results, 10,000

coalescent simulations were performed by considering local recombination and mutation rates and under the assumption of a neutral model of evolution. Then, nominal  $p$ -values were corrected with Bonferroni procedure using the R package *multtest* in order to account for the adopted multiple testing procedure and by considering a significance threshold of  $\alpha = 0.01$ .

#### *Haplotype and Integrated haplotype score (iHS) analyses*

Focusing on the gene windows showing significant  $p$ -values for the computed neutrality tests, pairwise LD for each SNP within each window was performed with the PLINK package v.1.07 (Purcell *et al.*, 2007). SNPs in high LD ( $r^2 > 0.90$ ) were then used to infer haplotypes from unphased genotypes using the PHASE software v.2.1 (Scheet & Stephens, 2006), which implements a Bayesian algorithm for haplotype reconstruction from population genotype data. Evolutionary relationships among the inferred haplotypes were then explored by drawing a median joining network with the Network package v.4.6.1.2 (<http://www.fluxus-engineering.com>).

To further investigate genetic signatures related to the action of potential selective pressures on the examined populations, statistical methods based on the evaluation of extended haplotype homozygosity (EHH) were applied. The *REHH* R package was used to estimate the relative extended haplotype homozygosity compared with the EHH of all other haplotypes on the gene and its progressive decay, defined as integrated EHH (iHH). The lengths of the haplotypes extending around core SNPs are deducible from the decay of EHH values until they drop below 0.2. Therefore, it is possible to calculate the haplotypes maximum dimension taking into account the distance between the core SNP and the first point to the left and to the right with an EHH  $> 0.2$ . On the basis of the decay level, iHS scores were then calculated as the log-ratio of iHH computed at the derived and ancestral alleles (Voight *et al.*, 2006).

The iHS statistic was finally obtained for each SNP using an homozygosity threshold of

**Tab. 1 - List of the examined genes.**

GENE	CHR*	FUNCTION
UCP1	4	uncoupling protein 1
UCP2	11	uncoupling protein 2
UCP3	11	uncoupling protein 3
ADRB3	8	adrenoceptor beta 3
DIO2	14	deiodinase, iodothyronine, type 2
PRDM16	1	PR domain containing 16
ADRA1A	8	adrenoceptor alpha 1A
FTO	16	fat mass and obesity associated
PLINK5	19	perilipin 5
INSIG2	2	insulin induced gene 2
APOA5	11	apolipoprotein A-V
MC4R	18	melanocortin 4 receptor
GNB3	12	guanine nucleotide binding protein (G protein), beta polypeptide 3
VDR	12	vitamin D
MTRR	5	5-methyltetrahydrofolate-homocystein methyltransferase reductase
PLRP2	10	pancreatic lipase-related protein 2
NAT2	8	N-acetyltransferase 2
PPARG	3	peroxisome proliferator-activated receptor gamma
TCF7L2	10	transcription factor 7-like 2 (T-cell specific, HMG-box)
MMP20	11	matrix metalloproteinase 20
ENAM	4	enamelin

\* Chromosome

After a first literature survey, 21 genes have been selected for the analysis according to their involvement in nutritional or thermoregulation processes.

EHH > 0.05 and allele frequency bins of 0.025 to standardize iHS scores. SNPs showing |iHS| values > 2 were considered as possible candidates loci to have undergone recent positive selection.

Bonferroni correction was then applied to account for the adopted multiple testing procedures. Several graphical tools of the REHH package were also exploited to better visualize linkage breakdown. Finally, to obtain a rough estimate of the age of selective sweeps detected by iHS analysis, the algorithm developed by Voight *et al.* (2006) was applied to each SNP showing significant signals of positive selection and taking advantage from the Li and Stephens' algorithm to evaluate recombination rates (Li & Stephens, 2003).

## Results

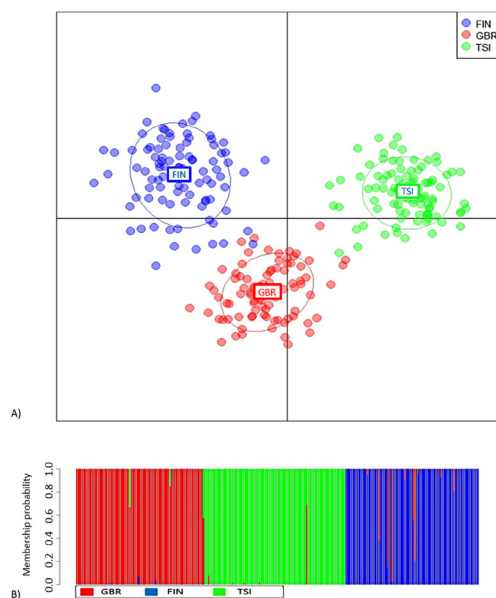
In order to evaluate the impact of climate and the potentially related dietary shifts on recent evolution of three western European populations, the present study explored variation at 21 genes selected by means of an extensive literature and public databases survey according to their involvement in various thermoregulation and nutritional processes (Tab.1).

Patterns of population structure were investigated by means of DAPC, by first considering the whole dataset, and then via a gene-by-gene approach. Appreciable intra-continental variability was thus observed, with Northern European populations showing slightly closer genetic affinities with each other than with the examined Southern European group (Supplementary Fig. 1, online). To dig up whether the observed structure was due mainly to one or more genes, the same analysis was repeated separately for each locus. While no population structure was found as regards almost all the explored genes, *PRDM16* turned out to be characterized by considerable heterogeneity among the studied populations (Fig. 1a,b).

That being so, subsequent analyses focused on the 2,720 SNPs annotated on the 369 kb genomic interval covering the *PRDM16* locus (chr1: 2,985,732-3,355,185; GRCH37), to test whether the observed pattern of population structure could be influenced or not by the action of natural selection at such a gene.

Different neutrality statistics were computed to explicitly test this hypothesis. First of all, site frequency spectrum-based neutrality tests were calculated by applying a sliding window approach and by considering genomic windows of 10 kb. Several windows showing significant results for the Fu & Li's D and F tests were identified after comparison of the calculated statistics with distributions obtained by 10,000 coalescent simulations, as reported in Figure 2 and on Supplementary Table 1 (online). In particular, ten windows showed significant values for the Fu & Li's D test and one for the Fu & Li's F test. To shortlist the most plausible candidate windows, Bonferroni correction was applied pointing to only four highly significant genomic intervals: two for Finn (FIN) and one for each of the other considered populations (i.e. people from Great Britain, GBR, and Tuscans from Italy, TSI). All these windows showed largely positive values ( $\geq 2$ ) indicating that such signatures could be related to the action of balancing selection on the investigated genomic regions.

To deepen investigation of the *PRDM16* evolutionary history, as well as to provide possible additional evidence supporting the hypothesis that balancing selection has acted on it, haplotypes were inferred by focusing exclusively on SNPs included in the most promising identified candidate window. In particular, haplotype structure related to the most reliable candidate genomic interval (330,007-340,006), which showed significant results for both Fu & Li' tests in GBR, was investigated also according to its high functional relevance. In fact, this chromosomal segment encodes for a Zinc finger 1 domain essential for the functionality of the *PRDM16* protein. LD analysis focused on a genomic interval including 100 kb downstream and upstream regions with respect to such a window was performed to exclude hitchhiking effects due to the presence of adaptive variants on the *PRDM16* nearby genes. Accordingly, 15 SNPs in high LD ( $r^2 \geq 0.90$ ) with the considered variants and belonging to the described extended chromosomal region were used to reconstruct haplotypes. A median joining network was

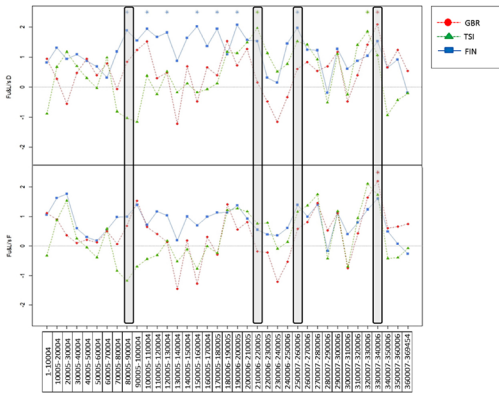


**Fig. 1 - Population structure analysis. a) DAPC analysis shows a clear discrimination among the three examined European populations. Populations are indicated by different colors and ellipses. b) Membership probabilities of each individual to belong to the tested group indicate high inter-group homogeneity. The colour version of this figure is available at the JASs website.**

then used to explore evolutionary relationships among the inferred haplotypes, pointing to the distinction of two highly divergent haplotype clades. The former was characterized by a main common haplotype carrying predominantly ancestral alleles, while the latter presented three frequent haplotypes mostly made up of derived alleles (Fig.3a). Looking at the haplotype frequencies (Fig.3b), GBR showed similar proportions for the two most common “ancestral” and “derived” haplotypes, while TSI and FIN exhibited a major abundance of haplotypes belonging to the “ancestral” and “derived” clades, respectively. This suggests that balancing selection have actually acted on this region exclusively in the GBR population.

In addition to this, also a potential more recent action of natural selection on *PRDM16* was tested by computing integrated haplotype





**Fig. 2 - Neutrality Tests. Results for the Fu & Li' D and F statistics. Each line indicates a different population: FIN (blue), GBR (red), and TSI (green). On the x-axis are reported the 37 10 kb windows in which the PRDM16 gene was divided. Asterisks indicate significant results obtained after 10,000 coalescent simulations and their color is relative to the population for which the significant result was obtained. The black columns highlight windows that remained significant after Bonferroni correction for multiple testing. The colour version of this figure is available at the JASs website.**

scores (iHS) for each SNP. Accordingly, several loci have been identified as potential targets of positive selection in the studied populations (Supplementary Fig.2, online). In particular, two limited genomic intervals within *PRDM16* showed a series of consecutive SNPs characterized by high |iHS| values. A first one spanned from 3,082,126 to 3,084,761 nucleotide positions, while the second one covered the 3,129,182-3,147,970 region, with the former including most of the top 1% and top 90% (thus concerning both the highest positive and negative values) of iHS significant SNPs observed for all populations. When a more stringent significance threshold was considered by applying Bonferroni correction for multiple testing, a very small number of SNPs still showed statistically significant iHS values, being limited to the FIN (rs112682827 and rs144090205) and TSI (rs2817126) populations (Supplementary Fig.3, Tab. 2, online).

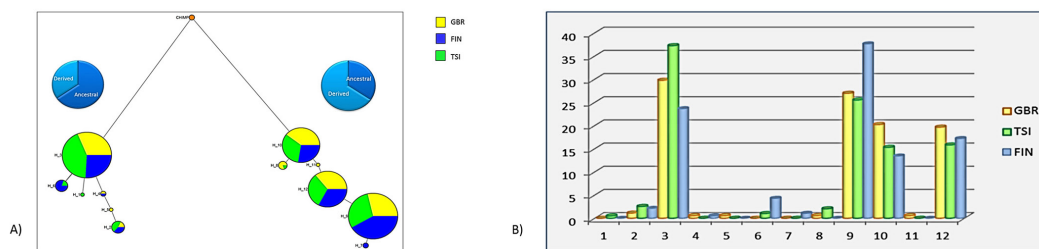
Both rs2817126 and rs112682827 showed highly positive iHS values (iHS = 4.07 in TSI

and iHS = 4.12 in FIN), suggesting positive selection having potentially acted on their ancestral alleles, while rs144090205 was characterized by a strong negative iHS score (iHS = -3.85 in FIN) indicating that selection probably acted on the derived allele of this SNP (Tab. 3). Extension of haplotype homozygosity (EHH) around these core SNPs was also found to be appreciably larger in TSI than FIN (Supplementary Fig.4, online), indicating that positive selection has plausibly acted on the former more recently than on the latter. Branches thickness displayed in the calculated bifurcation plots finally proved that the haplotype carrying the rs144090205 derived allele, as well as those carrying the ancestral alleles at rs2817126 and, especially, at rs112682827, are shared by the majority of individuals constituting the respective populations, thus providing a further evidence for their supposed adaptive role (Supplementary Fig.6, online).

To get further insights into the possible causes responsible for these multiple signatures of recent positive selection on FIN and TSI, age estimation for the detected selective sweeps was performed as described in Voight *et al.* (2006). In particular, average genetic distance between each core SNP and the first nucleotide position to the left and to the right whose EHH value dropped below 0.20 was computed. Within this range, values of 0.39 cM and 1.15 cM were obtained for rs112682827 and rs144090205, respectively, as regards the FIN group, while a value of 1.97 cM was observed for rs2817126 in TSI. Although it is hard to obtain rigorous age estimations, assuming a generation time of 25 years rough ages of approximately 3,742 years, 1,429 years and 1,092 years could be inferred for the putative selective events occurred on rs112682827 (FIN), rs144090205 (FIN) and rs2817126 (TSI), respectively.

## Discussion

After the *H. sapiens* exit from Africa, populations of anatomically modern humans have been subjected to a range of evolutionary shifts



**Fig. 3 - Median Joining Network.** a) Network of haplotypes carrying SNPs in high LD ( $r^2 \geq 0.90$ ) with variants included in the 330,007-340,006 window significant in GBR. The blue pies represent the percentage of derived and ancestral alleles at the two clades. b) Haplotype distribution among populations. The colour version of this figure is available at the JASs website.

in terms of occupied habitats, exploited food resources, experienced climate conditions, as well as demography and pathogen exposure. These events have certainly left sizeable marks on the human genome, as suggested by an increasing number of studies (Voight *et al.*, 2006; Barreiro *et al.*, 2008; Hernandez *et al.*, 2011; Stewart & Stringer, 2012). Nonetheless, the actual impact of natural selection on the processes especially related to human nutrition and energy expenditure has been to date poorly understood (Heyer & Quintana-Murci, 2009; Ye & Gu, 2011; Hancock *et al.*, 2011b; Breslin, 2013).

This study was thus aimed at detecting possible signatures in the gene pool of western European populations that could be associated to the above-mentioned evolutionary events, in the attempt to understand how relatively recent ecological and cultural shifts related to nutritional habits and environmental conditions could have shaped current patterns of genetic variation within the European continent.

For this purpose, 21 genes differently implicated in several nutritional and thermoregulation processes have been selected and related variation data have been retrieved for three western European non-admixed populations from different latitudes sequenced by the *1000Genomes Project*.

Evidence of overall intra-continental variability associated to this subset of genes was provided by results of DAPC applied on the whole dataset. However, when DAPC was repeated gene-by-gene, extremely reduced population structure was

pointed out for almost all loci except *PRDM16*, which instead showed a considerable level of differentiation between the examined populations, with Northern European groups (FIN and GBR) showing closer genetic affinities with each other than with the Southern European one (TSI). The *PRDM16* gene encodes for a 140 kDa protein characterized by a SET-like positive regulatory (PR) domain and 10 zinc finger sequences, which modulate the transcriptional activity of the brown fat (BAT) thermogenic genes programme. *PRDM16* plays also a major role in the mechanisms of transcription and regulation of adipocyte differentiation (Frühbeck *et al.*, 2009). In fact, it could be definitely considered a master regulator of the processes of BAT adipogenesis, differentiation and function, as well as of the so-called “browning” process of white adipose tissue (WAT). In the last years, a large number of studies have indeed confuted the old binomial view concerning the existence of only two adipose tissues, distinct from both histological and functional characteristics, by the discovery of a new kind of tissue named “beige” or “brite” (Wu *et al.*, 2012). This type of adipocytes constitutes a link between BAT and WAT adipose tissues, although its role in the physiological metabolism is still far to be elucidated. *PRDM16* thus exerted its function at the crossroads between the physiological processes involved in adaptation to climate and those regulating energy metabolism. In fact, it is mainly expressed in both BAT and beige tissues and controls the cell fate switch

Tab. 2 - Summary statistics for iHS.

POP	NUMBER OF SNPS WITH  IHS >2.0	MAX  IHS	LOWEST IHS P-VALUE	BON. THR.*	SIGN. SNPS**
GBR	42	3.75	0.00015	0.00011	0
TSI	32	4.07	0.00004	0.00009	1
FIN	40	4.12	0.00003	0.00013	2

\* Bonferroni Threshold

\*\* Significant Snps

Summary statistics for the iHS test performed on the analysed populations.

between skeletal myoblast and brown fat cells, also enhancing thermoregulation. Brown and beige adipose tissues share many biochemical characteristics and both are activated in response to cold exposure and to modulate energy balance (Ouellet *et al.*, 2012). Several studies indeed highlighted that climate changes could affect BAT activity. For instance, reduced seasonal cold exposure could influence energy balance and obesity risk (Johnson *et al.*, 2011), as well as BAT amount around arteries to assure survival at extreme cold temperatures (Huttunen *et al.*, 1981). Furthermore, *PRDM16* serves as positive regulator of the whole pathway expressed within the BAT and beige tissues, by mediating environment-induced beige adipocyte development and thus testifying its possible major role in evolutionary processes underlying human adaptation to climate (Ohno *et al.*, 2012, 2013; Cohen *et al.*, 2014).

Although the observed pattern of population structure at *PRDM16* could basically result from the demographic histories of such populations, absence of appreciable differentiation at the other surveyed loci suggests that natural selection might have substantially contributed to shape variation at *PRDM16* within western Europe. To corroborate this hypothesis, several neutrality tests have been applied on full sequence data from the whole *PRDM16* gene. After Bonferroni correction for multiple testing,

four genomic windows showing significant statistics were identified as concerns the Fu & Li' D test (two for FIN, one for GBR and one for TSI) and one (the same previously observed for GBR) as concerns also the Fu & Li' F test. For all these windows, largely positive values were obtained indicating the possibility that balancing selection has acted on these genomic regions. Moreover, since they included several intronic SNPs, as well as variants related to non-coding transcripts, it could be hypothesized that balancing selection could have acted mainly on *PRDM16* regulatory loci. Among these chromosomal intervals, that showing significant results for both Fu & Li' tests was pinpointed as the most promising candidate window, suggesting that balancing selection could have targeted more plausibly the GBR population. This hypothesis is supported also by the haplotype structure observed at such *PRDM16* region. In fact, two considerably divergent haplotype clades were inferred, one including a main common haplotype carrying predominantly ancestral alleles and another presenting frequent haplotypes mostly made up of derived alleles. Among the investigated populations, only GBR actually showed similar proportions for the two most common "ancestral" and "derived" haplotypes, as expected under evolution shaped by balancing selection. Interestingly, the candidate region putatively subjected to selection in GBR encodes for a DNA binding Zinc finger 1 domain of the *PRDM16* protein, corresponding to amino acids 224-454 (Kajimura *et al.*, 2008), implicated in activation of the gene thermogenic function by specifically binding the *PGC-1 $\alpha$*  and the *PGC-1 $\beta$*  loci and thus increasing their transcriptional activity (Puigserver *et al.*, 1998). In fact, these two proteins regulate both the development of brown fat cells identity (Seale *et al.*, 2007, 2008) and the thermogenic gene programme of BAT adipocytes by promoting BAT-specific gene induction (Hondares *et al.*, 2011). Through the activation of *PGC-1 $\alpha$* , which then co-activates *PPAR $\alpha$*  and *PPAR $\gamma$* , *PRDM16* actually regulates both mitochondrial biogenesis and thermogenesis processes (Barbera *et al.*, 2001). According to this view, it is highly plausible that



**Tab. 3 - Candidate SNPs under positive selection.**

RS	POSITION	POP	A	AA	P-VALUE	EHH(BP)
rs2817126	3166682	TSI	G	T	0.000043	132596
rs112682827	3165673	FIN	T	T	0.000036	33508
rs144090205	3165687	FIN	T	T	0.00011	40186

Candidate SNPs that have plausibly undergone positive selection **POP**= Population; **A**= selected allele; **AA**= Ancestral Allele; **P-value**= p-value after Bonferroni correction for multiple testing; **EHH**= dimension of the region showing extended haplotype homozygosity and in which the core SNPs are located.

selective pressures responsible for the described signature of balancing selection in GBR were related to climate-driven adaptive scenarios.

Further information about the evolutionary processes that could have shaped the observed *PRDM16* genetic patterns were provided also by testing for a potential more recent action of natural selection on the examined populations through the application of the iHS test. Accordingly, two SNPs for FIN (rs112682827 and rs144090205) and one for TSI (rs2817126) were pointed out as plausible targets of recent positive selection. Ancestral alleles at both rs2817126 and rs112682827 seem to have been subjected to positive selection respectively in TSI and FIN, while the derived allele of rs144090205 could have exerted an adaptive role solely in FIN. Interestingly, these SNPs are distributed in a restricted genomic range spanning around 2.6 kb and in which many other variants showed high absolute, even if not significant after Bonferroni correction, iHS values. This could indicate that probably the entire genomic interval has been subjected to recent positive selection in the considered European populations. Moreover, these SNPs did not show any appreciable LD with each other, and are all intronic substitutions related to a non-coding transcript variant (ENST00000607632). In particular, the two potentially adaptive SNPs observed in FIN lie in a *PRDM16* regulatory region associated with the gene promoter (ENSR00001516847) indicating such a population as the most plausible target of natural selection.

To get deeper insights into *PRDM16* recent evolutionary history and to better understand which selective pressures could have mainly triggered the observed selective events, age estimation for the putative adaptive haplotypes was obtained. First of all, significantly smaller EHH extension around the core SNPs putatively selected in FIN was found with respect to what observed for the candidate TSI variant, suggesting that positive selection has plausibly acted earlier in the former population. More specifically, the presence of larger haplotype homozygosity around the TSI SNP could be explained by a very recent selective event, while FIN are supposed to be subjected to selective pressures for a longer time interval, thus enabling recombination to erode LD around their adaptive SNPs. The algorithm developed by Voight *et al.* (2006) was then applied to formally obtain age estimations and confirmed these findings by pointing to very recent selective sweeps (3.7-1.1 kya). In accordance to the specificity of the iHS test to detect recent positive selection, it is plausible that selective pressures underlying the observed signals have actually started to act long time before the obtained dating. Plausibly, their intensity then strongly increased in more recent time, thus determining the signatures detected by the applied test. In fact, it is interesting to note that the obtained age estimations reflect the timing of different major climate changes occurred in Europe from the Bronze Age until more recent historical periods. These temperature shifts were linked to natural circumstances that seem

to have strongly influenced the demography of European populations, finally leading to the end of advanced civilizations or to intense migrations (Büntgen *et al.*, 2011; Li *et al.*, 2011; Miller, 2012; Kaniewski *et al.*, 2013). Particularly, recent studies have highlighted that the Bronze Age was a period of great changes in climate conditions, which probably triggered massive migrations from the Steppe to Central/Northern Europe and Asia, as suggested by both genetic and archaeological records (Itan *et al.*, 2009; Anthony & Ringe, 2015; De Fanti *et al.*, 2015; Cassidy *et al.*, 2016). These migrations seem to have deeply influenced the Neolithic European genetic background, contributing to the distinction of northern European populations from southern ones as regards several genomic traits (Allentoft *et al.*, 2015). Therefore, it is plausible to hypothesize that these major climate changes have exerted strong selective pressures on the genetic background of western, and especially northern European populations, as concerns loci involved in thermoregulation processes.

## Conclusions

In accordance with other studies demonstrating that European populations have been exposed to strong selective pressures even within the last 10,000 years (Sabeti *et al.*, 2002; Voight *et al.*, 2006), different populations-specific footprints of natural selection were detected on different regions of the *PRDM16* locus in the GBR, FIN and TSI groups. These signals probably reflect multiple local adaptive events occurred with different intensities among populations and also in different times during the recent evolutionary history of western European populations. In particular, balancing selection could have maintained different haplotypes in the GBR gene pool useful to fine-tuning BAT activity in response to both cold temperatures (i.e. improving thermogenesis) and possible periods of famine, in which instead the processes of energy storage should be privileged. On the other hand, the identified candidate SNPs that plausibly undergone recent

positive selection in FIN and TSI could be barely associated to unique events, but more probably increased their frequencies in response to multiple shifts in climate conditions and nutritional habits occurred over the last 5,000 years.

## Acknowledgements

*This work was supported by the PRIN (Progetti di Ricerca di Interesse Nazionale) Project 2010 "Biological and cultural heritage of the Central-Southern Italian population through 30 thousand years - EPIC" EL8TXP\_006 grant to DL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank all the members of the EPIC project, PRIN 2010-2011, for their collaboration and support to the project.*

## References

- Allentoft M.E., Sikora M., Sjogren K-G., Rasmussen S., Rasmussen M., Stenderup J., Damgaard P.B., Schroeder H., Ahlstrom T., Vinner L. *et al.* 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522: 167.
- Altshuler D.M., Durbin R.M., Abecasis G.R., Bentley D.R., Chakravarti A., Clark A.G., Donnelly P., Eichler E.E., Flicek P, Gabriel S.B. *et al.* 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 49: 56-65.
- Anthony D.W., Ringe D. 2015. The Indo-European Homeland from Linguistic and Archaeological Perspectives. *Annu. Rev. Linguist.*, 1: 199-219.
- Barbera M.J., Schluter A., Pedraza N., Iglesias R., Villarroya F. & Giralt M. 2001. Peroxisome proliferator-activated receptor alpha activates transcription of the brown fat uncoupling protein-1 gene. A link between regulation of the thermogenic and lipid oxidation pathways in the brown fat cell. *J. Biol. Chem.*, 276: 1486-1493.
- Barreiro L.B., Laval G., Quach H., Patin E. & Quintana-Murci L. 2008. Natural selection has

- driven population differentiation in modern humans. *Nat. Genet.*, 40: 340-345.
- Breslin P.A. 2013. An evolutionary perspective on food and human taste. *Curr. Biol.*, 23:R409-418.
- Brown E.A. 2012 Genetic explorations of recent human metabolic adaptations: hypotheses and evidence. *Biol. Rev. Camb. Philos. Soc.*, 87: 838-855.
- Büntgen U., Tegel W., Nicolussi K., McCormick M., Frank D., Trouet V., Kaplan J.O., Herzig F., Heussner K.U., Wanner H. et al. 2011. 2500 years of European climate variability and human susceptibility. *Science*, 331: 578-582.
- Cassidy L.M., Martiniano R., Murphy E.M., Teasdale M.D., Mallory J., Hartwell B. & Bradley D.G. 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl. Acad. Sci. USA*, 113: 368-373.
- Cohen P., Levy J.D., Zhang Y., Frontini A., Kolodin D.P., Svensson K.J., Lo J.C., Zeng X., Ye L., Khandekar M.J. et al. 2014. Ablation of PRDM16 and beige adipose causes metabolic dysfunction and a subcutaneous to visceral fat switch. *Cell*, 156: 304-316.
- Cordain L., Eaton S.B., Miller J.B., Mann N. & Hill K. 2002. The paradoxical nature of hunter-gatherer diets: meat-based, yet non-atherogenic. *Eur. J. Clin. Nutr.*, 56: S42-S52.
- Cordain L., Eaton S.B., Sebastian A., Mann N., Lindeberg S., Watkins B.A., O'Keefe J.H. & Brand-Miller J. 2005. Origins and evolution of the Western diet: health implications for the 21st century. *Am. J. Clin. Nutr.*, 81: 341-354.
- De Fanti S., Sazzini M., Giuliani C., Frazzoni F., Sarno S., Boattini A., Marasco E., Mantovani V., Franceschi C., Moral P. et al. 2015. Inferring the genetic history of lactase persistence along the Italian peninsula from a large genomic interval surrounding the LCT gene. *Am. J. Phys. Anthropol.*, 158: 708-718.
- Frühbeck G., Sesma P. & Burrell M.A. 2009. PRDM16: the interconvertible adipo-myocyte switch. *Trends Cell Biol.*, 19: 141-146.
- Hancock A.M., Alkorta-Aranburu G., Witonsky D.B. & Di Rienzo A. 2010a. Adaptations to new environments in humans: the role of subtle allele frequency shifts. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 365: 2459-2468.
- Hancock A.M., Clark V.J., Qian Y. & Di Rienzo A. 2011a. Population genetic analysis of the uncoupling proteins supports a role for UCP3 in human cold resistance. *Mol. Biol. Evol.*, 28: 601-614.
- Hancock A.M., Witonsky D.B., Alkorta-Aranburu G., Beall C.M., Gebremedhin A., Sukernik R., Utermann G., Pritchard J.K., Coop G. & Di Rienzo A. 2011b. Adaptations to climate-mediated selective pressures in humans. *PLoS Genet.*, 7: e1001375.
- Hancock A.M., Witonsky D.B., Ehler E., Alkorta-Aranburu G., Beall C., Gebremedhin A., Sukernik R., Utermann G., Pritchard J., Coop G. et al. 2010b. Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proc. Natl. Acad. Sci. USA*, 107: 8924-8930.
- Hancock A.M., Witonsky D.B., Gordon A.S., Eshel G., Pritchard J.K., Coop G. & Di Rienzo A. 2008. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet.*, 4: e32.
- Hernandez R.D., Kelley J.L., Elyashiv E., Melton S.C., Auton A., McVean G., Sella G., Przeworski M. & Project G. 2011. Classic selective sweeps were rare in recent human evolution. *Science*, 331: 920-924.
- Heyer E. & Quintana-Murci L. 2009. Evolutionary genetics as a tool to target genes involved in phenotypes of medical relevance. *Evol. Appl.*, 2:71-80.
- Hondares E., Rosell M., Díaz-Delfin J., Olmos Y., Monsalve M., Iglesias R., Villarroya F. & Giralt M. 2011. Peroxisome proliferator-activated receptor  $\alpha$  (PPAR $\alpha$ ) induces PPAR $\gamma$  coactivator 1 $\alpha$  (PGC-1 $\alpha$ ) gene expression and contributes to thermogenic activation of brown fat: involvement of PRDM16. *J. Biol. Chem.*, 286: 43112-43122.
- Huttunen P., Hirvonen J. & Kinnula V. 1981. The occurrence of brown adipose tissue in outdoor workers. *Eur. J. Appl. Physiol. Occup. Physiol.*, 46:339-345.

- Itan Y., Powell A., Beaumont M.A., Burger J. & Thomas M.G. 2009. The origins of lactase persistence in Europe. *PLoS Comput. Biol.*, 5: e1000491.
- Jablonski N.G., Chaplin G. 2010. Colloquium paper: human skin pigmentation as an adaptation to UV radiation. *Proc Natl Acad Sci USA*, 107 Suppl 2: 8962-8968.
- Johnson F., Mavrogianni A., Ucci M., Vidal-Puig A. & Wardle J. 2011. Could increased time spent in a thermal comfort zone contribute to population increases in obesity? *Obes. Rev.*, 12: 543-551.
- Jombart T., Devillard S. & Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.*, 11:94.
- Kajimura S., Seale P., Tomaru T., Erdjument-Bromage H., Cooper M.P., Ruas J.L., Chin S., Tempst P., Lazar M.A. & Spiegelman B.M. 2008. Regulation of the brown and white fat gene programs through a PRDM16/CtBP transcriptional complex. *Genes Dev.*, 22: 1397-1409.
- Kaniewski D., Van Campo E., Guiot J., Le Burel S., Otto T. & Baeteman C. 2013. Environmental roots of the late bronze age crisis. *PLoS One*, 8: e71004.
- Laland K.N. & Brown G.R. 2006. Niche construction, human behavior, and the adaptive-lag hypothesis. *Evol. Anthropol.*, 15: 95-104.
- Lazaridis I., Patterson N., Mittnik A., Renaud G., Mallick S., Kirsanow K., Sudmant P.H., Schraiber J.G., Castellano S., Lipson M. *et al.* 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513: 409.
- Lewis S.L. & Maslin M.A. 2015. Defining the Anthropocene. *Nature*, 519: 171-180.
- Li H., Zhao X., Zhao Y., Li C., Si D., Zhou H. & Cui Y. 2011. Genetic characteristics and migration history of a bronze culture population in the West Liao-River valley revealed by ancient DNA. *J. Hum. Genet.*, 56: 815-822.
- Li N. & Stephens M. 2003. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, 165: 2213-2233.
- Librado P. & Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.
- Lopez Herraez D., Bauchet M., Tang K., Theunert C., Pugach I., Li J., Nandineni M.R., Gross A., Scholz M. & Stoneking M. 2009. Genetic variation and recent positive selection in worldwide human populations: evidence from nearly 1 million SNPs. *PLoS One*, 4: e7888-e7888.
- Mathieson I., Lazaridis I., Rohland N., Mallick S., Patterson N., Roodenberg S.A., Harney E., Stewardson K., Fernandes D., Novak M. *et al.* 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528: 499.
- Miller G.H., Geirsdóttir Á., Zhong Y., Larsen D.J., Otto-Bliesner B.L., Holland M.M., Bailey D.A., Refsnider K.A., Lehman Scott J., Southon J.R., Anderson C., Björnsson H. & Thordarson T. 2012. Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophysical Research Letters*, 39.
- Ohno H., Shinoda K., Ohyama K., Sharp L.Z. & Kajimura S. 2013. EHMT1 controls brown adipose cell fate and thermogenesis through the PRDM16 complex. *Nature*, 504: 163-167.
- Ohno H., Shinoda K., Spiegelman B.M. & Kajimura S. 2012. PPAR $\gamma$  agonists induce a white-to-brown fat conversion through stabilization of PRDM16 protein. *Cell Metab.*, 15: 395-404.
- Ouellet V., Labbé S.M., Blondin D.P., Phoenix S., Guérin B., Haman F., Turcotte E.E., Richard D. & Carpentier A.C. 2012. Brown adipose tissue oxidative metabolism contributes to energy expenditure during acute cold exposure in humans. *J. Clin. Invest.*, 122: 545-552.
- Puigserver P., Wu Z., Park C.W., Graves R., Wright M. & Spiegelman B.M. 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell*, 92: 829-839.
- Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., de Bakker P.I., Daly M.J. *et al.* 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, 81: 559-575.

- Raj S.M., Pagani L., Gallego Romero I., Kivisild T. & Amos W. 2013. A general linear model-based approach for inferring selection to climate. *BMC Genet.*, 14:87.
- Rivas-Martínez S., Penas A. & Díaz T.E. 2006. *Biogeographic map of Europe*. Universidad de Leon, Secretariado de Publicaciones.
- Ruddiman W.F. 2013. The Anthropocene. *Annu. Rev. Earth Planet Sci.*, 41:45-68.
- Sabeti P.C., Reich D.E., Higgins J.M., Levine H.Z.P., Richter D.J., Schaffner S.F., Gabriel S.B., Platko J.V., Patterson N.J., McDonald G.J. et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419: 832-837.
- Sazzini M., Schiavo G., De Fanti S., Martelli P.L., Casadio R. & Luiselli D. 2014. Searching for signatures of cold adaptations in modern and archaic humans: hints from the brown adipose tissue genes. *Heredity* (Edinb.), 113: 259-267.
- Sazzini M., Gnechchi Ruscone G.A., Giuliani C., Sarno S., Quagliariello A., De Fanti S., Boattini A., Gentilini D., Fiorito G., Catanoso M. et al. 2016. Complex interplay between neutral and adaptive evolution shaped differential genomic background and disease susceptibility along the Italian peninsula. *Sci. Rep.*, 6: 32513. doi: 10.1038/srep32513.
- Scheet P. & Stephens M. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.*, 78: 629-644.
- Skoglund P., Malmstrom H., Raghavan M., Stora J., Hall P., Willerslev E., Gilbert M.T.P., Gotherstrom A. & Jakobsson M. 2012. Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. *Science*, 336: 466-469.
- Stewart J.R. & Stringer C.B. 2012. Human evolution out of Africa: the role of refugia and climate change. *Science*, 335: 1317-1321.
- Sturm R.A. 2009. Molecular genetics of human pigmentation diversity. *Hum. Mol. Genet.*, 18: R9-17.
- Turchin M.C., Chiang C.W., Palmer C.D., Sankararaman S., Reich D., Hirschhorn J.N. & Consortium GloATG. 2012. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat. Genet.*, 44: 1015-1019.
- Voight B.F., Kudaravalli S., Wen X. & Pritchard J.K. 2006. A map of recent positive selection in the human genome. *PLoS Biol*, 4: e72.
- Wright J.D. 1993. Deglaciation triggered by the resumption of North Atlantic Deep Water. *EOS*, 74: 24
- Wu J., Boström P., Sparks L.M., Ye L., Choi J.H., Giang A.H., Khandekar M., Virtanen K.A., Nuutila P., Schaart G. et al. 2012. Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. *Cell*, 150: 366-376.
- Yamane M., Oida Y., Ohnishi N., Matsumoto T. & Kitagawa K. 2010. Effects of wind and rain on thermal responses of humans in a mildly cold environment. *Eur. J. Appl. Physiol.*, 109: 117-123.
- Ye K. & Gu Z. 2011. Recent advances in understanding the role of nutrition in human genome evolution. *Adv. Nutr.*, 2: 486-496.
- Zhang D.D., Brecke P., Lee H.F., He Y-Q. & Zhang J. 2007. Global climate change, war, and population decline in recent human history. *Proc. Natl. Acad. Sci. USA*, 104: 19214-19219.

Associate Editor, Andrea Manca

