Received: 15 September 2022 | Revised: 18 August 2023 | Accepted: 5 December 2023

DOI: 10.1111/mec.17257

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Multiple genomic solutions for local adaptation in two closely related species (sheep and goats) facing the same climatic constraints

Badr Benjelloun[1,2] | Kevin Leempoel[3] | Frédéric Boyer[2] | Sylvie Stucki[3] | Ian Streeter[4] | Pablo Orozco-terWengel[5,6] | Florian J. Alberto[2] | Bertrand Servin[7] | Filippo Biscarini[8] | Adriana Alberti[9,10] | Stefan Engelen[11] | Alessandra Stella[8] | Licia Colli[12,13] | Eric Coissac[2] | Michael W. Bruford[5,6] | Paolo Ajmone-Marsan[12,13] | Riccardo Negrini[12,14] | Laura Clarke[4] | Paul Flicek[4] | Abdelkader Chikhi[1] | Stéphane Joost[3] | Pierre Taberlet[2] | François Pompanon[2]

[1]Livestock Genomics Laboratory, Regional Center of Agricultural Research Tadla, National Institute of Agricultural Research INRA, Rabat, Morocco

[2]Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, LECA, Grenoble, France

[3]Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[4]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, Cambridge, UK

[5]School of Biosciences, Cardiff University, Wales, UK

[6]Sustainable Places Research Institute, Cardiff University, Cardiff, UK

[7]GenPhySE, Université de Toulouse, INRAE, INPT, ENVT, Castanet-Tolosan, France

[8]Institute of Agricultural Biology and Biotechnology, Consiglio Nazionale delle Ricerche (CNR), Milan, Italy

[9]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ. Evry, Université Paris-Saclay, Evry, France

[10]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

[11]Genoscope, Institut de biologie François-Jacob, Commissariat à l'Energie Atomique CEA, Université Paris-Saclay, Evry, France

[12]Dipartimento di Scienze Animali, della Nutrizione e degli Alimenti, Facoltà di Scienze Agrarie, Alimentari e Ambientali, Università Cattolica del S. Cuore, Piacenza, Italy

[13]BioDNA – Centro di Ricerca sulla Biodiversità e sul DNA Antico, Facoltà di Scienze Agrarie, Alimentari e Ambientali, Università Cattolica del S. Cuore, Piacenza, Italy

[14]AIA Associazione Italiana Allevatori, Roma, Italy

Correspondence
Badr Benjelloun, Livestock Genomics Laboratory, Regional Center of Agricultural Research Tadla, National Institute of Agricultural Research INRA, Rabat, Morocco.
Email: badr.benjelloun@inra.ma

François Pompanon, Université Grenoble Alpes, Université Savoie Mont Blanc,

Abstract

The question of how local adaptation takes place remains a fundamental question in evolutionary biology. The variation of allele frequencies in genes under selection over environmental gradients remains mainly theoretical and its empirical assessment would help understanding how adaptation happens over environmental clines. To

Badr Benjelloun, Kevin Leempoel and Frédéric Boyer contributed equally to this study.

The authors express profound sorrow over the passing of their co-author Michael W. Bruford, who played a significant role in the project presented in this publication. Mike was a compassionate individual, an exceptional scientist and a motivating presence for all of us. We will fondly remember his insightful scientific conversations, his infectious enthusiasm and his genuine sincerity.

CNRS, LECA, F-38000 Grenoble, France.
Email: francois.pompanon@univ-grenoble-alpes.fr

Funding information
Seventh Framework Programme, Grant/Award Number: 244356

Handling Editor: Alana Alexander

bring new insights to this issue we set up a broad framework which aimed to compare the adaptive trajectories over environmental clines in two domesticated mammal species co-distributed in diversified landscapes. We sequenced the genomes of 160 sheep and 161 goats extensively managed along environmental gradients, including temperature, rainfall, seasonality and altitude, to identify genes and biological processes shaping local adaptation. Allele frequencies at putatively adaptive loci were rarely found to vary gradually along environmental gradients, but rather displayed a discontinuous shift at the extremities of environmental clines. Of the 430 candidate adaptive genes identified, only 6 were orthologous between sheep and goats and those responded differently to environmental pressures, suggesting different putative mechanisms involved in local adaptation in these two closely related species. Interestingly, the genomes of the 2 species were impacted differently by the environment, genes related to signatures of selection were most related to altitude, slope and rainfall seasonality for sheep, and summer temperature and spring rainfall for goats. The diversity of candidate adaptive pathways may result from a high number of biological functions involved in the adaptations to multiple eco-climatic gradients, and a differential role of climatic drivers on the two species, despite their co-distribution along the same environmental gradients. This study describes empirical examples of clinal variation in putatively adaptive alleles with different patterns in allele frequency distributions over continuous environmental gradients, thus showing the diversity of genetic responses in adaptive landscapes and opening new horizons for understanding genomics of adaptation in mammalian species and beyond.

KEYWORDS
contingency, domestic species, evolutionary convergence, landscape genomics

1 | INTRODUCTION

Local adaptation describes the adjustment or change in behaviour, physiology and morpho-anatomy of an organism to better fit its environment, and locally adapted populations exhibit higher fitness in their native habitats than populations from elsewhere (Kawecki & Ebert, 2004). Identifying genomic changes underlying local adaptation is critical in addressing fundamental questions in evolutionary biology. Most adaptive traits are affected by many segregating loci and are thus difficult to study (Ward & Kellis, 2012), explaining why relatively few studies have so far elucidated fine scale patterns of local adaptation. For example, adaptation of deer mice to a sand-coloured background is associated with independent selection on many Single Nucleotide Polymorphisms (SNP) within a single gene, each with a specific effect (Linnen et al., 2013). In contrast, a single mutation increases hair thickness, affects mammary fat pad size and increases eccrine gland number in mice (Kamberov et al., 2013). Such variation in the complexity of selection regime and genetic architecture of adaptive phenotype suggests the occurrence of multiple evolutionary trajectories at the levels of genetic and metabolic processes. These trajectories may be divergent (i.e. involving selection on different loci), parallel (repeated sorting of alleles identical by descent) or convergent (different mutation on the same locus affecting distinct lineages) between populations or species (Waters & McCulloch, 2021). In the case of repeated evolutionary trajectories, repeated sorting and convergence are mainly observed at the intra-specific and inter-specific levels respectively (Waters & McCulloch, 2021). Cases of genetic convergence and parallelism are frequently reported in the literature (e.g. Martin & Orgogozo, 2013; Christin et al., 2010; Elmer & Meyer, 2011; Conte et al., 2012; Bohutínská et al., 2021). Well-known examples are related to adaptation to altitude in Tibetans and domesticated and wild vertebrates, that is, dogs, sheep, goats and falcons (Cao et al., 2021; Gou et al., 2014; Lim et al., 2019; Simonson et al., 2010; Song et al., 2016; Wei et al., 2016; Yang et al., 2016; Yi et al., 2010; Zhan et al., 2013). However, other studies have reported similar adaptations based on divergent genomic architecture in closely related species/populations, such as a single pigmentation pattern controlled by different genes in beach mice (Manceau et al., 2010; Steiner et al., 2009) or different genomic regions involved in the adaptation of co-existing stickleback species despite substantial phenotypic parallelism (Raeymaekers et al., 2017). In addition to the influence of the genomic context, evolutionary trajectories also depend on the effects of

environmental factors and the interactions between the two. The most impacting environmental factors may differ between species or populations, as well as the threshold values of environmental drivers that affect genotype via phenotype, leading to differential responses along environmental gradients.

Previous studies addressing these key issues often compared isolated populations from contrasting environments under extremely divergent selective pressures (Gou et al., 2014; Linnen et al., 2013) or at wide geographical scales (Lv et al., 2014). Moreover, environmental factors may vary gradually along gradients rather than changing abruptly, raising the question of whether genomic variants related to adaptation, which are also affected by genetic drift and migration, vary in a similar way (Riesch et al., 2018). A few studies have examined genomic signatures of selective pressures along environmental gradients including latitudinal clines in wing size of *Drosophila subobscura* (Huey et al., 2000) and flowering time in *Arabidopsis thaliana* (Stinchcombe et al., 2004). Landscape genomics (Joost et al., 2007; Manel et al., 2003) is a well-suited approach to address such questions (Manel & Holderegger, 2013). To date, most landscape genomics studies have used population genomics approaches to detect the adaptive variation within genomes in a spatial context (Manel & Holderegger, 2013); however, specific methods have also been developed that directly correlate allele frequencies with environmental gradients (Frichot et al., 2013; Joost et al., 2007; Stucki et al., 2017).

Here, we used a large-scale landscape genomics design, relying on individual-based sampling and Whole Genome Sequencing (WGS) to examine how adaptive genomic variation changes spatially along environmental gradients and thus scrutinize adaptive trajectories in different species experiencing the same environmental heterogeneity. As repeated evolution is expected to occur more frequently in close lineages (Bohutínská et al., 2021; Conte et al., 2012), we sought to assess the similarity in adaptive pathways in closely related taxa, that is, sheep (*Ovis aries*) and goats (*Capra hircus*), under the same environments. The *Ovis*/*Capra* divergence occurred around 6 Mya (Hassanin et al., 2012), and both domestic species experienced similar histories with overlapping centres of domestication around 10 kya ago near the Fertile-Crescent (Naderi et al., 2008; Vigne et al., 2005; Zeder et al., 2006), and the same colonization routes led to their rapid spread over the old-world. The populations and breeds studied here have been managed extensively (i.e. in large areas with moderate human input and moderate productivity) with limited anthropogenic selection (and not subject to genomic selection), and have accumulated valuable adaptations to their local environment (climate, ecology and husbandry), while maintaining a high level of genetic diversity (Taberlet et al., 2008). We focused on indigenous co-distributed sheep and goat landraces in Morocco, and sequenced the genomes of 161 goats and 160 sheep representative of the whole diversity of populations and breeds inhabiting Morocco's high heterogeneity of geo-climatic conditions (e.g. Mediterranean, oceanic, high mountains, desert; Figure S1). We detected genomic signatures of selection associated with a wide variety of eco-climatic pressures that allowed us to identify sets of candidate genes and biological processes that may have shaped adaptation in sheep and goats. We characterized the patterns of variation in candidate adaptive allele frequencies that exhibited variation along environmental clines. Finally, we showed that the genetic architecture underlying the candidate adaptive responses to the same environmental gradients were mostly different between the two species.

2 | METHODS

A workflow of the whole analytic procedure is given in Figure S2.

2.1 | Sampling

The sampling area (~400,000 km^2; latitude range [28°–36°]; Figure 1) covered the whole range of contrasting environments across Morocco. Sheep and goats are in general traditionally managed in Morocco with no genomic selection. Grazing is the main source of feed for herds in most of the country and breeding programs cover a minor part of the herds. Large-scale herd movements are less and less used. Thus, a sampling grid consisting of 162 cells of 0.5° of longitude and latitude was established and a maximum of three animals was sampled per flock for three different flocks per cell for each species. We collected the samples between January 2008 and March 2012 for the NextGen European project (KBBE-2009-1-1-03) in accordance with ethical regulations of the European Union Directive 86/609/EEC. For each individual, tissue samples were collected from the distal part of the ear and placed in alcohol for 1 day, and then transferred to a silica-gel tube until DNA extraction. A total of 412 flocks were sampled from which we aimed to select 164 individuals for Whole Genome Sequencing while optimizing sample selection to include the widest possible range of environmental conditions. The extent of the environmental gradients (see Table S1) with areas of harsh conditions makes it unlikely that both species are at their optimum throughout the sampling area. We took the geographical location into account to maximize the spread of individuals across the region to ensure spatial representativeness for all environments. We, therefore, first performed a principal component analysis (PCA) on the 117 environmental variables extracted from the Climatic Research Unit (CRU) dataset (New et al., 2002). The PCA allowed us to maximize the ecological distance between the sample sites (separately for sheep and goats). Afterwards, we performed an Ascendent Hierarchical Classification (AHC) on the first 7 PCA-axes (96% of the variance) to group sampled sites according to their ecological distances. Using the Ward criterion, we reduced the number of classes to 164. After having grouped sampled sites, we selected one individual per class. In order to guarantee spatial representativeness, we performed 50 random samplings and chose the one with the maximum index of distribution (i.e. the maximum sum of distances between each site and its nearest neighbour). After sequencing, we removed seven individuals with low sequence quality, ending up with 161 goats and 160 sheep (Figure 1).

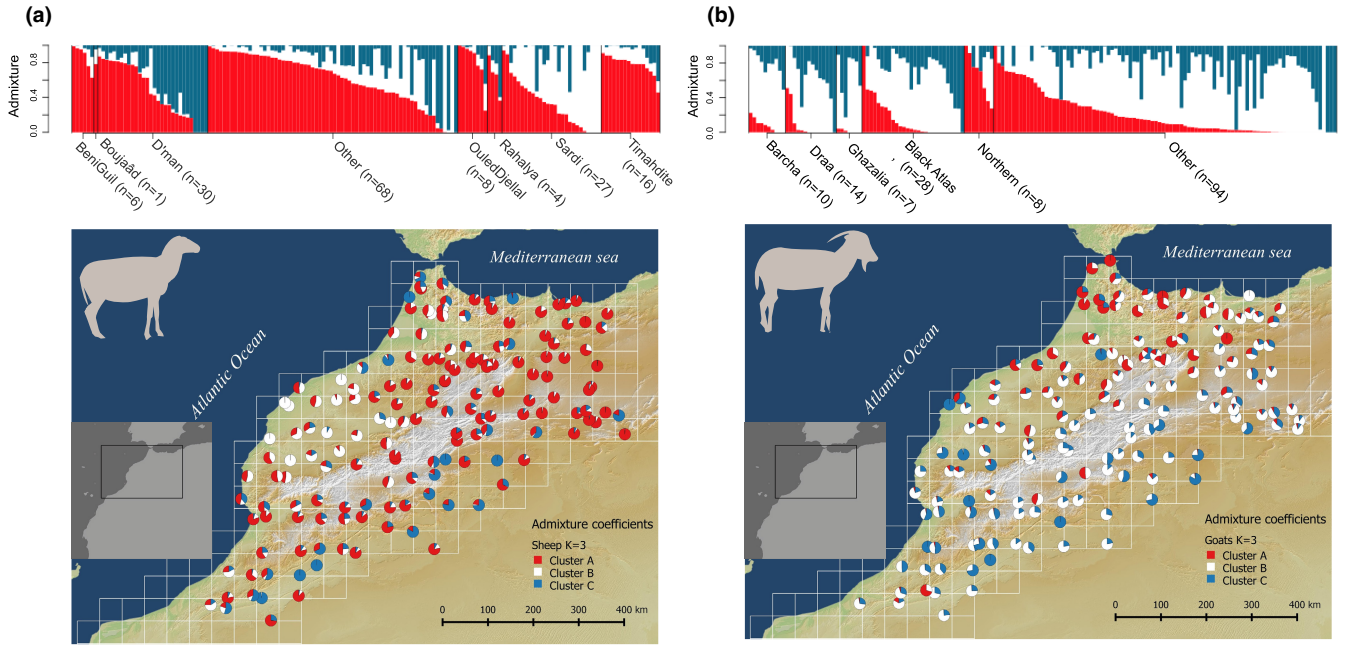

FIGURE 1 Genetic structure and sampling distribution for sheep (a) and goats (b). The upper charts show the proportions of genomes assigned to 3 genetic clusters according to the breed (i.e. admixture coefficients inferred by sNMF for the most likely number of clusters for both species). On the maps, each individual is represented by a pie chart located at the sampling site and showing the admixture coefficients.

2.2 | Production of WGS data

DNA extraction was done using the Puregene Tissue Kit from Qiagen® following the manufacturer's instructions and whole genome sequence data were produced and processed as described in Benjelloun et al., (2015). 500 ng of DNA was sheared to a 150–700-bp range using the Covaris® E210 (Covaris, Inc., USA). Sheared DNA was used for Illumina® library preparation with a semi-automated protocol. Briefly, end repair, A-tailing and Illumina® compatible adaptor (BiooScientific) ligation were performed using the SPRIWorks Library Preparation System and the SPRI TE instrument (Beckmann Coulter), according to the manufacturer's protocol. A 300–600-bp size selection was applied in order to recover most of fragments. DNA fragments were amplified in 12-cycle PCR using the Platinum Pfx Taq Polymerase Kit (Life® Technologies) and Illumina® adapter-specific primers. Libraries were purified with 0.8× AMPure XP beads (Beckmann Coulter). After library profile analysis using the Agilent 2100 Bioanalyzer (Agilent® Technologies, USA) and qPCR quantification, libraries were sequenced using 100 base-length read chemistry in paired-end flow cell on the Illumina HiSeq2000 (Illumina®, USA).

2.3 | WGS data processing

Illumina paired-end reads for sheep were mapped to the sheep reference genome (OAR v3.1, GenBank assembly GCA_000317765.1 (Jiang et al., 2014)) and those for goats were mapped to the goat reference genome (CHIR v1.0, GenBank assembly GCA_000317765.1 (Dong et al., 2013)) using BWA mem (Li & Durbin, 2009). The BAM files produced were then sorted using Picard SortSam and improved using Picard Markduplicates (http://picard.sourceforge.net), GATK RealignerTargetCreator, GATK IndelRealigner (DePristo et al., 2011) and Samtools calmd (Li et al., 2009). Variant calling was carried out using three different algorithms: Samtools mpileup (Li et al., 2009), GATK UnifiedGenotyper (McKenna et al., 2010) and Freebayes (Garrison & Marth, 2012).

There were two successive rounds of variant site filtering. Stage 1 merged calls together from the three algorithms, while filtering out the lowest-confidence calls. A variant site passed if it was called by at least two different calling algorithms with variant quality >30. An alternate allele at a site passed if it was called by any one of the calling algorithms, and the genotype count >0. Filtering stage 2 used Variant Quality Score Recalibration by GATK. First, we generated a training set of the highest confidence variant sites where (i) the site is called by all three variant callers with variant quality >100, (ii) the site is biallelic (Palti et al., 2015) the minor allele count is at least 3 while counting only samples with genotype quality >30. The training set was used to build a Gaussian model using the GATK VariantRecalibrator tool using the following variant annotations from UnifiedGenotyper: QD, HaplotypeScore, MQRankSum, ReadPosRankSum, FS, DP, Inbreeding Coefficient. The Gaussian model was applied to the full data set, generating a VQSLOD (log odds ratio of being a true variant). Sites were filtered out if VQSLOD < cut-off value. The cut-off value was set for each group by the following: Minimum VQSLOD = {the median value of VQSLOD for training set variants} − 3 * {the median absolute deviation VQSLOD of training set variants}. Measures of the transition/transversion ratio of SNPs suggest that this chosen cut-off criterion gave the best balance between selectivity and sensitivity. Genotypes were

improved and phased using Beagle 4 (Browning & Browning, 2013), and then filtered out where the genotype probability calculated by Beagle was less than 0.95.

In order to detect orthologous signals of selection between sheep and goats, we performed a cross-alignment between the two reference genomes as described in Alberto et al. (2018). First, we used the pairwise alignment pipeline from the Ensembl release 69 code base (Flicek et al., 2012) to align the reference genomes of sheep (OARv3.1) and goat (CHIR1.0). This pipeline uses LastZ (Harris, 2007) to align at the DNA level, followed by post-processing in which aligned blocks are chained together according to their location in both genomes. The LastZ pairwise alignment pipeline is run routinely by Ensembl for all supported species, but the goat was not included in Ensembl by the time of our analysis. We produced two different inter-specific alignments, using sheep as the reference genome and goat as non-reference and vice versa. This was done to avoid bias toward either species, given that genomic regions of the reference species are forced to map uniquely to a single locus of the non-reference species, whereas non-reference genomic regions are allowed to map to multiple locations of the reference species. For segments of chromosomes of one reference genome, we obtained the coordinates on the non-reference genome. Finally, for the SNPs discovered in one genus, we used the whole genome alignment with the reference genome of the other genus to identify the corresponding positions (more than 90% of each genome was covered by these alignments, see Alberto et al., 2018).

2.4 | Genetic diversity and structure

Neutral genomic variation was characterized to evaluate the level of genetic diversity present in Moroccan sheep and goats. The total number of variants and the number of variants within each species were calculated. The level of nucleotide diversity (π) was calculated in each species and averaged over all of the biallelic and fully diploid variants for which all individuals had a genotype called using Vcftools (Danecek et al., 2011). The observed frequency of heterozygous genotypes per individual (*Ho*) was calculated considering only the biallelic SNPs with no missing genotype calls. From *Ho*, inbreeding coefficients (*F*) were calculated for each individual using the allele frequencies calculated over all individual genomes. F-tests were applied to test for the significance of differences between sheep and goat *Ho* and *F*.

Genetic structure was assessed using two different methods. First, we performed a principal component analysis (PCA) using an *LD* pruned subset of bi-allelic SNPs. *LD* between SNPs in windows containing 50 markers was calculated before removing one SNP from each pair where r^2 exceeded .95. Subsequently, only 12,543,534 SNPs among a total of 29,427,980 bi-allelic SNPs were kept for this analysis in goats and 14,056,772 out of 30,069,299 of biallelic SNPs for sheep. We used Plink v1.90a (https://www.cog-genomics.org/plink2) for *LD* pruning and the R package adegenet v1.3–1 (Jombart & Ahmed, 2011) to run the PCA. Second, we carried out an analysis with the sNMF method of genetic clustering (Frichot et al., 2014). This method was specifically developed for fast analysis of large genomic datasets. It is based on sparse non-negative matrix factorization to estimate admixture coefficients of individuals. We used all bi-allelic variants and performed five runs for each number of genetic clusters (i.e. K value varying from 1 to 10) using a value of *alpha* parameter of 16. For each run, we calculated the cross-entropy criterion with 5% missing data to identify the most likely number of clusters. We considered the run showing the lowest cross-entropy (CE) as the most likely value of K. Similarly we considered the number of clusters associated with the lowest CE as the most likely representative of our data structure.

2.5 | Environmental variables

We extracted the values of 67 climatic variables from the WorldClim dataset (Hijmans et al., 2005) (http://www.worldclim.org/current) for the sampling locations. These variables are based on data collected over 30 years and provide temperature and rainfall measurements as well as bioclimatic indices (i.e. derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables) with an initial resolution of 1 × 1 km. Additionally, we used a Digital Elevation Model (DEM) with a resolution of 90 m (SRTM; http://earthexplorer.usgs.gov; courtesy of the U.S. Geological Survey) to obtain topography-related variables. We computed a total of 4 DEM-derived variables with SAGA GIS (Böhner et al., 2006): altitude, slope and solar radiation in June and December. Afterwards, we conducted a pairwise correlation analysis between all 71 variables to detect highly correlated ones ($|r| \geq .8$), and kept 10 representative variables to investigate signatures of selection (Tables S2 and S3).

2.6 | Analyses of signatures of selection

We applied two approaches to identify putative signatures of selection. The first was a correlative approach processing multiple parallel association models. Logistic regressions were performed between the frequencies at each locus and the value of the environmental variables at the sampled location. For this analysis, we used Samβada (Stucki et al., 2017), a method performing logistic regressions on individual genotypes. This is a variant of linear regression in which the binary genetic marker is either present or absent and correlates with a quantitative environmental variable. Therefore, it provides the probability of occurrence of a genotype for each individual as a function of environmental gradients (Joost et al., 2007). It also provides spatial statistics that are helpful for the interpretation of significant results regarding spatial autocorrelation and is efficient in terms of computing time in the context of several million SNPs to process with 10 environmental variables (Joost et al., 2007). We considered only SNPs with a MAF ≥ 0.05. Because Samβada requires binomial

data, we coded 0 or 1 for each allele of the biallelic SNPs studied. Each of the 3 possible genotypes was thus represented by 2 points. We then performed univariate regressions on each genotype with each of the 10 variables. Thereafter, the false discovery rate was calculated for each variable separately and Q-values were computed. The false discovery threshold was set to 0.1 on Samβada's results to identify candidate SNPs.

The second approach was group-based and compared, through genome scans, two groups from contrasting conditions for each environmental variable. More precisely, for each variable, each of the two groups was constituted by the 20 individuals subjected to the most extreme values of the variable, at each side of the environmental gradient. Then, we ran XP-CLR (Chen et al., 2010) to identify potential genomic regions differentially selected between groups. This method is designed to identify regions under positive selection in an object group, given the other group considered as the reference. Therefore, we carried-out each analysis twice, each group being in turn the reference and the object. XP-CLR is robust to detect selective sweeps and especially with regard to the uncertainty in the estimation of local recombination rate (Chen et al., 2010). Due to the absence of reliable information on genetic linkage for the whole genome in both species, the physical position (1 Mb ≈ 1 cM) was used. We used overlapped segments of a maximum of 27 Mb to estimate and assemble XP-CLR scores using the whole set of biallelic variants as described in Benjelloun et al., (2015). Segments were overlapping by 2 Mb and the scores computed for the extremes 1 Mb were discarded, except at the starting and end of chromosomes. XP-CLR scores were calculated using grid points spaced by 2500 bp in a window of 0.1 Mb containing more than 200 variants, and by down-weighting contributions of highly correlated variants (randomly removing one variant for each pair being correlated with $r^2 > .95$) in the reference group. Furthermore, a maximum of 250 SNPs were then randomly selected to calculate XP-CLR scores in each window. The 0.01% genomic regions with the highest XP-CLR scores were identified. Within these regions (0.1 Mb each), the top differentiated variants between the two groups were defined applying an *F*st (Weir & Cockerham, 1984) cut-off level of 0.05% of the genome-wide *F*st distribution. This threshold was chosen for having the highest enrichment values for outliers within XP-CLR windows compared to 0.1% and 0.01% (Student's *t*-test, $p = 7.3e\text{-}13$ for sheep and $p = 5.2e\text{-}13$ for goats and *F*-test, $p = 1.1e\text{-}37$ for sheep and $p = 1.6e\text{-}32$ for goats). In addition, overlapping top XP-CLR regions were merged and all regions were ranked according to their XP-CLR score. Then, we checked for the occurrence of orthologous outlier variants, genes and regions between sheep and goats. For that, outlier variants identified in one species were mapped on the genome assembly of the other species.

In order to test whether the association between outliers putatively under selection and environmental drivers differed according to the species, we calculated for each environmental parameter the following index: Δg = (number of associated sheep genes − number of associated goat genes)/(number of associated sheep genes + number of associated goat genes). Obviously, this index varied from −1 (only goat genes associated with the parameter) to +1 (only sheep genes associated with the parameter). Δg was calculated for genes detected with XP-CLR/Fst (Δgx) and a similar index was calculated for genomic regions (Δgr) and genes detected with Samßada (Δgs).

In order to detect whether the environmental variables impacted both species in the same way, the distributions of the number of outlier genes or genomic regions with regard to environmental variables were compared between species using χ^2 tests.

We used the Variant Effect Predictor (VEP) tools (McLaren et al., 2010) to classify the variants as intronic, exonic, synonymous, missense and intergenic. Genes including at least one candidate variant or located at less than 5 kb away from it (downstream 5′-end and upstream 3′-end) were used for Gene Ontology (GO) enrichment analyses. These analyses were done for each species using the sets of genes identified for each variable separately.

For each species we computed the profiles of allelic frequency variation for each putatively selected region along the corresponding environmental gradient. For this analysis, we subdivided the range of values in 8 groups corresponding each to 20 contiguous individuals (except one group of 21 goats) for each environmental variable, and calculated the allelic frequency of the most differentiated candidate variant for each group (on the basis of its *F*st value) associated with each genomic region. A neutral profile of allelic frequencies was represented by calculating for each case the allelic frequency of 100 variants randomly selected over the genome (leading to 41 for goats and 56 for sheep after removing rare variants). We classified all profiles by comparing them to synthetic profiles corresponding to 3 different classes of pattern of variation: (i) a gradual linear variation of the frequency along the gradient, (ii) a uniform frequency distribution all along the gradient with a punctual shift at one extreme end of the gradient and (iii) a uniform frequency distribution along the gradient with punctual shifts at the two extremes. We compared the frequency profiles with the synthetic ones using the Pearson correlation coefficient and attributed each profile to the synthetic profile exhibiting the highest absolute value for Pearson correlation.

2.7 | Gene ontology enrichment analyses

To explore the biological processes involving the whole set of candidate genes identified by at least one method (i.e. Samßada and/or XP-CLR/*F*st), Gene Ontology (GO) enrichment analyses were performed using GOrilla (Eden et al., 2009). The 12,669 and 14,620 genes associated with a GO term in goats and sheep, respectively, were used as background references. We assessed the significance for each individual GO-identifier with *p*-values corrected using a FDR q-value according to Benjamini and Hochberg's method (Benjamini & Hochberg, 1995). GO terms identified for each variable were clustered into homogenous groups using REVIGO by allowing medium similarity (0.7) (Supek et al., 2011). Low similarity option among GO terms in a group was applied and the weight of each GO term was assessed by its *p*-value.

3 | RESULTS

3.1 | Genetic diversity and structure

Whole genome sequences at 12× coverage were generated for 160 sheep and 161 goats and enabled the identification of more than 38.2 and 31.6 million variants respectively. Among these variants, approximately 6.7% comprised small insertions/deletions (indels), and a low proportion (2.1% in sheep and 0.7% goats) of SNPs showed more than two alleles (mainly tri-allelic). Rare variants characterized by a minor allele frequency (MAF) of less than 5% comprised approximately 17 and 18 million SNPs in sheep and goats respectively. Whole genome nucleotide diversity (π) was higher but not significantly so (*t*-test; $p > .05$) in sheep than in goats. The average heterozygosity (*Ho*) was not significantly different between sheep and goats although the latter displayed a significantly higher inbreeding coefficient (*F*) ($p < .05$; Table 1; see Methods).

Genetic structure was analysed with sNMF (Frichot et al., 2014) using only biallelic SNPs (see Methods) and showed weak geographical patterns arising when setting three genetic clusters, with one cluster prevalent in the North for goats, and one cluster prevalent in the West for sheep (Figure 1). However, cross-entropy analysis showed the most likely clustering corresponded to a single unit, suggesting no significant genetic structure (Figures S3 and S4). This was consistent with PCA analysis where the first and second principal components explained less than 2% of the global variation and showed no obvious structure in either species, regardless of breed identification (Figure S5).

3.2 | Signatures of selection

We investigated signatures of selection associated with the 10 least correlated environmental variables (with |Pearson's $r| < .8$) representing altitude, topography, solar radiation, temperature, rainfall, temperature seasonality and rainfall seasonality (see Methods and Table 2), using two approaches: (1) A correlative approach, Samβada (q-value threshold of 0.1; Stucki et al., 2017), was used to identify changes in allele presence/absence of biallelic SNPs along environmental variables using logistic regression for all individuals. In addition, a group-based approach was conducted by contrasting 2 groups comprising 20 individuals each from locations with the most extreme environmental values for each environmental gradient. We then calculated haplotype-based statistics (XP-CLR, Chen et al., 2010) and single-locus *F*st values (Weir & Cockerham, 1984) along the genome, and defined SNPs among the top 0.01% XP-CLR scores and the top 0.05% *F*st values as outlier variants.

Samβada identified 1242 significant allele/environment associations involving 758 SNPs from 108 genomic regions (i.e. regions containing adjacent outlier variants no more than 100 kb apart from each other) in sheep (Table 2; Figure S6). Of these SNPs, 93% clustered in 15 regions on 10 chromosomes, each region bearing at least 10 significant associations. SNPs belonging to the same cluster were associated with one or more environmental variables and displayed similar spatial patterns of allelic variation. Most of the genomic regions were associated with rainfall (prec4), temperature (tmean7) and annual temperature range (bio7). The largest region was centred on the *MC5R* gene on chromosome 23 and carried 839 significant allele/environment associations related to 383 SNPs that were mainly associated with rainfall (prec4), and to a lesser extent with temperature (tmax4, tmean7 and bio7). In goats, Samßada identified 497 significant allele/environment associations for 311 SNPs from 61 genomic regions (Table 2; Figure S6). Most associations were with annual temperature range (bio7), rainfall seasonality (bio15) and altitude. As with sheep, >90% of the significant SNPs were clustered, with 7 regions containing at least 10 significant associations. We observed a major cluster of 50 outliers mainly associated with rainfall seasonality (bio15) and altitude, centred on a non-annotated gene on chromosome 6. A search for its orthologue using Ensembl showed a high similarity (99.6% match) with the bovine gene *CAMK2D*, which is associated with the development of mammary tissue (Nguyen & Shively, 2016) and cardiac function (Little et al., 2007). We can hypothesize a possible association between this gene and goat dairy production even if goats in Morocco are generally raised for both meat and milk production. Four additional regions on chromosome 4, 24, 13 and 27 were associated with rainfall seasonality (bio15), and four regions on chromosome 4, 6, 9 and 11 were associated with temperature seasonality (bio7).

The group-based approach (XP-CLR/Fst) identified 9433 outlier variants in sheep from 210 genomic regions, which were associated with 183 genes. Fourteen candidate genes were associated with altitude, 24 with rainfall (prec4), 25 with annual temperature range (bio7), 27 with temperature (tmean7) and 8 genes with slope (Table 2; Figure S7). In goats, we found 5292 variants from 297 genomic regions linked to 198 genes. Thirty candidate genes were associated with altitude, 17 with rainfall (prec4), 16 with temperature annual range (bio7), 17 with temperature (tmean7) and 19 with slope (Table 2). Outlier loci associated with given environmental variables clearly showed different prevalence of alleles under both extremities of the variable gradient (compared to the distribution of neutral variants, *F*-test of equality

TABLE 1 Whole Genome Diversity in Moroccan sheep and goats.

Species	#variants	#SNPs	Nucleotide diversity (π)	Heterozygosity (*Ho*)[a]	Inbreeding coefficient (*F*)[a]
Sheep ($n = 160$)	38,278,356	35,992,193	2.5×10^{-3}	0.166 ± 0.014	0.045 ± 0.081
Goats ($n = 161$)	31,743,850	29,606,371	1.6×10^{-3}	0.119 ± 0.012	0.056 ± 0.096

[a]Heterozygosity and inbreeding coefficient are presented as "average ± standard deviation".

TABLE 2 Number of outlier variants and related genes identified for each environmental parameter in sheep and goats.

		Number of variants/genes					
		Sheep			Goat		
Environmental parameter	WorldClim and DEM-derived variables	Correlative approach	Group-based approach	Common	Correlative approach	Group-based approach	Common
Altitude	Altitude (m)	0/0	650/14	–	80/6	587/30	3/1
Slope	Slope	0/0	672/8	–	0/0	440/19	–
Solar radiations	ti2112 (sunshine duration on December 21st)	0/0	1066/25	–	0/0	452/20	–
Temperature	tmin8 (average monthly minimum temperature in August)	0/0	1262/39	–	0/0	585/16	–
	tmean7 (average monthly mean temperature in July)	276/13	1653/27	173/7	0/0	427/17	–
Rainfall	Prec4 (average monthly precipitation in April)	839/32	1500/23	290/5	5/0	718/17	0/0
Temperature seasonality	bio7 (temperature annual range)	141/8	904/25	0/0	135/8	466/16	2/0
	bio8 (mean temperature of the wettest quarter)	0/0	1145/21	–	0/0	421/19	–
	bio3 (isothermality)	0/0	1203/33	–	0/0	567/26	–
Rainfall seasonality	bio15 (precipitation seasonality, coefficient of variation)	0/0	785/8	–	127/14	797/26	51/1
Overall (# of associations/# of variants # of genomic-regions/# of genes)		1242/758 108/43	10,840/9434 210/183	463/317 4/8	347/311 61/16	5460/5292 297/198	56/56 5/2

Note: The number of outliers detected via the correlative approach (Samβada), the group-based approach (XP-CLR/Fst) and both. The total number ('Overall') is not the exact sum of the previous lines when same variants/genes are found in several categories.

of variances, $p=1.29E-35$ in sheep and $p=3.89E-12$ in goats) as illustrated in Figure 2c. The spatial distribution of candidate alleles is also represented for orthologous candidate genes in Figure S8.

In total we detected 9875 and 5547 outlier variants with at least one of the two approaches (324 and 56 common to both methods), corresponding to 291 and 346 genomic regions (4 and 5 regions common to both methods, Table 2) in sheep and goats respectively. These regions are distributed over all chromosomes (see Figure S9–S12). Among these regions, 19 were longer than 200 kb in sheep (largest 648 kb) with 16 such loci in goats (largest 427 kb). A total of 218 genes in sheep and 212 genes in goats were associated with at least one environmental variable (Table S4). Among these genes, eight were detected by both methods in sheep, related to rainfall (prec4) and temperature (tmean7). Among them, *MC5R* on chromosome 23, showed 2 missense alleles, which were present in the driest zones of the sampling area (Figure 2c). In goats, two genes were detected by both methods, that is, *CDH2*, located on chromosome 24, associated with rainfall seasonality (bio15) and *PXK* on chromosome 22, associated with altitude.

Most of the variants detected were intergenic (63% and 67% for sheep and goats respectively) and intronic (27% and 25% for sheep and goats respectively). Sixty-eight outlier SNPs were exonic (21 missenses and 47 synonymous) in sheep, with only one (a single missense mutation) in goats (Table S5). These proportions were similar to that of the entire genome for sheep but there were more intronic and intergenic variants among outliers for goats (χ^2 test, $df=9$, $p<.05$).

When considering the possible drivers of selection, the number and proportion of candidate genomic regions and candidate genes associated with most of the environmental drivers differed between species (Figure 3, χ^2, $p<.001$). Genes putatively under

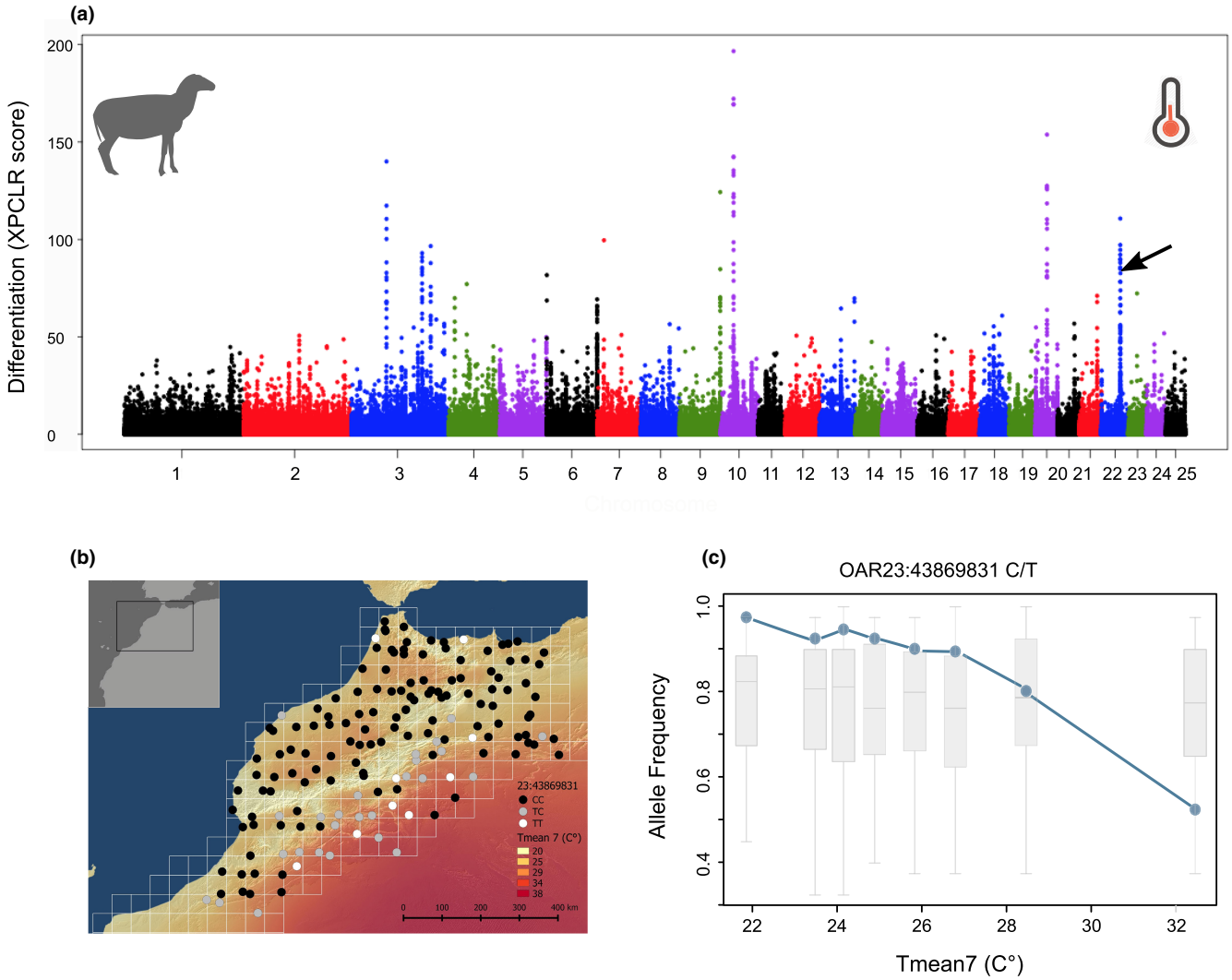

FIGURE 2 Genetic differentiation associated with temperature in sheep. (a) Manhattan plot of XP-CLR scores across autosomes in relation to the mean temperature of July (tmean7). The peak on chromosome 23 (arrow) comprises a missense SNP (OAR23:43869831 C/T). (b) Geographical variations of genotypes at this locus reported on the map of temperature (tmean7) variations. (c) Variation of allele frequency at this locus for groups of 20 sheep along the temperature gradient. The grey box-plots represent neutral variations of allelic frequencies for 100 variants randomly selected over the genome. This SNP was selected for displaying because it is a missense variant on the MC5R gene.

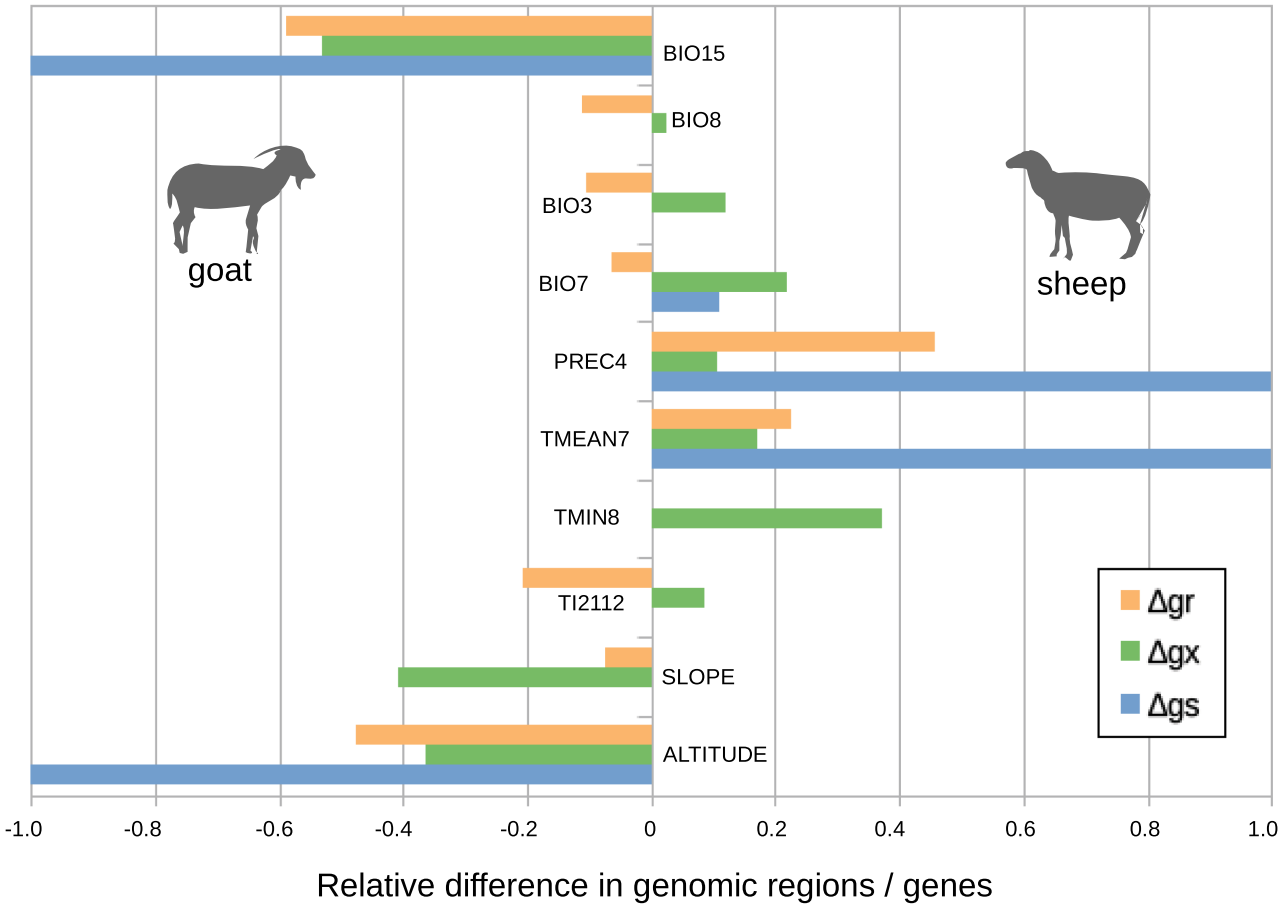

FIGURE 3 Relative difference between sheep and goats for the number of outlying genomic regions (Δgr) and genes detected with XP-CLR/*F*st (Δgx) and Samßada (Δgs). The three indexes were calculated as the difference between the number of regions/genes found in sheep and goats divided by the total number of regions/genes. They vary between −1 and +1: regions/genes associated with an environmental parameter only in goats or sheep respectively. See Table 2 for the codes referring to environmental parameters. The number of genomic regions and genes putatively under selection due to environmental variables differed for each species (χ^2 tests, df = 9, $p < .001$). See Table S3 for the list of genes.

selection were most related to altitude, slope and rainfall seasonality for goats, and summer temperature and spring rainfall for sheep (Table S4 and Figure S3).

3.3 | Patterns of allelic variation at selected loci

We used an exploratory approach to evaluate genomic responses to environmental variation by characterizing the variation of allele frequencies along environmental variables for the most differentiated SNP locus within each candidate selected genomic region for the 10 environmental parameters (Figure 4; Figures S13–S16). We categorized each variation profile as gradual (i.e. linear) or punctual at one or both extremities of the gradient (Figure 4a,b). For both species, all variation patterns were observed, with 69% of outliers in sheep and 80% of outliers in goats corresponding to punctual shifts occurring at extreme environmental values (i.e. low, high or both, e.g., Figure 4; Figures S13–S16). For most of the variables the distribution of patterns across categories was similar, except for temperature seasonality (bio7) and rainfall (prec4) that had a higher proportion of candidates with a linear variation in sheep. Thus, different environmental gradients were causing the strongest genetic response in sheep and goats, resulting in generally different candidate genes/genomic regions being impacted and patterns of variation along the gradients that vary by gene and species.

3.4 | Gene ontology enrichment analysis

We performed Gene Ontology (GO) enrichment analyses with GOrilla (Eden et al., 2009) followed by a clustering with REVIGO (Supek et al., 2011) using the entire set of candidate genes. In sheep, four GO categories were enriched in genes associated with altitude, including two GO terms linked to lung development related to the gene *NFIB*. Fourteen GO terms were enriched in genes associated with temperature (tmean7) and annual temperature range

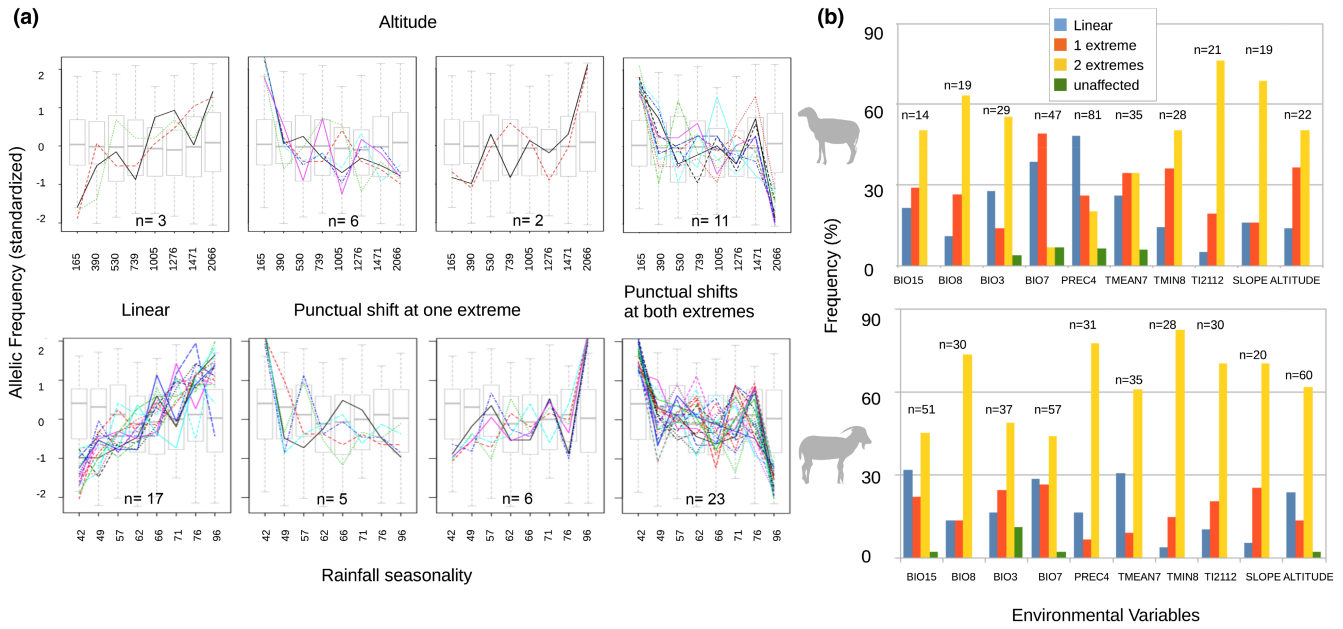

FIGURE 4 Patterns of allele frequencies variations for candidate variants along environmental gradients. (a) Examples for candidate variants related to altitude in sheep (upper charts) and rainfall seasonality in goats (lower charts). These two variables were chosen as examples. For each SNP, the variation profile of allele frequency (reduced centred values) is assigned to one of the following categories: linear variation, uniform with punctual shifts at one or both extremes of the gradient (see methods). The grey box-plots represent neutral variations of allelic frequencies for a set of random variants. (b) Distribution of patterns of allele frequency variation for the different environmental variables for sheep (upper) and goats (lower). See Table 2 for the definition of environmental variations.

(bio7), depending on one main category related to the regulation of cAMP biosynthesis. One GO term was associated with temperature (tmin8), rainfall seasonality (bio15) and solar radiation (ti2112) (Table S6). For goats, 24 GO terms were enriched in genes associated with temperature seasonality (bio8), in five distinct categories, that is, stem cell division, regulation of hydrolase activity, cellular component morphogenesis, locomotory behaviour and neuromuscular process. In addition, 21 GO terms associated with temperature (tmin8) corresponded mainly to the regulation of transmembrane cellular transport. We also found four GO terms associated with isothermality (bio3) and related to the cell response to superoxide and mitochondrial transport. Two GO terms were associated with rainfall seasonality (bio15). Slope, annual temperature range (bio7), rainfall (prec4) and solar radiation (ti2112) were associated with one GO term each (Table S7). No GO terms overlapped between the two species.

3.5 | Orthologous genomic regions related to environmental variations in sheep and goats

We mapped all variants identified in sheep to the goat genome and vice versa but found no orthologous outlier SNPs between species. Furthermore, of the 430 putatively adaptive genes in sheep and goats, only six were orthologous (i.e. *BICC1*, *CA2*, *DENND4C*, *DUSP2*, *FAT3*, *TCP11L2*; Table S4), which is very low but significantly different from random (Fisher's Exact Test, $p = .0071$, see Table S8). Additionally, only a single genomic region was orthologous in both species where 26 variants associated with winter sunshine duration (ti2112) in sheep were within 100 kb of 13 variants identified for the same variable in goats. It is a region covering the *CA2* gene where there were three downstream, three intronic and six upstream *CA2* variants in the sheep genome. Otherwise, *CA2* and *BICC1* were the only genes to be associated with the same variables (i.e. solar radiation for *CA2* and rainfall in April for *BICC1*) in both species. *CA2* had one downstream outlier variant in goats and four intronic, three downstream and six upstream outlier variants in sheep. In contrast, *BICC1* had 24 intronic outlier variants in goats and 2 intronic variants in sheep. When considering the four remaining orthologous genes, outlier variants in *FAT3* (1 in sheep and 16 in goats) and *TCP11L2* (four in sheep and four in goats) were all intronic in both species. The variants associated with *DUSP2* (1 in sheep and 11 in goats) were all upstream gene variants. Lastly, the variants associated with *DENND4C* (5 in sheep and 11 in goats) were mainly intronic (five and three, respectively, in sheep and goats), upstream or downstream (two in sheep and four in goats) and two UTR variants in goats.

4 | DISCUSSION

4.1 | Performance of the experimental design for assessing convergent adaptive trajectories

Our study was specifically designed to assess the diversity and uniqueness of the adaptive response at two levels: (i) within-species by characterizing patterns of candidate adaptive genetic variation along environmental gradients, (ii) between-species by comparing the

candidate adaptive responses of two closely related species along the same environmental gradients. Analysing signatures of the adaptive response of genomes along environmental gradients has mainly been studied via discrete sampling in contrasting extreme environments (Riesch et al., 2018), leaving large parts of these gradients unexplored. Our sampling strategy maximized the coverage of multiple gradients by sequencing 160 whole genomes per species across an environmentally heterogeneous region of approximately 400,000 km^2. As a consequence, the environmental clines explored, covering wide ranges of variations, could not be superimposed on latitudinal gradients or colonization routes, which are potential sources of confounding effects. We found high genomic variation in both species (38.6 and 31.7 million SNPs in sheep and goat respectively), with no significant genetic structure over the area studied and among breeds (Figure 1; Figures S3, S4). This suggests that no strong bottlenecks would have occurred during breed evolution and that gene-flow has been maintained (Benjelloun et al., 2015) and/or related to successive waves of colonization along the Mediterranean routes of domestication (Pereira et al., 2009). This absence of significant structure has a clear advantage when looking for selective sweeps due to local adaptation, allowing us to minimize confounding demographic factors (Holderegger et al., 2008).

4.2 | Limitations in detecting signatures of selection

Whole genome sequences (WGS) together with the use of many dedicated and powerful algorithms have greatly increased the power to identify selection signatures (Benjelloun et al., 2019). However, the identification of causal mutations is still challenging even when using WGS data and applying appropriate tools/algorithms. Indeed not all SNPs are typed given the quality thresholds defined in different steps of variant calling and filtering. Moreover, annotation of intergenic regulatory sequences is still poor in non-human species and this causes substantial limitations in explaining selection signatures identified in intergenic sequences despite the strong evidence they are under selection. Another limitation is also encountered when doing comparisons between results from different studies using different datasets, algorithms and approaches, for example, difference in threshold of distance of genes to outlier SNPs; difference between Individual versus breed distribution to define the level of environmental variables; the type of data when comparing WGS versus SNP arrays or other genome sampling approaches. This last issue was also raised by previous studies (Alexander et al., 2022).

4.3 | Genetic differentiation along environmental gradients

Changes in adaptive allele frequencies in relation to the spatial variation of their relative fitness are usually expected to be either linear or sigmoidal (Novembre & Di Rienzo, 2009), corresponding to patterns of either gradual or punctual genetic variation respectively. The combination of methods used here (i.e. Samβada, XP-CLR and *Fst*) should in principle allow the detection of both types of patterns. We detected putatively adaptive genomic regions showing either a gradual or more frequently punctual frequency variation along multiple environmental gradients (69% in sheep and 80% in goats). This may be explained by stronger selection at environmental extremes, in accordance with the expectation that extreme values of environmental parameters are responsible for higher functional stress and thus stronger selective pressures (Grant et al., 2017). For some variables (i.e. temperature seasonality and rainfall), the most represented patterns were different according to the species. Together with the fact that the number of candidate regions/genes is very different between species for 4 of the environmental parameters, this suggests that different environmental gradients are causing the strongest genetic response in sheep and goats, resulting in generally different candidate genes/genomic regions impacted and patterns of variation along the gradients that vary by gene and species.

As our sampling maximized the spatial spread of individuals and the representativity of diverse breeds, the impact of gene flow and drift on shaping the patterns of allelic variations at candidate SNPs might be reduced with regard to that of selection. However, a significant influence of neutral drivers could exist, for example, when a breed is overrepresented at one extremity of a gradient (e.g. abrupt change at one extreme related to drift leading the selected allele to fixation faster within this breed). In our study, a potential risk exists with the D'man sheep which is restricted to desert areas, but this breed also shows specific traits in harsh desert conditions (Lahlou-Kassi et al., 1988), which could be responsible for candidate regions identified.

The frequent punctual variation of allelic frequency at one extremity of the gradient would explain why Samβada identified substantially fewer SNPs than XP-CLR/*Fst*, even if we cannot exclude a lower power for this approach.

4.4 | Candidate targets of selection and adaptive pathways

Spatial patterns of allelic frequencies reflect variation in adaptive phenotypes (Novembre & Di Rienzo, 2009). Thus identifying genomic signatures of local adaptation and their associated environmental drivers is a first step towards understanding the genetic architecture underlying adaptive traits, and identifying the conditions responsible for divergence (Tiffin & Ross-Ibarra, 2014). We identified candidate genomic regions mainly responding to average temperature and rainfall, and to a lesser extent to seasonality, altitude, solar radiation and topography. Most of the selective sweeps identified affected only non-coding variants as reported for other species (Carneiro et al., 2014; Ai et al., 2015; Boitard et al., 2016). Functional annotation of non-coding variants is challenging as the gene(s) they affect can be located far away and might not even be the physically closest one. When located close to genes, they could

be influencing transcription (Ward & Kellis, 2012). Interestingly, Naval-Sanchez et al. (2018) showed that domestication and subsequent selection have preferentially impacted proximal gene regulatory elements in sheep. Non-coding variants could also constitute a part of sequences regulating translation, stability and localization (Chan et al., 2010) (GTEx Consortium, 2017). Moreover, they could contribute more generally to functional regions that are not closely located to genes under regulation such as regulatory non-coding RNAs (e.g. miRNAs) (Dunham et al., 2012; Johnson & Voight, 2018; Noonan & McCallion, 2010). These observations are in accordance with theories suggesting that evolution may be facilitated more through changes in gene regulation than through changes in gene sequence (Carroll, 2008). The current lack of knowledge on the identification and role of regulatory sequences in non-human mammals limits the interpretation of selective sweeps in adaptation to eco-climatic pressures and emphasizes the need for a comprehensive and integrated resource on functional elements of animal genomes such as the FAANG initiative (https://www.animalgenome.org/community/FAANG/index).

We identified multiple gene-bearing regions affected by selective sweeps and searched for insight in their functional meaning by analysing enrichment of GO terms. In sheep, adaptation to altitude involves GO terms related to the growth and differentiation of lung cells whose proliferation is related to altitude (Heath et al., 1976) and hypoxia (Uhlik et al., 2005), and GO terms related to altitude also include heart rate regulation via the gene *TPM1*. More specifically, the gene *MCM3* was a top-candidate gene related to altitude. It regulates the activity of *HIF-1* (hypoxia-inducible factor 1) (Hubbi et al., 2011; Semenza, 2011), which has been shown with its paralogue *HIF-2* (*EPAS1*) to be under selection at high altitude in several species including sheep and goats (Beall et al., 2010; Gou et al., 2014; Pan & Shen, 2017; Simonson et al., 2010; Wei et al., 2016; Yi et al., 2010). Even if the environmental pressures related to altitude encountered here (highest altitudes sampled at 2300 m with likely transhumance—summer moving of flocks—up to 3300 m ASL) might differ from that of previous studies targeting higher altitudes (e.g. Tibetan populations, >3000 m ASL), our findings illustrate that different genes of the same pathway are candidates for different adaptive responses to similar environmental pressures. This illustrates an interesting case with similar adaptations by selection on different genes of the same network at the intraspecific level (i.e. sheep).

In goats, the GO term "Regulation of respiratory gaseous exchange" (GO:0043576) was related to temperature annual range via the genes *PASK* and *GLS*. This is consistent with a thermoregulation via panting in goats experiencing highly fluctuating temperatures (Dmiel & Robertshaw, 1983); (Baker, 1989). Indeed, goats have two different thermoregulation mechanisms, panting or sweating, according to the level of dehydration (Dmiel & Robertshaw, 1983) (Baker, 1989) depending on the breed (Dmiel et al., 1979). This result might be partly driven by the presence of Draa goats in our samples (7 Draa among the 20 goats in the group with the highest annual temperature range) even in the absence of breed-specific genetic clustering. The same GO category has already been shown to be enriched for candidate genes in this breed which is adapted to oases and desert areas in Morocco (Benjelloun et al., 2015).

GO terms linked to cellular component morphogenesis and stem cell division and differentiation were associated with temperature seasonality. Among the outlier genes identified, *PAFAH1B1* is involved in stem cell morphogenesis, *ABL1* in B cell differentiation, *NOTCH1* in angiogenesis (Scott et al., 2015) and *ALS2* in several cellular processes, for example, endocytosis, development of axons and dendrites (Ratti & Berry, 2016). The distribution patterns of pathogens, parasites and their vectors are strongly dependent on temperature and humidity (Altizer et al., 2006; Vajana et al., 2018), with, for example, a higher prevalence of the liver fluke under wet conditions or of several ectoparasites under hot and humid environments (Taylor, 2012). In this context, selection signatures for genes involved in stem cells and immune B cell differentiation are likely to be related to environmental variation in pathogen selective pressures.

Another interesting case is that of the "Regulation of bone remodelling" category (GO:0046850), which is enriched in genes putatively selected in goats in association with solar radiation, namely *LRP5* and *CA2*. The latter, which is one of the six genes putatively selected in both sheep and goats, was also associated with solar radiation in sheep. We can, therefore, predict a possible role of these genes in preventing the decrease in bone density following lower levels of vitamin D at low UV radiation exposure. Indeed, vitamin D levels which affect bone density (Reid et al., 2014) depend on food intake and cutaneous production in sheep and goats (Kohler et al., 2013). Similarly, BICC1 that was identified as an outlier in association with rainfall in both sheep and goats has been proved to be a genetic determiner of osteoblastogenesis and bone mineral density (Liu et al., 2022; Mesner et al., 2014).

Besides these two orthologous candidate genes (*CA2* and *BICC1*), the remaining four genes are involved in many biological processes. *DENND4C* which was associated with isothermality in sheep and altitude in goats has been proved to be a key component for insulin-stimulated translocation of the glucose transporter GLUT4 to the cell surface in fat and muscle cells, which is the basis for insulin-stimulated glucose transport (Sano et al., 2011). *DUSP2* was putatively associated with isothermality in sheep and slope in goats. Its downregulation has been proved to be associated with the enhancement of immune functions of monocytes (Wang et al., 2015). *FAT3* which was identified in putative association with rainfall in sheep and altitude in goats can be involved in the regulation of dendrite development, neuron migration and retina layer formation (Cheng et al., 2016). Lastly, *TCP11L2*, which was putatively associated with temperature variation in both species (bio7 in sheep and bio3 in goats), has recently been proved to be involved in muscle-derived satellite cell (MDSC) differentiation and migration (Li et al., 2020). These cells mediate the life-long maintenance of muscle tissues and are activated by stimuli such as physical trauma or growth signals under homeostatic conditions (Dumont et al., 2015). We can thus hypothesize a possible role of this gene in facing constraints induced by temperature variation in sheep and goats. However, we cannot predict the exact roles of these genes in sheep and goat adaptation

and this would require further investigations using approaches adapted to some specific associations.

For example, future research would require studying the variation of gene expression along the environmental gradients (e.g. via RNA sequencing) for validating candidate genes. Transplant experiments would also be of interest for future validations. Even if such standardized transplants have not been performed so far, human displacements of animals in Morocco have shown that they exhibit lower performance outside the climate of the breed's cradle, and several breeds are known to have specific production traits to a limited range of environmental conditions (e.g. for D'man sheep: Lahlou-Kassi et al., 1988).

4.5 | Co-occurrence of anthropogenic and natural drivers of selection

Besides natural selection, human-induced selection can also have a strong impact on the genomes of domestic animals and disentangling both is not straightforward. In sheep, *MC5R* and *MC2R*, which are located in the largest selective sweep identified (~650 kb), are related to temperature and rainfall. In *MC5R*, we found a non-synonymous polymorphism spatially segregated in regions with the lowest rainfall and highest temperature. *MC5R* is expressed in a variety of peripheral tissues, especially in the skin (Switonski et al., 2013), and is a candidate gene for fat deposition in domestic animals because it down-regulates leptin secretion (Kristiansen & Mandrup-Poulsen, 2005). Fat deposition, which is advantageous under dry conditions, could be directly selected by climatic pressures but also by anthropogenic selection in breeds such as the D'man, which is currently predominant in the dry and warm regions of Morocco. We also report several genes associated with rainfall, which may have been anthropogenically selected. They include *RXFP2* involved in horn morphology (Johnston et al., 2011) and reproductive success (Johnston et al., 2013), *RANBP17* associated with sperm maturation (Koch et al., 2000) (Bao et al., 2011), *SEMA5A* related to milk production traits and mammary gland inflammation in cattle (Sugimoto et al., 2006) (Pareek, 2010), and *RBM19* involved in mammary gland morphology in cattle (Pausch et al., 2012). However, we currently cannot disentangle whether these observations are a combined effect of human-induced and natural selection on the same traits, or the result of a confounding effect based on breeds adapted to a specific climate.

4.6 | Differences in adaptive pathways between species

Previous studies report that adaptation may involve convergence at different biological levels, from the same mutation to the production of different phenotypes for fulfilling the same organismal function (Losos, 2011) (Riesch et al., 2018). Out of the 218 and 212 putatively adaptive genes identified in sheep and goats, respectively, only six were orthologous between species. Similarly, among these sets of genes, only 18 have been reported to be under selection in previous studies describing genetic bases of local adaptation to climatic and/or altitude conditions in sheep and goats (Table S9). However, each study has used different approaches to define genes under selection related to environmental variation as stated above (e.g. WGS vs. SNP arrays; threshold of distance of genes to outlier SNPs; individual vs. breed distribution to define the level of environmental variables; and so on). Otherwise, all GO categories identified for the same environmental variable were different with no clear links between them. This supports the predominance of distinct candidate genes and biological pathways for adaptation to the same environment in these species. This low level of gene reuse (i.e. independent selection of the same gene in distinct evolutionary lineages) could result from the occurrence of false negative/positive drowning out the true signal, for example, the occurrence of large mutational target (Conte et al., 2012) or the level of shared standing variation (Bohutínská et al., 2021; Chaturvedi et al., 2022; Montejo-Kovacevich et al., 2021).

The different adaptive pathways may also result from environmental drivers shaping differentially adaptive divergence in each species. While sheep and goats were sampled in the same geographical area and along the same environmental gradients, we found that most drivers impacted differently their genomes. The candidate variants and genomic regions were most related to altitude, slope and rainfall seasonality for sheep, and summer temperature and spring rainfall for goats. However, these variables are highly correlated with other ones which makes it difficult to identify the true drivers of selection. For example, altitude is correlated with many temperature variables ($|r| \geq .8$) (Table S3) and we could not exclude that another environmental force led to the signature of selection detected for altitude. Otherwise, the different relative impact of drivers of selection between species could be explained by many factors. First, despite the fact that sheep and goats are raised under similar eco-climatic conditions in Morocco, they have morphological and physiological differences that may affect the way animals face climatic constraints (e.g. goats are hairy and sheep woolly). Secondly, the two species have different diets under similar eco-climatic conditions and can be subjected to different husbandry practices, which can lead to different impacts of environmental pressures and thus different selective responses. Finally, the landscape genomics approach does not consider breed-specific effects but detects marks of adaptation common to several breeds, preventing from detecting some breed-specific adaptation within each species, such as the different mechanisms of thermoregulation used by two Moroccan goat breeds to face warm/desert environment (Benjelloun et al., 2015).

Besides the focus on inter-specific gene reuse, our study also illustrates the occurrence of similar adaptations by selection on different genes of the same network at the intraspecific level. For example, the gene *MCM3* associated with altitude in Moroccan sheep has been shown to regulate *HIF* genes regulating haemoglobin concentration (Hubbi et al., 2011) (Semenza, 2011). We did not find

evidence for selection of *HIF* paralogues here, but they are widely reported to be under selection at high altitude, especially in Tibetan populations (Simonson et al., 2010) (Yi et al., 2010) (Wei et al., 2016) (Song et al., 2016) (Gou et al., 2014). Other candidate genes involved in the *HIF* pathway have already been found under selection in sheep at higher altitudes than the present study (>3300 m ASL; (Yang et al., 2016)).

5 | CONCLUSION

Our study documented well-supported candidate genes and their possible role in adaptation and found that environmental drivers would have a differential impact on sheep and goat genomes both in terms of magnitude and type of response (pattern of variation along the gradients). This could at least partly explain the differences in adaptive pathways observed between sheep and goats. This contrasts also with a high level of gene reuse observed between these two species with regard to key changes related to the early phases of domestication (Alberto et al., 2018). This situation would result from similar drivers of selection during the early phase of domestication, targeting traits offering fewer options to selection (smaller mutational targets) in comparison with adaptation to ecology and climate for which the pressure can be theoretically softer than domestication.

From a more global perspective, this study analysed the variation patterns of allele frequencies of candidate SNPs across environmental gradients and described typical empirical examples of clinal variation over many continuous environmental gradients. The diversity of genetic responses reported in "adaptive landscapes" would open new horizons for understanding genomic mechanisms of adaptation in mammalian species and beyond.

AUTHOR CONTRIBUTIONS

The paper represents the joint efforts of several research groups, most of whom were involved in the NEXTGEN project (coordinated by P.T.). P.T., F.P., E.C. and R.N. designed the study. P.T. and F.P. supervised the joint work in NEXTGEN. B.B., A.S., A.C., M.W.B., S.J., R.N., P.A.-M. and P.F. supervised the work of their research group. B.B. and A.C. coordinated the collection of the samples. A.A. and S.E. produced whole-genome sequences. S.J. developed the sampling design and supervised Geographic Information Systems. I.S., L.Cl., E.C., S.E., F.Bo. and B.B. contributed to bioinformatic analyses. B.B., K.L., S.S., B.S., F.J.A., F.Bo., P.L., L.Co., F.Bi., F.P. and M.B. did the analyses. B.B., K.L., S.J., F.Bo. and F.P. produced the figures. B.B., K.L. and F.P. wrote the text with input from all authors and especially F. Bo, S.J. and P.O.-t.W.

ACKNOWLEDGEMENTS

This work was supported by the European Union 7th framework project NEXTGEN (Grant Agreement no. 244356, coordinated by P.T.), the FACCE ERA-NET Plus project CLIMGEN (grant ANR-14-JFAC-0002-01, coordinated by M.W.B.), the LabEx OSUG@2020 (Investissements d'avenir – ANR10LABX56), the Wellcome Trust [WT108749/Z/15/Z] and the European Molecular Biology Laboratory (I.S., L.Cl., P.F.). The SRTM V3.0 data product used was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool. The authors are grateful to Abdelmajid Bechchari, Mustapha Ibnelbachyr, Mouad Chentouf, Mohammed BenBati, Rachid Hadria, Mouloud Laghmir, Lahbib Haounou, El Mekki Hafiani, El Mustapha Sekkour, Ali Dadouch, Abdelkader Lberji, Chérif Errouidi, Mohamed Ouatiq, Btissam Hilal and Mustapha Sadek who contributed to the field sampling. They also thank Ludovic Orlando for fruitful discussions.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data are accessible at ftp://ftp.ebi.ac.uk/pub/databases/nextgen/.

DATA AVAILABILITY STATEMENT

The variant calls and genotype calls used in this paper are archived in the European Variation Archive with accession ID ERZ019290 for sheep and ID ERZ020631 For goats. The data are accessible at ftp://ftp.ebi.ac.uk/pub/databases/nextgen/. The code to reproduce analysis is also available on the Nextgen ftp site and some scripts are available on GitHub https://github.com/BadrBenjelloun/Nextgen_Local_adaptation_paper_202308.git. This paper is one of the main papers produced within the frame of the EU FP7 NextGen project and result from a joint effort of scientists from all contributing countries, who have been involved in sampling, analysing data and writing the manuscript. The paper is led as first author by a scientist from the country where the genetic samples come from.

ORCID

Badr Benjelloun https://orcid.org/0000-0003-4310-7281
Kevin Leempoel https://orcid.org/0000-0001-7335-7930
Sylvie Stucki https://orcid.org/0000-0002-1624-1576
Pablo Orozco-terWengel https://orcid.org/0000-0002-7951-4148
Adriana Alberti https://orcid.org/0000-0003-3372-9423
Licia Colli https://orcid.org/0000-0002-7221-2905
Stéphane Joost https://orcid.org/0000-0002-1184-7501

REFERENCES

Ai, H., Fang, X., Yang, B., Huang, Z., Chen, H., Mao, L., Zhang, F., Zhang, L., Cui, L., He, W., Yang, J., Yao, X., Zhou, L., Han, L., Li, J., Sun, S., Xie, X., Lai, B., Su, Y., … Huang, L. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nature Genetics*, *47*(3), 217–225.

Alberto, F. J., Boyer, F., Orozco-terWengel, P., Streeter, I., Servin, B., de Villemereuil, P., Benjelloun, B., Librado, P., Biscarini, F., Colli, L., Barbato, M., Zamani, W., Alberti, A., Engelen, S., Stella, A., Joost, S., Ajmone-Marsan, P., Negrini, R., Orlando, L., … Pompanon, F. (2018). Convergent genomic signatures of domestication in sheep and goats. *Nature Communications*, *9*(1), 813.

Alexander, A., Robbins, M. B., Holmes, J., Moyle, R. G., & Peterson, A. T. (2022). Limited movement of an avian hybrid zone in relation to regional variation in magnitude of climate change. *Molecular Ecology*, *31*, 6634–6648. https://doi.org/10.1111/mec.16727

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484.

Baker, M. A. (1989). Effects of dehydration and rehydration on thermoregulatory sweating in goats. *Journal of Physiology (London)*, *417*, 421–435.

Bao, J., Wu, Q., Song, R., Jie, Z., Zheng, H., Xu, C., & Yan, W. (2011). RANBP17 is localized to the XY body of spermatocytes and interacts with SPEM1 on the manchette of elongating spermatids. *Molecular and Cellular Endocrinology*, *333*(2), 134–142.

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 (HIF2 alpha) 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.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*(1), 289–300.

Benjelloun, B., Alberto, F. J., Streeter, I., Boyer, F., Coissac, E., Stucki, S., BenBati, M., Ibnelbachyr, M., Chentouf, M., Bechchari, A., Leempoel, K., Alberti, A., Engelen, S., Chikhi, A., Clarke, L., Flicek, P., Joost, S., Taberlet, P., Pompanon, F., & NextGen Consortium. (2015). Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS data. *Frontiers in Genetics*, *6*, 107.

Benjelloun, B., Boyer, F., Streeter, I., Zamani, W., Engelen, S., Alberti, A., Alberto, F. J., BenBati, M., Ibnelbachyr, M., Chentouf, M., Bechchari, A., Rezaei, H. R., Naderi, S., Stella, A., Chikhi, A., Clarke, L., Kijas, J., Flicek, P., Taberlet, P., & Pompanon, F. (2019). An evaluation of sequencing coverage and genotyping strategies to assess neutral and adaptive diversity. *Molecular Ecology Resources*, *19*, 1497–1515. https://doi.org/10.1111/1755-0998.13070

Böhner, J., McCloy, K. R., & Strobl, J. (2006). SAGA: Analysis and modelling applications. *Göttinger Geographische Abhandlungen*, *115*, 130.

Bohutínská, M., Vlček, J., Yair, S., Laenen, B., Konečná, V., Fracassetti, M., Slotte, T., & Kolář, F. (2021). Genomic basis of parallel adaptation varies with divergence in Arabidopsis and its relatives. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(21), e2022713118. https://doi.org/10.1073/pnas.2022713118

Boitard, S., Rodríguez, W., Jay, F., Mona, S., & Austerlitz, F. (2016). Inferring population size history from large samples of genome-wide molecular data – An approximate Bayesian computation approach. *PLoS Genetics*, *12*(3), e1005877.

Browning, B. L., & Browning, S. R. (2013). Detecting identity by descent and estimating genotype error rates in sequence data. *American Journal of Human Genetics*, *93*(5), 840–851.

Cao, Y.-H., Xu, S.-S., Shen, M., Chen, Z.-H., Gao, L., Lv, F.-H., Xie, X.-L., Wang, X.-H., Yang, H., Liu, B.-B., Zhou, P., Wan, P.-C., Zhang, Y.-S., Yang, J.-Q., Pi, W.-H., Hehua, E. E., Berry, D. P., Barbato, M., Esmailizadeh, A., … Li, M.-H. (2021). Historical introgression from wild relatives enhanced climatic adaptation and resistance to pneumonia in sheep. *Molecular Biology and Evolution*, *38*(3), 838–855.

Carneiro, M., Rubin, C.-J., Di Palma, F., Albert, F. W., Alfoeldi, J., Barrio, A. M., Pielberg, G., Rafati, N., Sayyab, S., Turner-Maier, J., Younis, S., Afonso, S., Aken, B., Alves, J. M., Barrell, D., Bolet, G., Boucher, S., Burbano, H. A., Campos, R., … Andersson, L. (2014). Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science*, *345*(6200), 1074–1079.

Carroll, S. B. (2008). Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell*, *134*(1), 25–36.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Jr., Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Petrov, D., Jónsson, B., Schluter, D., Bell, M. A., & Kingsley, D. M. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science*, *327*(5963), 302–305.

Chaturvedi, S., Gompert, Z., Feder, J. L., Osborne, O. G., Muschick, M., Riesch, R., Soria-Carrasco, V., & Nosil, P. (2022). Climatic similarity and genomic background shape the extent of parallel adaptation in *Timema* stick insects. *Nature Ecology & Evolution*, *6*, 1952–1964. https://doi.org/10.1038/s41559-022-01909-6

Chen, H., Patterson, N., & Reich, D. (2010). Population differentiation as a test for selective sweeps. *Genome Research*, *20*(3), 393–402.

Cheng, H., Burroughs-Garcia, J., Birkness, J. E., Trinidad, J. C., & Deans, M. R. (2016). Disparate regulatory mechanisms control Fat3 and P75NTR protein transport through a conserved Kif5-interaction domain. *PLoS ONE*, *11*(10), e0165519. https://doi.org/10.1371/journal.pone.0165519

Christin, P. A., Weinreich, D. M., & Besnard, G. (2010). Causes and evolutionary significance of genetic convergence. *Trends in Genetics*, *26*(9), 400–405.

Conte, G. L., Arnegard, M. E., Peichel, C. L., & Schluter, D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 5039–5047.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498.

Dmiel, R., & Robertshaw, D. (1983). The control of panting and sweating in the black Bedouin goat – A comparison of two modes of imposing a heat load. *Physiological Zoology*, *56*(3), 404–411.

Dmiel, R., Robertshaw, D., & Choshniak, I. (1979). Sweat gland secretion in the black Bedouin goat. *Physiological Zoology*, *52*(4), 558–564.

Dong, Y., Xie, M., Jiang, Y., Xiao, N., Du, X., Zhang, W., Tosser-Klopp, G., Wang, J., Yang, S., Liang, J., Chen, W., Chen, J., Zeng, P., Hou, Y., Bian, C., Pan, S., Li, Y., Liu, X., Wang, W., … Wang, W. (2013). Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology*, *31*(2), 135–141.

Dumont, N. A., Bentzinger, C. F., Sincennes, M. C., & Rudnicki, M. A. (2015). Satellite cells and skeletal muscle regeneration. *Comprehensive Physiology*, *5*(3), 1027–1059. https://doi.org/10.1002/cphy.c140068

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B.-K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, *10*, 48.

Elmer, K. R., & Meyer, A. (2011). Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution*, *26*(6), 298–306.

Flicek, P., Ahmed, I., Amode, M. R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., Gil, L., García-Girón, C., Gordon, L., Hourlier, T., Hunt, S., Juettemann, T., Kähäri, A. K., Keenan, S., … Searle, S. M. J. (2012). Ensembl 2013. *Nucleic Acids Research*, *41*(D1), D48–D55.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & Francois, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973–983.

Frichot, E., Schoville, S. D., Bouchard, G., & Francois, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, *30*(7), 1687–1699.

Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing*. arXiv:1207.3907v2 [q-bio.GN].

Gou, X., Wang, Z., Li, N., Qiu, F., Xu, Z., Yan, D. W., Yang, S. L., Jia, J., Kong, X. Y., Wei, Z. H., Lu, S., Lian, L., Wu, C., Wang, X., Li, G., Ma, T., Jiang, Q., Zhao, X., Yang, J., … Li, Y. (2014). Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Research*, *24*(8), 1308–1315.

Grant, P. R., Grant, B. R., Huey, R. B., Johnson, M. T. J., Knoll, A. H., & Schmitt, J. (2017). Evolution caused by extreme events. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*(1723), 20160146.

GTEx Consortium GT: Lead analysts. (2017). Genetic effects on gene expression across human tissues. *Nature*, *550*, 204–213.

Harris, R. S. (2007). *Improved pairwise alignment of genomic DNA*. PhD dissertation. The Pennsylvania State University. 74 pp.

Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., Jansen van Vuuren, B., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskoug, V., Nguyen, T. T., & Couloux, A. (2012). Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *Comptes Rendus Biologies*, *335*(1), 32–50.

Heath, D., Smith, P., & Harris, P. (1976). Clara cells in llama. *Experimental Cell Biology*, *44*(2), 73–82.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978.

Holderegger, R., Herrmann, D., Poncet, B., Gugerli, F., Thuiller, W., Taberlet, P., Gielly, L., Rioux, D., Brodbeck, S., Aubert, S., & Manel, S. (2008). Land ahead: Using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, *1*(2), 273–283.

Hubbi, M. E., Luo, W., Baek, J. H., & Semenza, G. L. (2011). MCM proteins are negative regulators of hypoxia-inducible factor 1. *Molecular Cell*, *42*(5), 700–712.

Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D., & Serra, L. (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science*, *287*(5451), 308–309.

Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J. F., Faraut, T., Wu, C., Muzny, D. M., Li, Y., Zhang, W., Stanton, J. A., Brauning, R., Barris, W. C., Hourlier, T., Aken, B. L., Searle, S. M. J., Adelson, D. L., Bian, C., Cam, G. R., … Dalrymple, B. P. (2014). The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, *344*(6188), 1168–1173.

Johnson, K. E., & Voight, B. F. (2018). Patterns of shared signatures of recent positive selection across human populations. *Nature Ecology & Evolution*, *2*(4), 713–720.

Johnston, S. E., Gratten, J., Berenos, C., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M., & Slate, J. (2013). Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*, *502*(7469), 93–95.

Johnston, S. E., McEwan, J. C., Pickering, N. K., Kijas, J. W., Beraldi, D., Pilkington, J. G., Pemberton, J. M., & Slate, J. (2011). Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, *20*(12), 2555–2566.

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071.

Joost, S., Bonin, A., Bruford, M. W., Despres, L., Conord, C., Erhardt, G., & Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: Towards a landscape genomics approach to adaptation. *Molecular Ecology*, *16*(18), 3955–3969.

Kamberov, Y. G., Wang, S., Tan, J., Gerbault, P., Wark, A., Tan, L., Yang, Y., Li, S., Tang, K., Chen, H., Powell, A., Itan, Y., Fuller, D., Lohmueller, J., Mao, J., Schachar, A., Paymer, M., Hostetter, E., Byrne, E., … Sabeti, P. C. (2013). Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell*, *152*(4), 691–702.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241.

Koch, P., Bohlmann, I., Schäfer, M., Hansen-Hagge, T. E., Kiyoi, H., Wilda, M., Hameister, H., Bartram, C. R., & Janssen, J. W. G. (2000). Identification of a novel putative ran-binding protein and its close homologue. *Biochemical and Biophysical Research Communications*, *278*(1), 241–249.

Kohler, M., Leiber, F., Willems, H., Merbold, L., & Liesegang, A. (2013). Influence of altitude on vitamin D and bone metabolism of lactating sheep and goats. *Journal of Animal Science*, *91*(11), 5259–5268.

Kristiansen, O. P., & Mandrup-Poulsen, T. (2005). Interleukin-6 and diabetes. *Diabetes*, *54*(suppl 2), S114.

Lahlou-Kassi, A., Benlamlih, S., & Boukliq, R. (1988). Reproduction and adaptation characteristics in D'man sheep. *Agricultural and Food Science*, *60*(6), 566–575.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Genome Project Data P. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Li, S., Wang, Z., Tong, H., Li, S., & Yan, Y. (2020). TCP11L2 promotes bovine skeletal muscle-derived satellite cell migration and differentiation via FMNL2. *Journal of Cellular Physiology*, *235*, 7183–7193. https://doi.org/10.1002/jcp.29617

Lim, M., Witt, C. C., Graham, C. H., & Dávalos, L. M. (2019). Parallel molecular evolution in pathways, genes, and sites in high-elevation hummingbirds revealed by comparative Transcriptomics. *Genome Biology and Evolution*, *11*(6), 1552–1572. https://doi.org/10.1093/gbe/evz101

Linnen, C. R., Poh, Y.-P., Peterson, B. K., Barrett, R. D. H., Larson, J. G., Jensen, J. D., & Hoekstra, H. E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, *339*(6125), 1312–1316.

Little, G. H., Bai, Y., Williams, T., & Poizat, C. (2007). Nuclear calcium/Calmodulin-dependent protein kinase IIδ preferentially transmits signals to histone deacetylase 4 in cardiac cells. *Journal of Biological Chemistry*, *282*(10), 7219–7231.

Liu, Z., Ou, C., Xiang, S., Xie, Q., Che, S., Zhang, D., Meng, L., Liu, Z., Zhang, G., Wang, Y., Huang, Y., Zhang, W., & Deng, Y. (2022). Expression and regulatory network analysis of BICC1 for aged Sca-1-positive bone narrow mesenchymal stem cells. *Disease Markers*, *2022*, 4759172. https://doi.org/10.1155/2022/4759172

Losos, J. B. (2011). Convergence, adaptation and constraint. *Evolution*, *65*(7), 1827–1840.

Lv, F.-H., Agha, S., Kantanen, J., Colli, L., Stucki, S., Kijas, J. W., Joost, S., Li, M.-H., & Ajmone Marsan, P. (2014). Adaptations to climate-mediated selective pressures in sheep. *Molecular Biology and Evolution*, *31*, 3324–3343. https://doi.org/10.1093/molbev/msu264

Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B., & Hoekstra, H. E. (2010). Convergence in pigmentation at multiple levels: Mutations, genes and function. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *365*(1552), 2439–2450.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614–621.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189–197.

Martin, A., & Orgogozo, V. (2013). The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution; International Journal of Organic Evolution*, *67*(5), 1235–1250. https://doi.org/10.1111/evo.12081

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., & Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics*, *26*(16), 2069–2070.

Mesner, L. D., Ray, B., Hsu, Y. H., Manichaikul, A., Lum, E., Bryda, E. C., Rich, S. S., Rosen, C. J., Criqui, M. H., Allison, M., Budoff, M. J., Clemens, T. L., & Farber, C. R. (2014). Bicc1 is a genetic determinant of osteoblastogenesis and bone mineral density. *The Journal of Clinical Investigation*, *124*(6), 2736–2749. https://doi.org/10.1172/JCI73072

Montejo-Kovacevich, G., Salazar, P. A., Smith, S. H., Gavilanes, K., Bacquet, C. N., Chan, Y. F., Jiggins, C. D., Meier, J. I., & Nadeau, N. J. (2021). Genomics of altitude-associated wing shape in two tropical butterflies. *Molecular Ecology*, *30*, 6387–6402. https://doi.org/10.1111/mec.16067

Naderi, S., Rezaei, H.-R., Pompanon, F., Blum, M. G. B., Negrini, R., Naghash, H.-R., Balkiz, O., Mashkour, M., Gaggiotti, O. E., Ajmone-Marsan, P., Kence, A., Vigne, J.-D., & Taberlet, P. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 17659–17664.

Naval-Sanchez, M., Nguyen, Q., McWilliam, S., Porto-Neto, L. R., Tellam, R., Vuocolo, T., Reverter, A., Perez-Enciso, M., Brauning, R., Clarke, S., McCulloch, A., Zamani, W., Naderi, S., Rezaei, H. R., Pompanon, F., Taberlet, P., Worley, K. C., Gibbs, R. A., Muzny, D. M., … Kijas, J. (2018). Sheep genome functional annotation reveals proximal regulatory elements contributed to the evolution of modern breeds. *Nature Communications*, *9*(1), 859.

New, M., Lister, D., Hulme, M., & Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, *21*(1), 1–25.

Nguyen, T., & Shively, J. E. (2016). Induction of lumen formation in a three-dimensional model of mammary morphogenesis by transcriptional regulator ID4: Role of CaMK2D in the epigenetic regulation of ID4 gene expression. *Journal of Biological Chemistry*, *291*, 16766–16776.

Noonan, J. P., & McCallion, A. S. (2010). Genomics of long-range regulatory elements. *Annual Review of Genomics and Human Genetics*, *11*, 1–24.

Novembre, J., & Di Rienzo, A. (2009). Spatial patterns of variation due to natural selection in humans. *Nature Reviews Genetics*, *10*, 745–755.

Palti, Y., Gao, G., Liu, S., Kent, M. P., Lien, S., Miller, M. R., Rexroad, C. E., III, & Moen, T. (2015). The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Molecular Ecology Resources*, *15*(3), 662–672.

Pan, Y., & Shen, W.-B. (2017). Contemporary crustal movement of southeastern Tibet: Constraints from dense GPS measurements. *Scientific Reports*, *7*, 45348.

Pareek, C. (2010). Emerging trends on functional genomics in farm animals. *Advances in Applied Biotechnology*, *23*, 34.

Pausch, H., Jung, S., Edel, C., Emmerling, R., Krogmeier, D., Götz, K.-U., & Fries, R. (2012). Genome-wide association study uncovers four QTL predisposing to supernumerary teats in cattle. *Animal Genetics*, *43*(6), 689–695.

Pereira, F., Queiros, S., Gusmao, L., Nijman, I. J., Cuppen, E., Lenstra, J. A., Davis, S. J. M., Nejmeddine, F., Amorim, A., & Econogene, C. (2009). Tracing the history of goat pastoralism: New clues from mitochondrial and Y chromosome DNA in North Africa. *Molecular Biology and Evolution*, *26*(12), 2765–2773.

Raeymaekers, J. A. M., Chaturvedi, A., Hablützel, P. I., Verdonck, I., Hellemans, B., Maes, G. E., De Meester, L., & Volckaert, F. A. M. (2017). Adaptive and non-adaptive divergence in a common landscape. *Nature Communications*, *8*(1), 267.

Ratti, E., & Berry, J. D. (2016). Chapter 42 – Amyotrophic lateral sclerosis 1 and many diseases. In T. Lehner, B. L. Miller, & M. W. State (Eds.), *Genomics, circuits, and pathways in clinical neuropsychiatry* (pp. 685–712). Academic Press.

Reid, I. R., Bolland, M. J., & Grey, A. (2014). Effects of vitamin D supplements on bone mineral density: A systematic review and meta-analysis. *The Lancet*, *383*(9912), 146–155.

Riesch, R., Plath, M., & Bierbach, D. (2018). Ecology and evolution along environmental gradients. *Current Zoology*, *64*(2), 193–196.

Sano, H., Peck, G. R., Kettenbach, A. N., Gerber, S. A., & Lienhard, G. E. (2011). Insulin-stimulated GLUT4 protein translocation in adipocytes requires the Rab10 guanine nucleotide exchange factor Dennd4C. *The Journal of Biological Chemistry*, *286*(19), 16541–16545. https://doi.org/10.1074/jbc.C111.228908

Scott, G. R., Elogio, T. S., Lui, M. A., Storz, J. F., & Cheviron, Z. A. (2015). Adaptive modifications of muscle phenotype in high-altitude deer mice are associated with evolved changes in gene regulation. *Molecular Biology and Evolution*, *32*(8), 1962–1976.

Semenza, G. L. (2011). Hypoxia. Cross talk between oxygen sensing and the cell cycle machinery. *American Journal of Physiology. Cell Physiology*, *301*(3), C550–C552.

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. L. (2010). Genetic evidence for high-altitude adaptation in Tibet. *Science*, *329*(5987), 72–75.

Song, S., Yao, N., Yang, M., Liu, X. X., Dong, K. Z., Zhao, Q. J., Pu, Y. B., He, X. H., Guan, W. J., Yang, N., Ma, Y., & Jiang, L. (2016). Exome sequencing reveals genetic differentiation due to high-altitude adaptation in the Tibetan cashmere goat (*Capra hircus*). *BMC Genomics*, *17*, 122.

Steiner, C. C., Rompler, H., Boettger, L. M., Schoneberg, T., & Hoekstra, H. E. (2009). The genetic basis of phenotypic convergence in beach mice: Similar pigment patterns but different genes. *Molecular Biology and Evolution*, *26*(1), 35–45.

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., Purugganan, M. D., & Schmitt, J. (2004). A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(13), 4712–4717.

Stucki, S., Orozco-terWengel, P., Forester, B. R., Duruz, S., Colli, L., Masembe, C., Negrini, R., Landguth, E., Jones, M. R., NEXTGEN

Consortium, Bruford, M. W., Taberlet, P., & Joost, S. (2017). High performance computation of landscape genomic models including local indicators of spatial association. *Molecular Ecology Resources*, *17*(5), 1072–1089.

Sugimoto, M., Fujikawa, A., Womack, J. E., & Sugimoto, Y. (2006). Evidence that bovine forebrain embryonic zinc finger-like gene influences immune response associated with mastitis resistance. *Proceedings of the National Academy of Sciences*, *103*, 6454–6459.

Supek, F., Bosnjak, M., Skunca, N., & Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, *6*(7), e21800.

Switonski, M., Mankowska, M., & Salamon, S. (2013). Family of melanocortin receptor (MCR) genes in mammals—Mutations, polymorphisms and phenotypic effects. *Journal of Applied Genetics*, *54*(4), 461–472.

Taberlet, P., Valentini, A., Rezaei, H. R., Naderi, S., Pompanon, F., Negrini, R., & Ajmone-Marsan, P. (2008). Are cattle, sheep, and goats endangered species? *Molecular Ecology*, *17*(1), 275–284.

Taylor, M. A. (2012). Emerging parasitic diseases of sheep. *Veterinary Parasitology*, *189*(1), 2–7.

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, *29*(12), 673–680.

Uhlik, J., Konradova, V., Vajner, L., & Adaskova, J. (2005). Normobaric hypoxia induces mild damage to epithelium of terminal bronchioles in rabbits (ultrastructural study). *Veterinární Medicína*, *50*(10), 432–438.

Vajana, E., Barbato, M., Colli, L., Milanesi, M., Rochat, E., Fabrizi, E., Mukasa, C., Del Corvo, M., Masembe, C., Muwanika, V., Kabi, F., Sonstegard, T. S., Huson, H. J., Negrini, R., NextGen Consortium, Joost, S., & Ajmone-Marsan, P. (2018). Combining landscape genomics and ecological modelling to investigate local adaptation of indigenous Ugandan cattle to East Coast fever. *Frontiers in Genetics*, *9*, 385.

Vigne, J.-D., Helmer, D., & Peters, J. (2005). *New archaeological approaches to trace the first steps of animal domestication: general presentation, reflections and proposals*. In Vigne, J.D., Peters, J. and Helmer, D. (eds) New methods and the first steps of mammal domestication. Proceedings of the 9th International Council of Archeozoology. (Durham, 23rd-28th August 2002). Oxford: Oxbow Books: 1–16.

Wang, D., Cai, F., Ge, J., & Yin, L. (2015). Brief exercises affect gene expression in circulating monocytes. *Scandinavian Journal of Immunology*, *82*(5), 429–435. https://doi.org/10.1111/sji.12345

Ward, L. D., & Kellis, M. (2012). Interpreting noncoding genetic variation in complex traits and human disease. *Nature Biotechnology*, *30*(11), 1095–1106.

Waters, J. M., & McCulloch, G. A. (2021). Reinventing the wheel? Reassessing the roles of gene flow, sorting and convergence in repeated evolution. *Molecular Ecology*, *30*(17), 4162–4172.

Wei, C. H., Wang, H. H., Liu, G., Zhao, F. P., Kijas, J. W., Ma, Y. J., Lu, J., Zhang, L., Cao, J. X., Wu, M. M., Wang, G., Liu, R., Liu, Z., Zhang, S., Liu, C., & du, L. (2016). Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific Reports*, *6*, 26770.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370.

Yang, J., Li, W.-R., Lv, F.-H., He, S.-G., Tian, S.-L., Peng, W.-F., Sun, Y.-W., Zhao, Y.-X., Tu, X.-L., Zhang, M., Xie, X. L., Wang, Y. T., Li, J. Q., Liu, Y. G., Shen, Z. Q., Wang, F., Liu, G. J., Lu, H. F., Kantanen, J., … Liu, M. J. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution*, *33*(10), 2576–2592.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z. X. P., 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 50 human exomes reveals adaptation to high altitude. *Science*, *329*(5987), 75–78.

Zeder, M. A., Emshwiller, E., Smith, B. D., & Bradley, D. G. (2006). Documenting domestication: The intersection of genetics and archaeology. *Trends in Genetics*, *22*(3), 139–155.

Zhan, X., Pan, S., Wang, J., Dixon, A., He, J., Muller, M. G., Ni, P., Hu, L., Liu, Y., Hou, H., Chen, Y., Xia, J., Luo, Q., Xu, P., Chen, Y., Liao, S., Cao, C., Gao, S., Wang, Z., … Bruford, M. W. (2013). Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nature Genetics*, *45*, 563–566.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Benjelloun, B., Leempoel, K., Boyer, F., Stucki, S., Streeter, I., Orozco-terWengel, P., Alberto, F. J., Servin, B., Biscarini, F., Alberti, A., Engelen, S., Stella, A., Colli, L., Coissac, E., Bruford, M. W., Ajmone-Marsan, P., Negrini, R., Clarke, L., Flicek, P. … Pompanon, F. (2023). Multiple genomic solutions for local adaptation in two closely related species (sheep and goats) facing the same climatic constraints. *Molecular Ecology*, *00*, e17257. https://doi.org/10.1111/mec.17257