A haplotype-based evolutionary history of barley domestication

Dissertation
zur Erlangung des
Doktorgrades der Agrarwissenschaften (Dr. agr.)

der

Naturwissenschaftlichen Fakultät III
Agrar- und Ernährungswissenschaften,
Geowissenschaften und Informatik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von
Herrn M.Sc. Yu Guo

Gutachter:
Prof. Dr. Thorsten Schnurbusch
Prof. Dr. Klaus F.X. Mayer

Verteidigung am: 12.05.2025

Acknowledgments

Throughout the journey of completing this dissertation, I have received invaluable support and encouragement from many individuals. Their guidance, assistance, and steadfast belief in me have been instrumental in bringing this work to fruition. I would like to express my deepest gratitude to everyone who has supported me during this time.

First and foremost, I am profoundly grateful to my advisor, Dr. Martin Mascher, for his patient and meticulous guidance throughout my research. His insights and academic expertise have been crucial in shaping the direction of my study. I am also sincerely thankful to my co-advisors, Dr. Frank Blattner and Dr. Thorsten Schnurbusch, for their advice and constructive feedback, which have helped me refine my work.

I extend my heartfelt thanks to Dr. Nils Stein, head of the Department of Genebank at IPK, for his unwavering support. I am especially grateful to the generations of researchers who contributed to the genebank's diverse collection. The global diversity of barley samples they preserved formed the foundation of my research on the domestication history of barley. A special acknowledgment goes to Professor Brian Steffenson for providing the wild barley samples that were critical to my study.

I also want to express my gratitude to the technical teams that supported me during my research. My thanks go to Axel Himmelbach and Mary Ziems from the DNA sequencing department for their expertise, as well as Anne Fiebig from the data upload department for her assistance. Additionally, I deeply appreciate the administrative support provided by Bianka Jacobi and Britt Leps, whose efficiency and dedication facilitated the smooth progress of my work.

My gratitude also goes to my fellow lab members in the DG group, whose camaraderie and collaboration have been a constant source of inspiration and encouragement. I am thankful to Raz Avni, Amanda Souza Camara, Woohyeon Cho, Jia-Wu Feng, Max Haupt, Murukarthick Jayakodi, Marina Püpke Marone, Mona Schreiber, Hyeonah Shim, and Hailin Zhang. Your companionship, academic discussions, and shared experiences have enriched my journey and made it truly memorable.

Finally, I would like to thank my family—my wife, my parents, and my five-year-old daughter—for their unshakeable support throughout this journey. My wife's dedication and effort in managing family responsibilities allowed me to concentrate on my work. I am also grateful for the joy and positivity my daughter has brought into my life during this time.

With this acknowledgment, I extend my deepest gratitude to everyone who has supported and helped me along the way.

Abbreviations

%	Percent
℃	degree Celsius
θ_W	Watterson's estimator
π	nucleotide diversity
AFD	absolute allele frequency difference
AHG	ancestral haplotype groups
BAM	Binary Alignment Map
bp	base pair
CA	Central Asia
DNA	deoxyribonucleic acid
F_{ST}	Pairwise fixation indices
FRAC	fraction
GBS	genotyping by sequencing
iTOL	Interactive Tree of Life
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
ka BP	thousand years before present
K	number of ancestral populations
kb	kilobase
LD	linkage disequilibrium
LGM	Last Glacial Maximum
MAF	minor allele frequency
Mb	megabase
NE	Near East
NL	Northern Levant
NJ	neighbor-joining
NM	Northern Mesopotamia
PCA	principal component analysis

PSMC	pairwise sequentially Markovian coalescent
Q20	mapping quality≥20
SD	Syrian Desert
SL	Southern Levant
SNP	single nucleotide polymorphism
UDG	uracil-DNA glycosylase
VCF	variant call format file
WBDC	Wild Barley Diversity Collection
WGS	whole-genome sequencing

Contents

1. Introduction

1.1 Wild barley

Barley, belonging to the genus *Hordeum* within the grass family (*Poaceae*), is one of the oldest domesticated crops in the world. Archaeological evidence indicates that cultivated barley (*Hordeum vulgare* subsp. *vulgare*) was domesticated from its progenitor-wild barley (*Hordeum vulgare* subsp. *spontaneum*) in the Near East approximately 10,000 years ago (Zohary et al. 2012). Wild barley diverged from its closest relative, *Hordeum bulbosum*, approximately 3.7 million years ago in the Near East (Blattner 2018). Wild barley exhibits far greater genetic diversity than cultivated barley, containing a more extensive gene pool (Russell et al. 2016; Liu et al. 2020). The gene pool of wild barley is a valuable resource for cultivated barley, as it contains genes for traits such as disease resistance, drought tolerance, and salt tolerance that cultivated varieties have lost or never possessed. By studying wild barley, it is possible to identify genes beneficial for modern barley breeding. These genes can be introduced into cultivated barley through hybridization or gene editing, helping to improve crop yield and adaptability. Consequently, Ellis et al. proposed that "wild barley is a source of genes for crop improvement in the 21st century" (Ellis et al. 2000). Successful applications include enhancing tolerance to abiotic stresses, such as water stress (Ivandic et al. 2000; Suprunova et al. 2007) and drought (Nevo and Chen 2010; Lakew et al. 2011), as well as improving disease resistance (Morrell et al. 2005; Steffenson et al. 2007; Ames et al. 2015). There have also been efforts to introduce beneficial agronomic traits from wild barley into cultivated varieties through introgression lines (von Korff et al. 2006; Schmalenbach et al. 2009; Naz et al. 2014), such as those for malting quality (Schmalenbach and Pillen 2009).

The extant wild barley is widely distributed across the Near East, North Africa, Central Asia, and extends as far east as Tibet in China (Harlan and Zohary 1966).

Some studies suggest that the global population structure of wild barley is largely shaped by geographical factors. As wild barley spread into new habitats, adaptation to local environments through natural selection contributed to its present distribution patterns (Hübner et al. 2009; Fang et al. 2014; Jakob et al. 2014; Russell J et al. 2016). However, fundamental questions regarding the evolutionary history of wild barley remain unresolved. These include the timing and pathways of its dispersal from the Near East to other regions, as well as changes in population sizes throughout its spread. Understanding the evolutionary history of wild barley is essential for several reasons. First, it sheds light on how plant populations adapt to diverse environments over time, offering valuable insights into evolutionary biology. Second, this knowledge may uncover genetic traits that enhance resilience to environmental stresses, which could be applied to cultivated barley to improve its sustainability and performance under changing climate conditions. Lastly, studying the dispersal and adaptation patterns of wild barley can inform conservation strategies for preserving genetic diversity – a priority as agricultural biodiversity faces increasing threats from habitat loss and climate change. In my study, I aim to contribute to this field by expanding upon current research into the evolutionary history of wild barley.

1.2 Barley domestication

Barley (*Hordeum vulgare*) is one of the oldest cultivated crops, with mentions of it appearing in some of the earliest written records, dating back to around 3100 BCE (Hasselbach-Andee 2020). By that time, the practice of plant cultivation had already been established long before the invention of writing. Much of what is known about the early domestication and spread of barley and other crops is derived from archaeological findings, with the earliest evidence dating back approximately 10,000 years (BP) (Zohary et al. 2012; Fuller et al. 2023). These findings primarily consist of charred grains, which archaeobotanists use to identify key traits of domestication, such as the loss of spike brittleness (Zohary

et al. 2012). Molecular geneticists have since isolated the genes responsible for these and other domestication traits, studying the sequence diversity at surrounding loci in both wild and domesticated forms to pinpoint the likely wild ancestors of cultivated barley (Pourkheirandish et al. 2015). Genome sequences have become a valuable resource for research on crop evolution (Schreiber et al. 2018). In the last two decades, as whole-genome sequencing has become more accessible, the field has expanded significantly. Many crop evolutionists, often implicitly, operate under the assumption that current population structures, particularly in wild relatives, can reflect ancient scenarios. However, this assumption is not necessary for methods like the pairwise sequentially Markovian coalescent (PMSC), which can infer historical population size trajectories from the genome sequences of current individuals (Li and Durbin 2011). A recent advancement in the field is the development of IntroBlocker (Wang et al. 2022), software designed to define ancestral haplotype groups (AHGs). These co-inherited blocks of single nucleotide polymorphisms (SNPs), or haplotypes, enable the identification of "closest wild relatives" not just at the population or whole-plant level, but also within specific megabase-sized genomic regions. Additionally, ancient DNA sequences provide valuable insights into past genetic diversity (Orlando et al. 2021). DNA extracted from well-preserved ancient specimens, primarily bones, has significantly altered our understanding of human prehistory (Posth et al. 2023). Although there have been some notable successes with ancient plant DNA (Mascher et al. 2016; Kistler et al. 2020), its impact on crop evolutionary studies has been limited due to the poor preservation of plant materials in most environments (Gutaker and Burbano 2017).

Determining where and when plants were first domesticated is a fundamental question in the study of crop evolution. Historically, the search for a single point in time and space where domestication occurred drove geneticists to map "centers of origin" using molecular markers. Efforts to do this in Einkorn wheat (Heun et al. 1997) were met with both conceptual and methodological challenges (Allaby

and Brown 2003). In the case of barley, however, the concept of a single monophyletic origin is not applicable. Multiple lines of evidence have shown that barley did not originate from a single domestication event. For example, two independent loss-of-function mutations that eliminate spike brittleness have been identified, with the closest wild relatives of these mutant alleles located in geographically and genetically distinct wild populations (Pourkheirandish et al. 2015). Further supporting evidence comes from genome-wide marker data, leading Poets et al. (2015) to propose a model of mosaic ancestry, where domesticated barley does not descend from a single wild population but has incorporated contributions from several different wild groups. Poets et al. (2015) and others (Russell et al. 2016; Pankin et al. 2018) conducted their studies before the availability of nearly complete and highly contiguous barley genome sequences (Mascher et al. 2021) and relied on markers identified in domesticated barley or reduced representation sequencing. In this study, we use whole-genome sequences from diverse wild and domesticated barley populations, including ancient specimens, to link haplotypes in both wild and domesticated populations. This approach allows us to address key questions such as: Which regions of the genome trace back to specific wild ancestors? Which wild barley genomes are the closest present-day relatives of domesticated haplotypes? How have haplotypes evolved and reshuffled after domestication?

1.3 Six-rowed wild growing barley

All domesticated crops originated from wild plant ancestors. Many of these crops are capable of interbreeding with their extant wild relatives (Harlan and de Wet 1971), and both genetic and ecological studies provide substantial evidence of natural hybridization between wild and domesticated populations in overlapping distribution area (Ellstrand et al. 2013; Janzen et al. 2019). Two primary traits distinguish domesticated barley from its wild relatives: non-shattering (or non-brittle) spikes and the presence of six-rowed spikes. In wild barley, the central

axis of the spike, or rachis, naturally disintegrates at maturity, breaking into individual dispersal units or spikelet triplets. These triplets consist of one central grain-bearing spikelet flanked by two lateral spikelets that remain sterile. In domesticated six-rowed barleys, however, the lateral spikelets are fertile, resulting in increased yield. Today, six-rowed barley varieties dominate cultivation in most regions worldwide (Milner et al. 2019).

Although *H. spontaneum* is a two-rowed species, reports have surfaced of "wild" six-rowed barleys that thrive without human cultivation. Whether these represent true wild descendants of *H. spontaneum* or are crop-wild hybrids has been debated. European explorers who first documented these six-rowed wild barleys in East Asia classified them as *H. agriocrithon* (Åberg 1938) and speculated that they were ancestral to all domesticated barleys. This hypothesis was based on the assumption that acquiring fertile lateral spikelets in two-rowed barley would have been highly improbable, leading Schiemann (1932) to propose that two-rowed barley forms likely evolved from six-rowed ancestors by losing lateral spikelet fertility.

Subsequent research, however, challenged this hypothesis. Mutation breeding experiments in the 1960s demonstrated that two-rowed barleys could, in fact, mutate into six-rowed forms, showing that the transition is more feasible than previously thought. Numerous six-rowed mutants were identified following irradiation of two-rowed barley (Lundqvist et al. 1989; Komatsuda et al. 2007), suggesting that mutations converting two-rowed to six-rowed types are readily achievable. The genetic mechanisms underlying non-brittle and six-rowed traits were further clarified through positional cloning. Komatsuda et al. (2007) revealed that all six-rowed barleys carry loss-of-function mutations in the *VRS1* homeobox domain transcription factor. Similarly, non-brittle barleys emerged through independent mutations in two closely linked genes, *NON-BRITTLE RACHIS 1* (*BTR1*) and *NON-BRITTLE RACHIS 2* (*BTR2*)

(Pourkheirandish et al. 2015). With this genetic foundation, we can envision three potential pathways for the emergence of six-rowed, wild-type barley. The first pathway is spontaneous mutation, where six-rowed types arise independently in *H. spontaneum* populations through mutations in *VRS1* or other genes without interaction with domesticated barley. The second is introgression, whereby a domesticated *vrs1* knock-out allele is introduced into a wild genomic background. The third pathway is de-domestication by recombination; in crosses between domesticated barley plants carrying distinct allelic combinations (*btr1Btr1* and *Btr1btr2*), rare crossover events can produce recombinant offspring with functional *Btr1Btr2* alleles, potentially resulting in a "de-domesticated" barley. In six-rowed backgrounds, this recombination would yield a barley with brittle spikes and fertile lateral florets.

Pourkheirandish et al. (2018) found evidence that all three of these scenarios may have contributed to the origin of H. agriocrithon. By resequencing *BTR1/2* and *VRS1*, they showed that six-rowed wild barleys from Tibet arose either through recombination or introgression, while "eu-agriocrithon" barleys from Central Asia appear to be direct descendants of an ancient six-rowed wild barley population from which modern six-rowed varieties were domesticated. This latter possibility, which implies that six-rowed barley predates the non-shattering trait, has drawn our interest and led us to explore this hypothesis further through population genomic analysis.

2. Materials and Methods

2.1 Wild barley

Our wild barley panel consists of 285 accessions from the Wild Barley Diversity Collection (WBDC), which represents a broad range of ecogeographic diversity. The whole-genome sequencing of the WBDC collection is comprehensively described in a paper by Sallam et al. (2024). Additionally, we incorporated 95 diverse barley accessions, primarily drawn from the panel described by Russell et al. (2016). These samples had previously been sequenced at approximately 3× coverage by Jayakodi et al. (2020). For the current study, we generated deeper sequencing data (~10× coverage) to enhance genomic resolution.

2.2 Domesticated barley

Milner et al. (2019) employed model-based ancestry estimation with ADMIXTURE (Alexander et al. 2009) to define 12 populations within a global diversity panel of 19,778 domesticated barley accessions, which had been genotyped using genotyping-by-sequencing (GBS) (Milner et al. 2019). We utilized their ADMIXTURE results and the GBS single-nucleotide polymorphism (SNP) matrix for sample selection. With the exception of the Near-Eastern (NE) population (depicted in orange of Milner et al. 2019), the following approach was used to select samples: First, unadmixed samples (those with an ADMIXTURE ancestry coefficient $q \geq 0.95$) were input into a principal component analysis (PCA) using smartpca (version 7.2.1) (Patterson et al. 2006). Samples were then chosen to ensure even coverage of the PCA diversity space (**Fig. 1**). Due to the NE population's higher genetic diversity and substructure, a more detailed approach was applied (**Fig. 2**). Specifically, ADMIXTURE (version 1.23) (Alexander et al. 2009) was rerun on 1,078 NE population samples identified by Milner et al. (2019), where the NE ancestry coefficient was the highest among all

populations (q ranging from 0.25 to 0.98). Before running ADMIXTURE, SNPs were thinned using PLINK (version 1.9) (Purcell et al. 2007) with the parameters `--indep-pairwise 50 10 0.1`. For *K* values (number of ancestral populations) ranging from 2 to 6, 15 replicate ADMIXTURE runs with varying random seeds were merged using CLUMPP (version 1.1.2) (Jakobsson and Rosenberg 2007) and visualized with Distruct (version 1.1) (Rosenberg 2004). Individuals with $q \geq 0.80$ for their primary ancestry component were classified as unadmixed. The results for *K*=6 were selected for further analysis, and the genetic separation of these subpopulations was validated using PCA. From the NE subpopulations, only samples originating from the Near East were included for sequencing. These selected samples were distributed evenly across the PCA space to capture genetic diversity. Overall, 302 samples representing 15 populations were selected. Populations were named based on their geographic origins and three key traits related to global population structure (**Table 1**): row type (two-rowed [T], six-rowed [S], mixed [M]), lemma adherence (hulled [H], naked [N]), and growth habit (winter-sown [W], spring-sown [S], mixed [M]). For instance, the ISR-THS population predominantly consists of two-rowed, hulled, spring barley accessions from Israel. For each population, approximately 20 accessions were selected for whole-genome sequencing (WGS). Of these, 7~10 accessions per population (116 samples in total) were sequenced at high coverage (~10×), while the remaining 186 samples were sequenced at lower coverage (~3×). Seeds for all selected accessions are available from the German Federal ex situ genebank at IPK Gatersleben.

Table 1: Explanation of the abbreviations of 15 domesticated barley populations

population	meaning	sample number (high coverage)	sample number (low coverage)	total
ISR-THS	Israel two-rowed hulled spring	6	11	17
SYR-THM	Syria two-rowed hulled spring & winter	10	14	24
GEO-THS	Georgia two-rowed hulled spring	8	13	21
IRN-THS	Iran two-rowed hulled spring	6	14	20
TUR-THM	Turkey two-rowed hulled spring & winter	6	11	17
EU-THM	Europe two-rowed hulled spring & winter	5	11	16
EU-THS	Europe two-rowed hulled spring	10	10	20
ME-SHS	Mediterranean six-rowed hulled spring	10	10	20
EU-SHW	Europe six-rowed hulled winter	7	13	20
EU-SHM	Europe six-rowed hulled spring & winter	9	12	21
ETH-MHS	Ethiopia six-rowed & two-rowed hulled spring	8	12	20
ETH-MNS	Ethiopia six-rowed & two-rowed naked spring	9	12	21
CA-SHS	Central Asia six-rowed hulled spring	7	14	21
CA-SNS	Central Asia six-rowed naked spring	7	16	23
EA-SHW	Eastern Asia six-rowed hulled winter	8	13	21
total		116	186	302

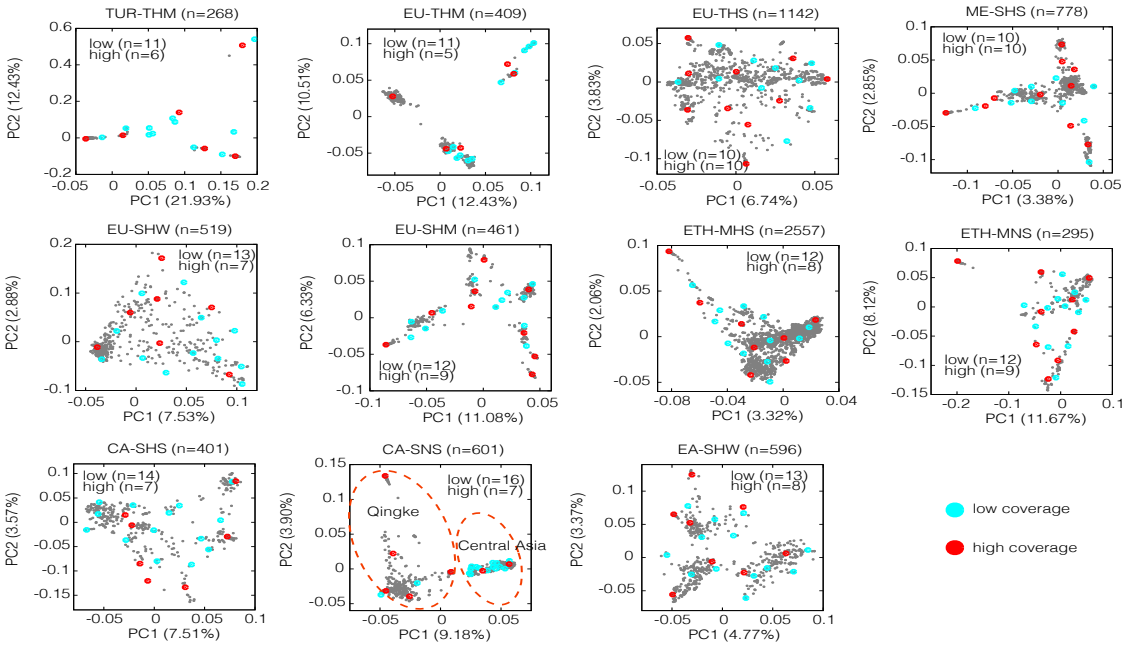


Figure 1: Selection of samples for whole-genome sequencing from 11 domesticated barley populations. Milner et al. (2019) defined 12 populations of domesticated barley. The selection of samples from their "orange" population is shown in **Fig. 2**. For each of the remaining 11 populations, PCAs were run for un-admixed samples, i.e. those with ADMIXTURE ancestry

coefficients ≥ 0.95. Then, about 20 samples (13 high-coverage, 7 low-coverage) were selected to cover the PCA diversity space of each population. The left-hand cluster in the PCA of CA-SNS (Central Asia 6-rowed naked barleys) contains Qingke (Tibetan hulless barleys). This population was studied by Zeng et al. (2018) in detail and was not the focus of the present study. Hence, comparatively fewer Qingke samples were selected.

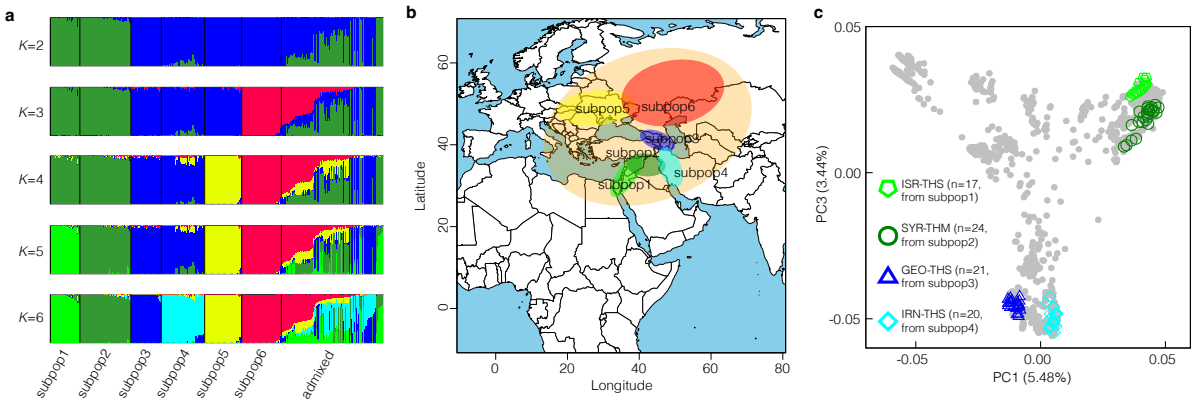


Figure 2: Selection of samples for whole-genome sequencing from the "orange" population of Milner et al. (2019)**.** This figure is a complement to the sample selection in **Fig. 1**. **(a)** Individual ADMIXTURE ancestry coefficients with the number of ancestral populations (*K*) ranging from 2 to 6. ADMIXTURE was run on member of the Milner et al.'s "orange" population using their GBS data. Individuals whose major ancestry coefficient was less than 0.8 were considered admixed. **(b)** Predominant geographical origins of the subpopulations of "orange". **(c)** Samples from four populations in the Near East and Causasus that were selected for whole-genome sequencing are highlighted in the PCA.

2.3 Plant growth, DNA isolation and Illumina sequencing

Plant cultivation and DNA extraction followed previously established protocols (Milner et al. 2019). Whole-genome sequencing (WGS) libraries were prepared using the Illumina Nextera DNA Flex kit and sequenced in paired-end mode (2 × 151 cycles) on an Illumina NovaSeq 6000 system at IPK Gatersleben, adhering to the manufacturer's guidelines (Illumina Inc., San Diego, CA, USA).

2.4 Read mapping and variant calling

The sequencing reads from 682 barley genotypes, comprising 380 wild and 302 domesticated accessions, were aligned to the MorexV3 genome assembly (Mascher et al. 2021) using Minimap2 (version 2.24) (Li 2021). The resulting BAM files were processed with novosort (version 3.06.05) (https://www.novocraft.com/products/novosort/) to sort and remove duplicates. Variant calling was performed with BCFtools (v1.15.1) (Danecek et al. 2021) using the command `mpileup -a DP,AD -q 20 -Q 20 --ns 3332`.

The initial raw SNP matrix underwent several filtering steps:

1. Retaining only bi-allelic SNP sites.
2. Marking genotype calls as valid if the read depth ranged between ≥2 and ≤50; otherwise, genotypes were set to missing.
3. Including SNP sites with less than 20% missing calls and fewer than 20% heterozygous calls for ADMIXTURE analysis with *K* values from 2 to 4.

At *K*=4, wild barley samples with ≥15% ancestry attributed to domesticated barley were classified as admixed. Consequently, 80 admixed wild samples were excluded from further analysis. Of these, 251 wild barley accessions with high coverage (~10x) and no domesticated admixture were selected for population genetic analyses.

Two SNP matrices, SNP1 and SNP2, were generated for downstream analysis:

SNP1: Derived from 367 high-coverage samples (251 wild and 116 domesticated) extracted from the raw SNP matrix. Filtering criteria included:

1. Retaining only bi-allelic SNP sites.
2. Considering homozygous calls valid if their read depth was ≥2 and ≤50; otherwise, set to missing.

3. Treating heterozygous calls as valid if the allelic depth of both alleles was ≥5; otherwise, set to missing.

SNP2: Constructed from 302 domesticated samples by running bcftools on a downsampled dataset. In this dataset, the read alignments of high-coverage samples (n=116) were thinned to match the sequence depth of low-coverage samples. Downsampling was performed using SAMtools (v1.16.1) (Danecek et al. 2021) with the command `samtools view -s 0.FRAC` (where FRAC is the sampling rate). The target number of uniquely mapped (Q20), deduplicated reads for downsampled high-coverage data was randomly set between 35M and 52M. All reads were 2 × 150 bp in length. SNP2 was filtered as follows:

1. Retaining only bi-allelic SNP sites.
2. Considering homozygous calls valid if their read depth was ≥2 and ≤20; otherwise, set to missing.
3. Setting all heterozygous calls to missing.

A flowchart detailing the construction of these SNP matrices is provided in **Fig. 3**. For analyses requiring an outgroup, WGS data (Mascher et al. 2013) of *Hordeum pubiflorum* was utilized. Read mapping and SNP calling followed the same protocol, with one exception: a VCF file was generated to include all genomic sites, even those identical to the reference genome. This VCF file was merged with other VCF files to infer ancestral states.

2.5 SNP-based genetic distances

The number of SNPs between any two high-coverage genotypes was calculated using the following approach. First, pairwise SNP counts were determined within specific genomic windows using PLINK2 (version 2.00a3.3LM) (Chang et al. 2015). The command employed was:

```
plink2 --from-bp x --to-bp y --sample-diff counts-only
counts-cols=ibs0,ibs1 ids=s1 s2 …
```

Here, `x` and `y` represent the start and end coordinates of a window, while `s1, s2, …` denote the list of sample IDs. Multiple window sizes were analyzed: 100 kb (with a shift of 20 kb), 500 kb (shift: 100 kb), 1 Mb (shift: 200 kb), 2 Mb (shift: 400 kb), and 5 Mb (shift: 5 Mb).

To account for variations in the effective coverage of short reads across different genomic regions, a normalized distance metric was calculated within each window (**Fig. 4**). For this purpose, the per-base read depth of each sample was determined across the reference genome using the command:

```
samtools view -q 20 -F 3332 | samtools depth
```

The effectively covered region of each window was defined as the union of sites with read depths between 2 and 50. For each pairwise sample comparison, the intersection of the effectively covered regions was calculated using a Perl script.

The pairwise distance in a genomic window was then computed as:

$$\text{distance} = \frac{\text{hom} + \frac{\text{het}}{2}}{\text{cov}}$$

where `hom` and `het` represent the counts of homozygous and heterozygous differences, respectively, and `cov` is the size of the intersection of the effectively covered regions for the two samples. Genomic windows were included in the analysis only if the effective coverage exceeded half the window size; otherwise, the distance for that window was set to missing.

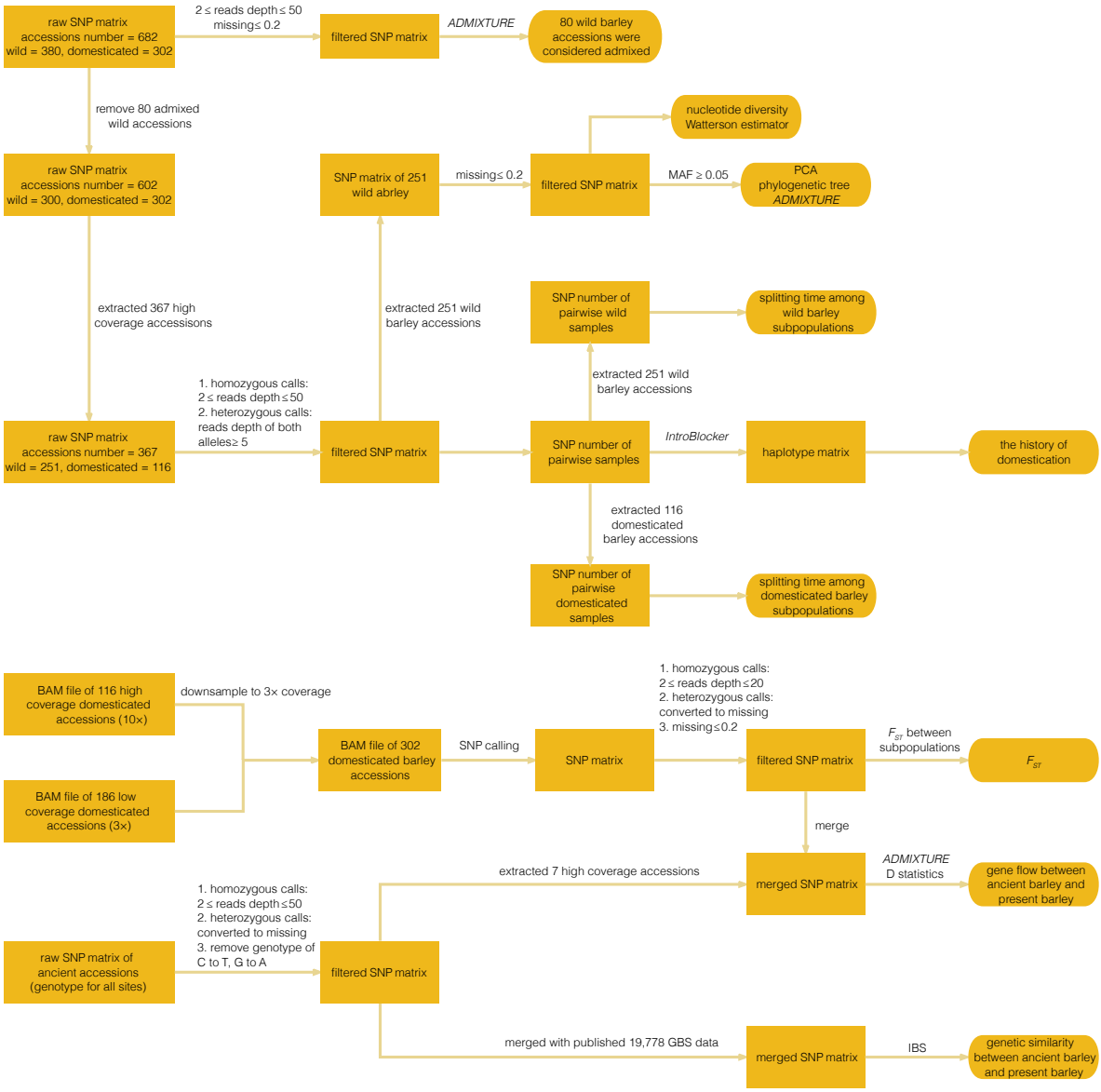


Figure 3: Workflow of analyses conducted on SNP matrices.

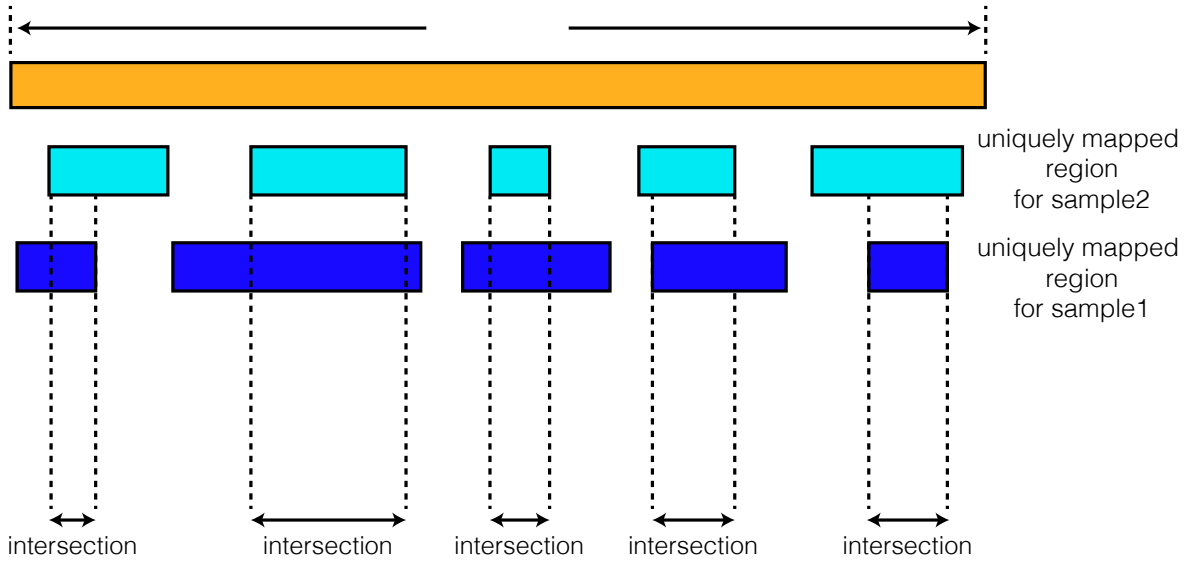


Figure 4: Normalization of SNP counts for read depth. To determine the number SNPs between two samples in a given genomic window of size, we first intersect the regions covered by at least two uniquely mapped (MAPQ ≥ 20) reads. Then, we calculated the normalized SNP number according to the formula: (raw SNP number / cumulative of the intersected intervals) × window size. The "raw" SNP number was determined with command "sample-diff counts-only" of Plink2.

2.6 LD decay

The barley genome was divided into three distinct compartments—distal, interstitial, and proximal—based on recombination rates as described in a previous study (Mascher et al. 2017) (**Table 2**). Linkage disequilibrium (LD) decay for both wild and domesticated barley was analyzed within each compartment using PopLDdecay (version 3.42) (Zhang et al. 2019). The command used for this analysis was:

```
-Het 0.99 -Miss 0.2 -MAF 0.01 -MaxDist 500
```

Table 2: Boundaries of the genomic compartments in MorexV3 (unit: Mb).

chromosome	distal	interstitial	proximal	interstitial	distal	centromere
chr1H	0-50	50-100	100-290	290-470	470-516	206.49
chr2H	0-80	80-200	200-400	400-590	590-665	301.29
chr3H	0-50	50-210	210-380	380-530	530-621	267.85
chr4H	0-60	60-100	100-400	400-540	540-610	276.15
chr5H	0-50	50-130	130-300	300-540	540-588	204.88
chr6H	0-50	50-170	170-330	330-510	510-561	256.32
chr7H	0-70	70-230	230-400	400-570	570-632	328.85

2.7 Evolutionary analyses in wild barley

Variant calls for 251 high-coverage wild barley samples were extracted from the SNP1 matrix described earlier. SNP sites were included in the population structure analysis if they had fewer than 20% missing calls, fewer than 20% heterozygous calls, and a minor allele frequency (MAF) ≥ 5%. Population structure was analyzed using ADMIXTURE (version 1.23) (Alexander et al. 2009) with the number of ancestral populations (*K*) ranging from 2 to 5. Principal component analysis (PCA) was conducted using smartpca (version 7.2.1) (Patterson et al. 2006).

The genotype data from the outgroup sample *Hordeum pubiflorum* was merged with the SNP matrix, and an identity-by-state (IBS) genetic distance matrix was calculated using PLINK (version 1.9) (Purcell et al. 2007). This distance matrix was used to construct a neighbor-joining (NJ) tree with Fneighbor (Rice et al. 2000) (part of the EMBOSS package), available at http://emboss.toulouse.inra.fr/cgi-bin/emboss/fneighbor. The resulting tree was visualized using Interactive Tree of Life (iTOL) (Letunic and Bork 2021).

For each of the five wild barley subpopulations, nucleotide diversity (π) (Tajima 1983) and Watterson's estimator (θ_W) (Watterson 1975) were calculated using the

SNP matrix without MAF filtering, employing a previously published Perl script (Zeng et al. 2018). Pairwise fixation indices (F_{ST}) were computed for genomic windows (1 Mb size, 500 kb overlap) between wild barley populations, using Hudson's estimator based on equation 10 in a prior reference (Bhatia et al. 2013).

Coverage-normalized SNP distances were determined in 1-Mb genomic windows (shift: 200 kb) as described earlier. The $\log_{10}$-transformed distributions of SNP distances across the distal, interstitial, and proximal genomic compartments were visualized for each wild barley population using R (version 3.5.1) (R Core Team 2013).

To infer divergence times, only SNPs from a 50-Mb region surrounding the centromeres (±25 Mb) were analyzed. Divergence times were calculated using the formula $g=d/2\mu$, where g represents the number of generations, d is the SNP distance per bp, and μ is the mutation rate. Assuming a generation time of one year for *Hordeum vulgare*, we used a mutation rate of 6.13×10^{-9} as estimated for *Brachypodium distachyon* by Wang et al. 2019. SNP frequency distributions were visualized using logarithmic binning (50 bins, range: 10^1 to $10^{4.5}$, corresponding to 31,622 SNPs).

Demographic history was inferred using the PSMC model (version 0.6.5-r67) (Li and Durbin 2011) with default parameters, based on pseudo-diploid genomes created by combining BAM files from two homozygous individuals, as previously described (Meyer et al. 2016; Cubry et al. 2018; Shah et al. 2020). A total of 38 pseudo-diploid genomes were generated, including one haplotype from the SL population and the other from one of the four remaining wild barley populations.

Comparative genomic analysis was performed using MUMmer (version 4.0.0) (Marçais et al. 2018) to align eight barley genome assemblies (Jayakodi et al. 2024), each representing different haplotypes. Alignments focused on

chromosome 5H (100–300 Mb region), with minimum identity and alignment length thresholds set to 90% and 2,000 bp, respectively.

2.8 Definition of ancestral haplotype groups

Ancestral haplotype groups (AHGs) were identified using IntroBlocker (Wang et al. 2022). To establish an appropriate threshold for distinguishing haplotypes, coverage-normalized SNP-based distances were calculated in 1-Mb windows (shift: 200 kb) for three pairwise sample sets: (1) within wild barley samples, (2) within domesticated barley samples, and (3) between wild and domesticated barley samples. All possible sample combinations within each set were analyzed. A threshold of 400 SNPs per Mb was selected for separating AHGs.

Coverage-normalized SNP-distance matrices, derived from 367 high-coverage samples, were used as input for IntroBlocker under the "semi-supervised" model, prioritizing wild samples over domesticated ones in AHG labeling. IntroBlocker was executed with multiple window sizes: 100 kb (shift: 20 kb), 500 kb (shift: 100 kb), 1 Mb (shift: 200 kb), 2 Mb (shift: 400 kb), and 5 Mb (shift: 5 Mb). Based on a detailed examination of the results (**Fig. 5**), the outputs from the 100-kb window analysis were selected for subsequent analyses.

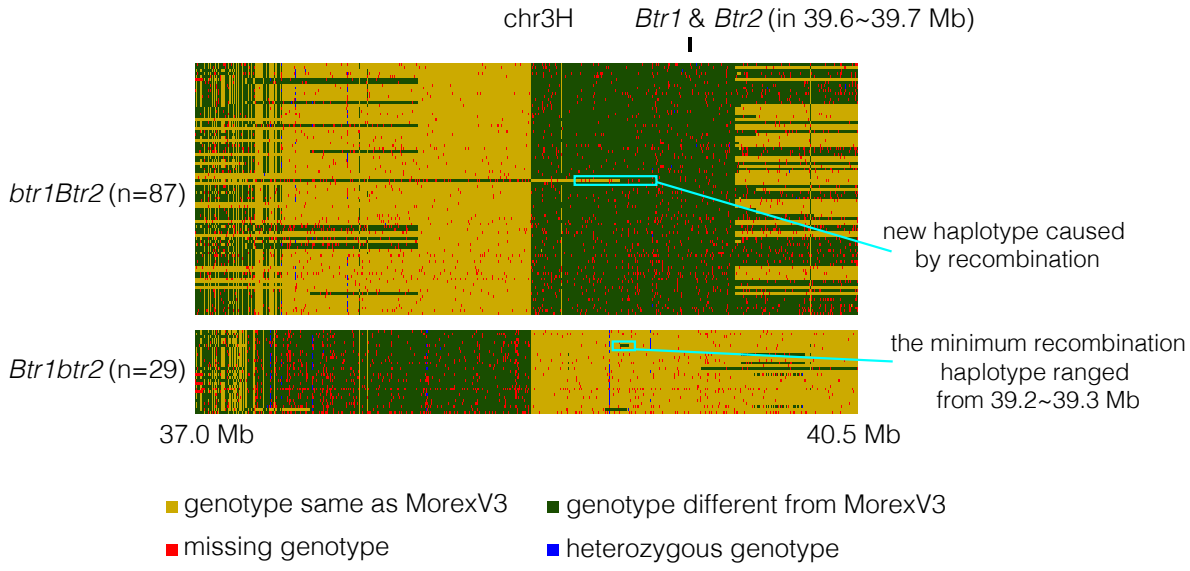


Figure 5: Effect of window size on haplotype definition. The haplotype structure at the *btr1/2* locus is shown as an example. The samples shown are cultivated barleys with recombinant haplotypes. If the window size is large, more recombinant haplotypes are considered novel. The length of the smallest sequence exchange we observed between the original haplotypes (*Btr1btr2* and *btr1Btr2*) was 100 kb. Hence, we used this window size (with a 20 kb shift) to compile the haplotype matrix for subsequent analyses.

2.9 Analysis of the AHG matrix

The proportions of shared and private AHGs in wild and domesticated barley populations were calculated using custom Perl scripts**.** Saturation curves were constructed by randomly selecting subsets of k wild barley samples (from a total of 251 samples), with k ranging from 1 to 250. For each k, 100 random subsets were generated. The proportion of haplotypes in domesticated barley shared with these subsets was determined for each sample. Mean values and 95% confidence intervals were calculated for each k using R (version 3.5.1) (R Core Team 2013) and a t-test. Two-dimensional haplotype frequency spectra were also computed using custom Perl scripts**,** excluding genomic windows with more than 20% missing data.

To estimate when wild haplotypes entered the domesticated gene pool, IntroBlocker was run with varying thresholds for haplotype separation: 400 SNPs (~32,000 years ago), 98 SNPs (~8,000 years), 73 SNPs (~6,000 years), 49 SNPs (~4,000 years), and 24 SNPs (~2,000 years). For each genomic window where haplotypes were shared between wild and domesticated samples, the temporal and spatial origins of this sharing were inferred, as shown in **Fig. 6**. For each domesticated haplotype, results from the different IntroBlocker thresholds (representing divergence time brackets) were compared. The most recent time bracket where haplotype sharing occurred was considered the terminus post quem (earliest possible time) for the incorporation of a wild haplotype into the domesticated gene pool. This approach does not assume the direction of gene flow.

To exclude recent introgressions from domesticated barley into wild populations, windows where multiple domesticated samples shared haplotypes with a small number of wild samples that diverged within the last ~8,000 years were removed. The spatial origin of domesticated haplotypes was determined by averaging the ADMIXTURE ancestry coefficients of all wild samples carrying that haplotype **Fig. 6**. If two wild samples sharing a domesticated haplotype were highly similar (pairwise IBS ≥ 0.95), only one was included in the calculation.

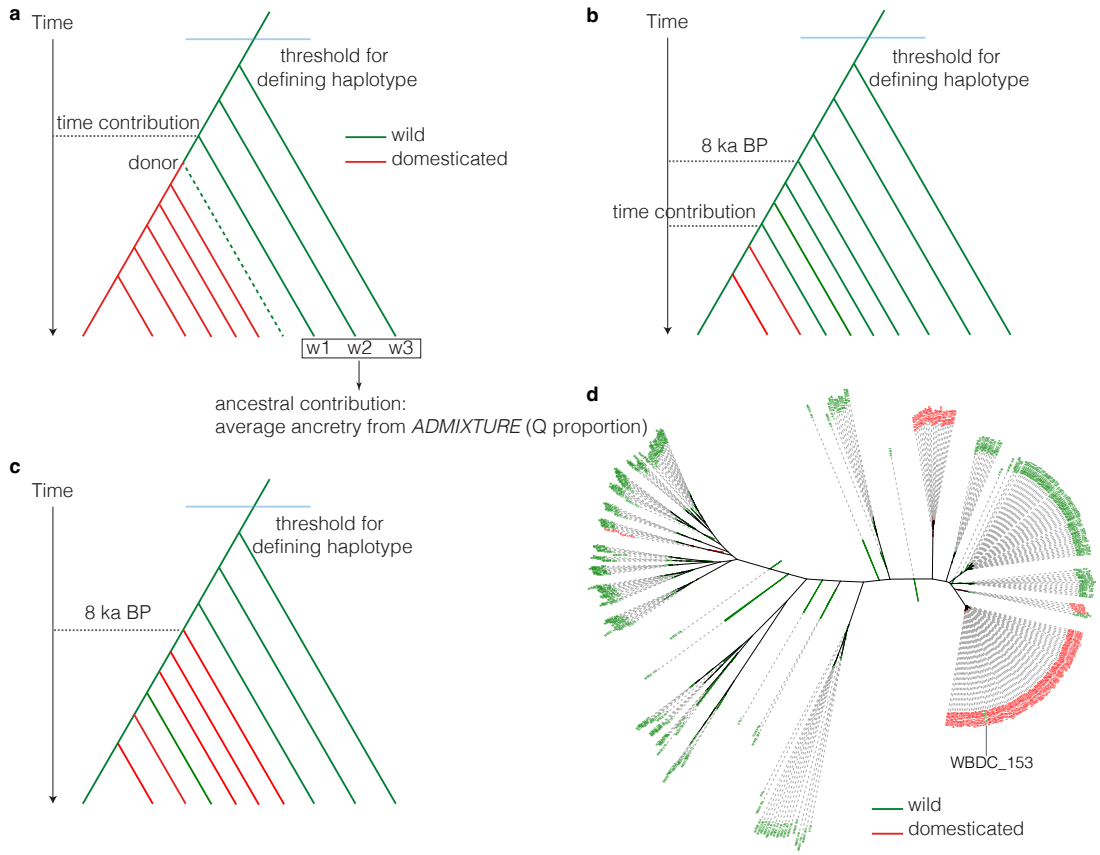


Figure 6: Inference of haplotype origins in time and space. Samples in trees have the same haplotype (H) in a genomic region of interest, i.e. their sequence divergence is below a chosen threshold. Red and green color stand for domesticated and wild samples, respectively. **(a)** Dating in the absence of post-domestication introgression. The time of origin of H was set to set time of divergence of the domesticated branch from the closest wild sample (w1). To find the most closely related wild sample, we compared the results of IntroBlocker runs with different thresholds: 400 SNPs (equivalent to an approximate divergence time of 32,000 years ago), 98 SNPs (8,000 years), 73 SNPs (6,000 years), 49 SNPs (4,000 years), and 24 SNPs (2,000 years). To determine the wild source population of population, the ancestry coefficients

of all wild barleys (w1, w2, w3) with haplotype H were averaged. In panels **(b)** and **(c)**, the divergence time between domesticated and wild carriers of H is less than 8,000 years, indicating recent gene flow, either in the direction wild > crop **(b)** and crop > wild **(c)**. **(d)**. Neighbor-joining tree constructed from SNPs on chromosome 4H, 250 Mb – 300 Mb. The sample WBDC 153 is the only wild barley that shared a haplotype with domesticated samples. The most likely explanation is geneflow in the direction crop > wild.

2.10 Haplotype-based genetic diversity and selective sweeps

Saturation curves for the average number of haplotypes per genomic window were generated as a function of sample size. This was done by randomly selecting k individuals, with k ranging from 1 to 115 for domesticated barley samples and from 1 to 250 for wild barley samples. For each k, 100 random subsets were selected, and the average haplotype number was calculated for each subset. Mean values and 95% confidence intervals were computed using R (version 3.5.1) (R Core Team 2013) with a t-test. Watterson's estimator (θ_W) (Watterson 1975) and the Shannon diversity index (Spellerberg and Fedor 2003) were calculated using custom Perl scripts on haplotype matrices that included only genomic windows with less than 20% missing data. The values of θ_W and the Shannon index across the seven barley chromosomes were visualized with Gnuplot (version 5.2, http://www.gnuplot.info), applying the "smooth bezier" function for trend smoothing.

To identify regions of reduced diversity in domesticated barley compared to wild barley and to search for candidate genes under selection, we avoided bias caused by using a domesticated reference genome (Morex). Instead, we performed read mapping, variant calling, and haplotype inference using the annotated genome of wild barley B1K-04-02 (FT11) (Jayakodi et al. 2024). Regions with a Shannon index ≤1 were considered selective sweeps. The functional impacts of SNPs and indels located within genes in these regions were predicted using SnpEff (version

4.3t) (Cingolani et al. 2012). Variants with high allele frequency differentiation were prioritized as potential selection targets.

The genetic differentiation between populations of domesticated barley was analyzed by calculating the absolute allele frequency difference (AFD) (Berner 2019) for the following population comparisons (**Fig. 7**): NE+EU vs. ETH, NE+EU vs. Asia, ETH vs. Asia, NE vs. EUT, NE vs. EUS, and EUT vs. EUS. In parallel, F_{ST} values were calculated in 100-kb genomic windows (shift: 20 kb) using the same method applied for wild barley. AFD analyses focused on haplotypes derived from high-coverage samples (SNP1), while F_{ST} calculations included all samples, including those with low coverage (SNP2).

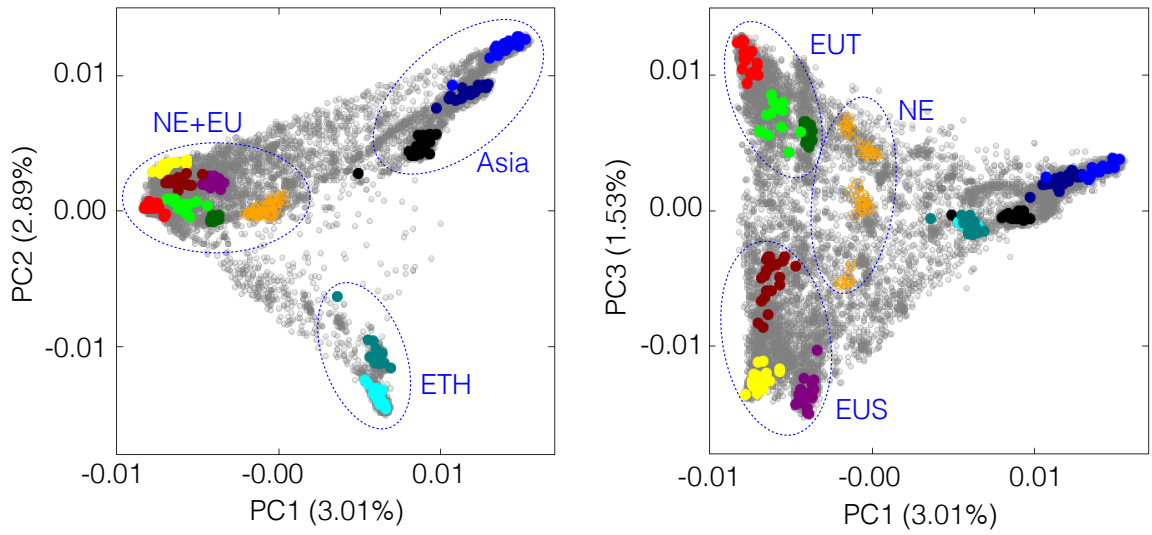


Figure 7: Principal component analysis of 19,778 based on genotyping-by-sequencing data of Milner et al. (2019) (62,888 bi-allelic SNPs). Samples analyzed in this study are shown in non-gray color. Blue circles delineate the groups used for the comparisons: NE – Near East; EU – Europe and Mediterranean Basin; ETH – Ethiopia; Asia – Central and East Asia; EUT – EU two-rowed; EUS – EU six-rowed.

2.11 Demographic history of domesticated barley

Trajectories of effective population size over time were inferred using PSMC (version 0.6.5-r67) (Li and Durbin 2011) with default parameters, based on pseudo-diploid genome sequences derived from pairs of homozygous

barley individuals. A generation time of one year and a mutation rate of 6.13×10^{-9} mutations per site per generation were applied.

1. Pseudo-Diploid Genome Construction for Domesticated Barley Groups:
   A total of 27 pseudo-diploid genomes were generated by combining sequence data from pairs of individuals across three domesticated barley groups: Europe and Near East, East and Central Asia, Ethiopia. Pseudo-diploid genomes from pairs of these groups were used to infer the average demographic history of the entire domesticated barley population.
2. Population-Specific Trajectories:
   Additionally, PSMC analyses were performed on 341 pseudo-diploid genomes created from all possible pairwise combinations of individuals within 15 domesticated barley populations. These analyses provided insights into the population histories of individual domesticated barley subpopulations.

Split times between pairs of domesticated barley populations were estimated by analyzing the distribution of SNP counts in genomic windows (size: 1 Mb, shift: 200 kb). Windows where sample pairs differed by fewer than 300 SNPs (corresponding to a divergence time of approximately 24,470 years) were considered. Only windows with an effective coverage of at least 90% (i.e., ≥900 kb with ≥2-fold coverage in both samples) were included. The SNP number distribution was visualized using frequency polygons with linear binning (50 bins, range: 0–300 SNPs). Divergence times were calculated using the formula:

$$\text{divergence time} = \frac{\text{SNP Number per Mb}}{10^6 \times (2 \times 6.13 \times 10^{-9})}$$

where 6.13×10^{-9} represents the mutation rate (μ) determined for *Brachypodium distachyon*.

The origins of three key genes (*BTR1/2*, *VRS1*, and *NUD*) were inferred by examining the distribution of SNPs within selective sweep haplotypes surrounding these loci. The emergence times of loss-of-function mutations in domesticated barley were also estimated (**Fig. 8**). For each gene, a Neighbor-Joining (NJ) tree was constructed using SNPs from specific intervals within their respective sweep regions:

- *BTR1/2*: 39.4–39.7 Mb on chromosome 3H,
- *VRS1*: 570.5–571.2 Mb on chromosome 2H,
- *NUD*: 525.3–525.7 Mb on chromosome 7H.

These analyses provided insights into the timing and genetic basis of domestication-related traits in barley.

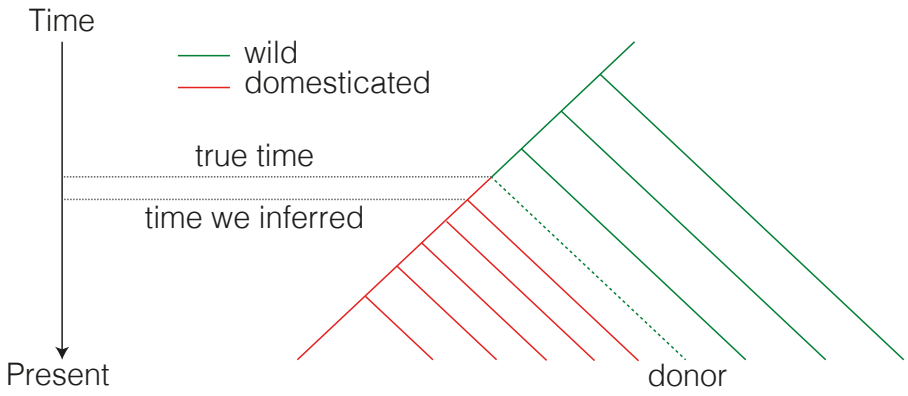


Figure 8: Dating the origin of haplotypes around domestication genes. The figure showed the distributions of sequence divergence (SNPs per Mb) in pairwise comparisons between all domesticated barleys at a locus of interest. The time inferred marks the earliest divergence between extant domesticated samples, but true time must be predating our estimate because the wild lineage (donor) in which the mutation arose is not represented in our data, for example because it became extinct or so rare as to avoid sampling.

2.12 Archaeological excavations

We analyzed ancient DNA sequences of 23 barley grains excavated at three archaeological sites in Israel (**Fig. 9**, **Table 3**). This number includes published

data of 5 barley grains from Yoram Cave (Mascher et al. 2016). The sites Yoram and Timna have been described by Mascher et al. (2016). Abi'or Cave is a medium-sized cave located on the eastern slopes of the Judean Desert, above Jericho, approximately 50 meters below sea level, across from the Karantal Monastery. The excavations at the cave were directed by the late H. Eshel in 1986. It is situated above a larger cave known as "The Spies Cave" and has three openings above it. The cave contains a main long tunnel, approximately 50 meters long, and has revealed archaeological material dating from the Chalcolithic period to the time of the Bar Kochba Revolt (2nd century CE). The cave was found to be heavily disturbed by animals, antiquities robbers, and monks who lived in it during the Islamic and more recent periods.

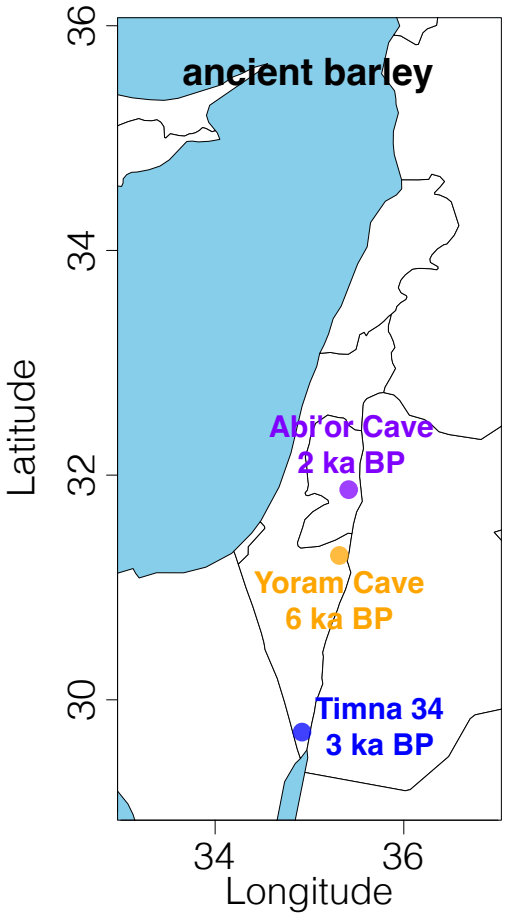


Figure 9: Archaeological sites in the Judean Desert at which ancient barley grains used for ancient DNA extractions were found. Ages of the samples, as determined by radiocarbon dating, are indicated in the figure.

Table 3: The information of ancient barley samples in three archaeological sites.

archaeological sites	period	accessions number	*btr1/2* genotype	*vrs1* genotype	row type
Yoram Cave	Late Chalcolithic (4,500-3,700 BCE)	9	*btr1Btr2*	*Vrs1.b2*	2-rowed
Timna 34	Early Iron Age (1,100-900 BCE)	10	*btr1Btr2*	*Vrs1.b2*	2-rowed
Abi'or Cave	Roman (132-136 CE)	4	*btr1Btr2*	*vrs1.a1*	6-rowed

2.13 Ancient DNA sequencing and analysis

All laboratory procedures, including sampling, DNA extraction, library preparation, and indexing, were conducted in specialized ancient DNA facilities at the University of Tübingen. Each seed was bisected prior to processing: one portion (ranging from 3.6 to 6.5 mg) was allocated for DNA extraction, while the remaining part (weighing 2.6 to 3.4 mg) was submitted for radiocarbon dating at the Klaus-Tschira-Archäometrie-Zentrum, Curt-Engelhorn-Zentrum Archäometrie gGmbH, in Mannheim, Germany. DNA was extracted following a standard protocol for ancient plant material. Dual-indexed, double-stranded DNA libraries were prepared for sequencing (Meyer and Kircher 2010; Kircher et al. 2012).

Six ancient DNA samples (TU697 and JK2281–JK3014) underwent treatment with uracil-DNA glycosylase (UDG) before sequencing (Rohland et al. 2015). Sequencing was carried out on Illumina sequencing devices at IPK Gatersleben, the University of Tübingen, and the Max-Planck Institute for the Science of Human History in Jena. Paired-end reads from each sample were merged using leeHom (Renaud et al. 2014), and then mapped to the MorexV3 genome assembly using Minimap2 (version 2.24) (Li 2021). BAM files were sorted, and

duplicates were marked with Novosort (version 3.06.05) (https://www.novocraft.com/products/novosort/). Nucleotide misincorporation patterns were analyzed with mapDamage2.0 (Jónsson et al. 2013). Variant calling was conducted with BCFtools (v1.15.1) (Danecek et al. 2021) using the command:

```
mpileup -a DP,AD -q 20 -Q 20 --ns 3332
```

The "--variants-only" parameter was omitted in the bcftools call step to output genotypes for all sites. Variants representing G->A and C->T transitions, where G and C are reference alleles and T and A are alternative alleles, were excluded to account for ancient DNA damage signatures.

The resulting SNP matrix was merged with three existing datasets: SNP1 (367 high-coverage samples), SNP2 (302 domesticated barley samples), and a published SNP matrix derived from genotyping-by-sequencing (GBS) data of 19,778 domesticated barleys (Milner et al. 2019). The GBS data had been pre-filtered for a site-level missing rate of less than 20% before merging. The merged SNP1 matrix was used for principal component analysis (PCA) with smartpca (version 7.2.1) (Patterson et al. 2006) using the parameter:

```
lsqproject: YES
```

A neighbor-joining tree was constructed based on SNPs from the pericentromeric region of chromosome 1H (150 to 200 Mb in MorexV3) and included seven ancient DNA samples with high coverage. The merged GBS matrix was used to calculate an identity-by-state (IBS) matrix with PLINK (version 1.9) (Purcell et al. 2007).

The merged SNP2 matrix was applied for two analyses: Ancestry inference using ADMIXTURE (version 1.23) (Alexander et al. 2009) with *K* (the number of ancestral populations) ranging from 2 to 11; and D-statistics calculation using qpDstat from ADMIXTOOLS (version 3.0) (Patterson et al. 2012).

2.14 Six-rowed wild growing barley

A total of 77 accessions from the IPK genebank were genotyped using genotyping-by-sequencing (GBS). DNA was extracted from seedling leaves following the protocol outlined by Milner et al. (2019), with the modification of using 2 ml polypropylene microfuge tubes instead of cluster tubes. GBS libraries were prepared as described by Wendler et al. (2014) and sequenced on an Illumina NovaSeq 6000 at IPK Gatersleben. Additionally, publicly available GBS data for 389 wild barley accessions and 400 domesticated barley accessions were included, as reported by Sallam et al. (2017) and Milner et al. (2019). The 400 domesticated barley accessions were selected from a pool of approximately 20,000 domesticated barley samples to capture a broad range of diversity, ensuring a balanced representation of both wild and domesticated accessions.

GBS reads were mapped to the MorexV3 pseudomolecule reference (Mascher et al. 2021) using Minimap2 (version 2.24) (Li 2021). Variant calling was performed using BCFtools (v1.15.1) (Danecek et al. 2021) with the command:

```
mpileup -q 20 -Q20
```

The resulting variant matrix was filtered as follows:

1. Only bi-allelic SNP sites were retained.
2. Homozygous genotype calls were kept if their read depth ranged from 2 to 50, with any calls outside this range set to missing.
3. Heterozygous genotype calls were retained if both alleles had a read depth of at least 2, with others set to missing.
4. SNP sites with fewer than 20% missing calls and fewer than 20% heterozygous calls were used in further analysis.

Principal component analysis (PCA) was performed with smartpca (version 7.2.1) (Patterson et al. 2006), and only variants with a minor allele frequency (MAF) greater than or equal to 5% were included in the analysis.

A subset of 17 accessions out of the 77 genotyped by GBS was selected for whole-genome sequencing (WGS). Illumina Nextera libraries were prepared using the same DNA as for the GBS and sequenced on an Illumina NovaSeq 6000 at IPK Gatersleben. Publicly available WGS data (Jayakodi et al. 2020; Zeng et al. 2018) from 298 wild and domesticated barley samples were also included in the analysis. Read mapping, variant calling, and filtering were performed as for the GBS data, with the following adjustments: (1) the parameter '--ns 3332' was added to the "bcftools mpileup" command to exclude duplicates and secondary alignments, and (2) homozygous genotype calls were accepted only if the read depth was at least 1×. Pairwise genetic distances, based on identity-by-state (IBS), were calculated using PLINK (version 1.9) (Purcell et al. 2007). These distance matrices were used to construct a Neighbor-Joining (NJ) tree with Fneighbor (Rice et al. 2000) (part of the EMBOSS package, available at http://emboss.toulouse.inra.fr/cgi-bin/emboss/fneighbor). The resulting tree was visualized using Interactive Tree of Life (iTOL) (Letunic and Bork 2021). Principal Component Analysis (PCA) was conducted using smartpca (version 7.2.1) (Patterson et al. 2006). Individual ancestry coefficients were assessed with ADMIXTURE (version 1.23) (Alexander et al. 2009). Prior to ADMIXTURE analysis, linkage disequilibrium pruning was performed using PLINK (version 1.9) (Purcell et al. 2007) with the parameters "--indep-pairwise 50 10 0.1". For each value of *K*, 20 replicate runs of ADMIXTURE were combined using CLUMPP (version 1.1.2) (Jakobsson and Rosenberg 2007) and visualized with Distruct (version 1.1) (Rosenberg 2004). Individuals with more than 80% ancestry from either WWA or WCA were classified as unadmixed wild barleys.

3. Results

3.1 Structure and divergence of wild barley populations

We began with the premise that the current population structure of wild barley reflects its historical structure at the time when humans first domesticated the species. Wild barley (*Hordeum vulgare* subsp. *spontaneum*), a highly genetically diverse taxon, is native to Western Asia. To investigate this, we sequenced 380 wild barley accessions (**Table S1**), primarily sourced from the Wild Barley Diversity Collection, to an average depth of 10-fold using Illumina short-read sequencing. Previous studies consistently identify isolation-by-distance as the primary mechanism driving population differentiation in wild barley (Russell et al. 2016; Jakob et al. 2014). To explore population structure, we employed model-based ancestry estimation (**Fig. 10a**) and principal component analysis (PCA) (**Fig. 10c**). These approaches revealed five distinct populations, whose geographic distributions form a gradient spanning the Southern Levant (SL), Syrian Desert (SD), Northern Levant (NL), Northern Mesopotamia (NM), and Central Asia (CA) (**Fig. 10b**). The populations exhibited varying levels of genetic diversity (**Fig. 10d**, **Table 4**). The SD population showed particularly low genetic diversity and high differentiation from other populations, likely due to elevated genetic drift in this region.

If there were no recombination and gene flow, the number of sequence variants between two genomes would inform directly about divergence times. However, three examplary cases (**Fig. 11**) demonstrate that this straightforward model does not hold for barley. When we analyzed single-nucleotide polymorphisms (SNPs) within 1 Mb windows and examined their distributions, we found notable local differences in divergence times between some sample pairs, particularly between distal and proximal regions (**Fig. 11**). In barley and its relatives, such as wheat and rye, proximal regions known as “genetic centromeres” are characterized by

extensive non-recombining segments, reduced gene content, and markedly suppressed recombination rates (Mascher et al. 2017; International Wheat Genome Sequencing Consortium (IWGSC). 2018; Rabanus-Wallace et al. 2021). In domesticated barley, these regions also exhibit reduced sequence diversity (Russell et al. 2016; Mascher et al. 2017) (**Fig. 12**). Wild barley presents a more complex picture. Between-population comparisons revealed unimodal distributions of divergence times in the distal regions of all chromosomes, peaking around 600 thousand years before present (ka BP) (**Fig. 13**). This period aligns with a significant reduction in effective population size during the same timeframe (**Fig. 13**). Historical fluctuations in population sizes were further corroborated by effective population size trajectories reconstructed using PSMC (**Fig. 13**). These analyses suggest that all wild barley populations experienced a recovery from a bottleneck that occurred between 2000 and 500 ka BP. A subsequent bottleneck, estimated to have occurred between 110 and 12 ka BP, coincided with the Last Glacial Maximum (LGM).

Divergence time distributions in proximal genomic regions were multimodal and varied across chromosomes (**Fig. 14**), presenting a complex pattern that resists simple explanations. This could be attributed to the scarcity of centromeric haplotypes and their persistence as intact linkage blocks over evolutionary timescales. Exploiting this unique characteristic, we explored the relationship between the divergence of extensive centromeric haplotypes, the genetic divergence among individuals, and the split times between wild barley populations (**Fig. 15**). To investigate, we focused on SNPs located within pericentromeric regions (centromere ±25 Mb) to estimate pairwise divergence times among wild barley individuals. We then constructed a tree (**Fig. 16c**) depicting the relationships among wild barley populations based on their most recent splits. Although this tree clearly simplifies the complexity of population relationships, the divergence time distributions are multimodal. (**Fig. 16a**). Peaks in these distributions aligned with fluctuations in global surface temperature (**Fig.**

16b), suggesting repeated cycles of habitat colonization, population contraction or extinction, recolonization, and secondary contact between populations. For instance, the shared ancestor of the Syrian Desert, Northern Mesopotamian, and Central Asian populations diverged from the Northern Levantine lineage approximately 120 thousand years before present, possibly during a warm period that opened new habitats. Subsequent divergence of the Northern Mesopotamian and Central Asian populations occurred around 17 ka BP, consistent with paleoclimate models (Jakob et al. 2014) indicating the absence of wild barley from Central Asia as recently as 21 ka BP. The early divergence of the Southern Levantine population supports its status as a potential glacial refugium (Jakob et al. 2014). Notably, some Southern Levantine wild barley harbors a centromeric haplotype that diverged from other centromeric haplotypes around 900 ka BP (**Fig. 14**, **Fig. 17a**), a much deeper split than observed within or between other wild barley populations. This "relict" haplotype may have escaped genetic drift due to larger effective population sizes in the Southern Levant or may have been maintained by selection for an adaptive advantage (**Fig. 17b-c**). Supporting the latter hypothesis, this relict haplotype is prevalent in many domesticated barley populations (**Fig. 17d**). Fang et al. (2014) suggested that the unusually high differentiation observed on chromosome 5H in wild barley might be linked to a large pericentric inversion in this region. While our data confirm the presence of inversions, they do not span entire haplotypes and are also found in other haplotypes (**Fig. 17e**). This indicates that structural variation alone is unlikely to account for the persistence of the relict haplotype over evolutionary timescales.

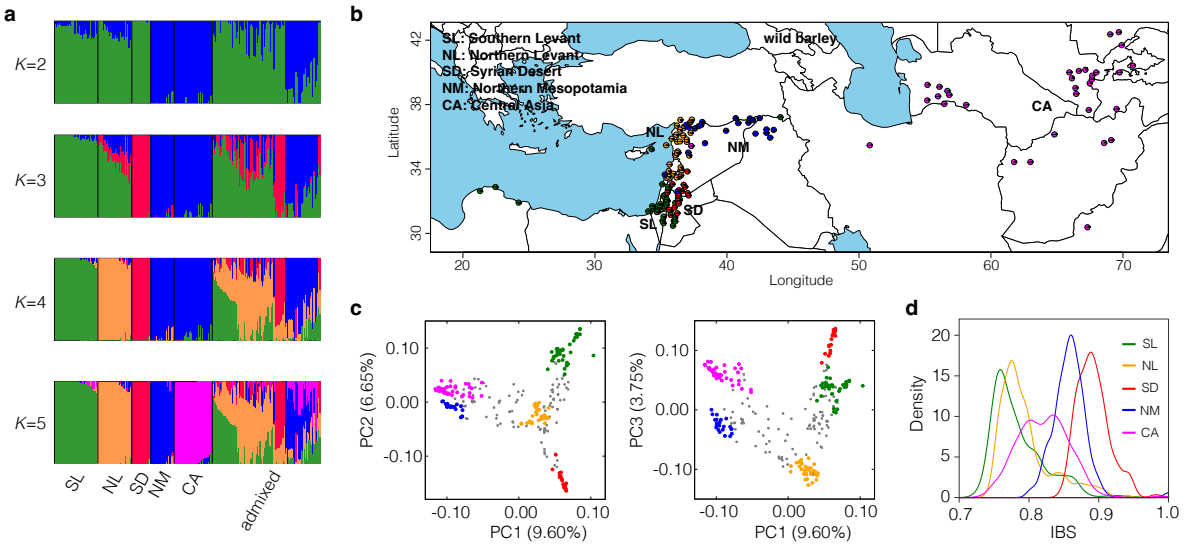


Figure 10: Diversity panel and population structure of wild barley. (a) Individual ancestry coefficients in ADMIXTURE with the number of ancestral populations (*K*) ranging from 2 to 5. Individuals with ≥ 85% ancestry form were considered un-admixed samples. **(b)** Collection sites and population structure of 143 wild barley genotypes with precise geographical locations. Pie charts show the results of model-based ancestry estimation with ADMIXTURE (*K*=5) and are plotted at approximate collection sites. Jitter was added to avoid overlaps between nearby accessions. Only unadmixed samples, i.e. those whose major ancestry components was ≥ 0.85 are shown. **(c)** Principal component analysis (PCA) based on 37.14 million bi-allelic SNPs with a MAF > 5%. The first three PCs are shown (PC1 vs. PC2 and PC1 vs. PC3). Unadmixed samples are colored to their major ancestry component in panel **(a)**. **(d)** Distribution of pairwise identity-by-state (IBS) in the five wild barley populations.

Table 4: Nucleotide diversity (π) and Watterson's estimator (θ_W) in wild barley populations.

	chromosome	SL (n=41)	NL (n=32)	SD(n=18)	NM (n=22)	CA (n=36)	all (n=251)
π (10^{-3})	chr1H	4.38	3.77	1.50	2.35	2.79	4.31
	chr2H	4.90	3.58	1.83	2.74	2.43	4.34
	chr3H	4.60	4.47	2.39	2.32	3.51	4.82
	chr4H	4.31	3.13	2.12	2.34	3.14	4.65
	chr5H	4.25	3.24	2.39	1.72	1.59	4.42
	chr6H	4.24	3.77	2.62	2.38	3.94	4.26
	chr7H	4.99	3.89	1.20	1.87	3.14	4.76
	entire genome	4.53	3.69	2.01	2.26	2.93	4.51
θ_W (10^{-3})	chr1H	6.28	4.42	1.88	2.22	2.61	8.39
	chr2H	6.98	4.28	2.01	2.87	2.18	8.56
	chr3H	6.37	5.03	2.47	2.38	3.62	8.98
	chr4H	6.03	3.58	1.90	2.48	2.78	8.63
	chr5H	6.27	4.12	2.44	1.94	1.64	8.27
	chr6H	5.85	4.51	2.85	2.58	3.67	8.44
	chr7H	7.66	4.45	1.83	2.12	3.11	9.02
	entire genome	6.50	4.33	2.19	2.38	2.80	8.62

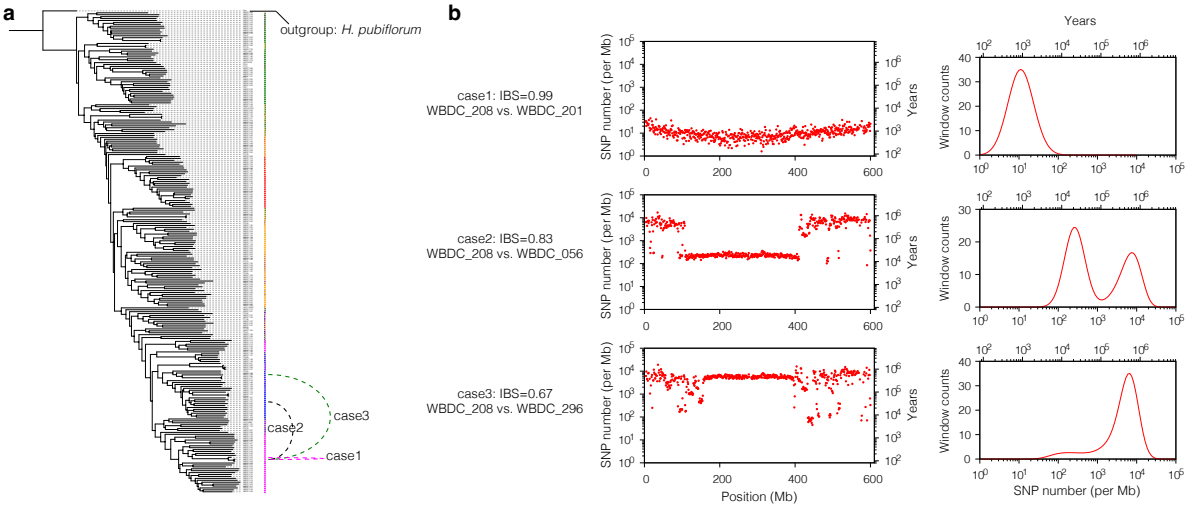


Figure 11: Sequence divergence between wild barleys in different genomic compartments. **(a)** Neighbor-joining tree of 251 wild barley genotypes computed from 37.14 M bi-allelic SNP markers with a MAF > 5%. *H. pubiflorum* was used an outgroup. Pie charts to the right show ancestry coefficients as determined by ADMIXTURE **(Fig. 10)**. **(b)** The left-hand panels show

the sequence divergence (SNPs per non-overlapping 1 Mb genomic windows) on chromosome 4H between three pairs of wild barleys. Case 1 each compares two members from the same population (CA, WBDC208 vs. WBDC201). In cases 2 and 3, WBDC 208 is compared to two members of the NM population, WBDC296 and WBDC056. The right-hand panels show the respective distribution densities.

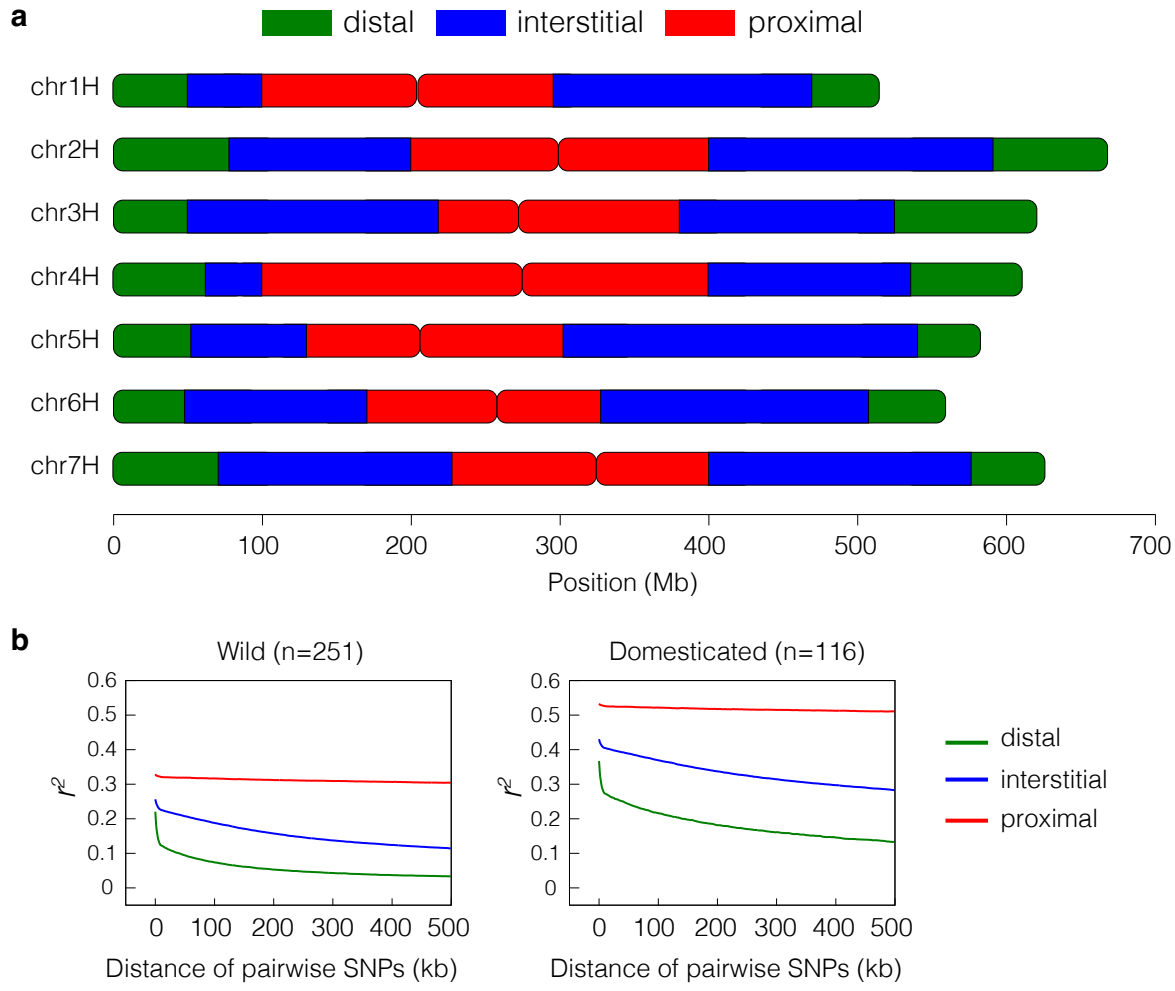


Figure 12: Decay of linkage disequilibrium (LD) in three genomic compartments. (a) The distal, interstitial and proximal regions were defined based on differences in recombination rate following Mascher et al. (2017). **(b)** LD decay in the three genomic compartments thus defined in wild and domesticated barley.

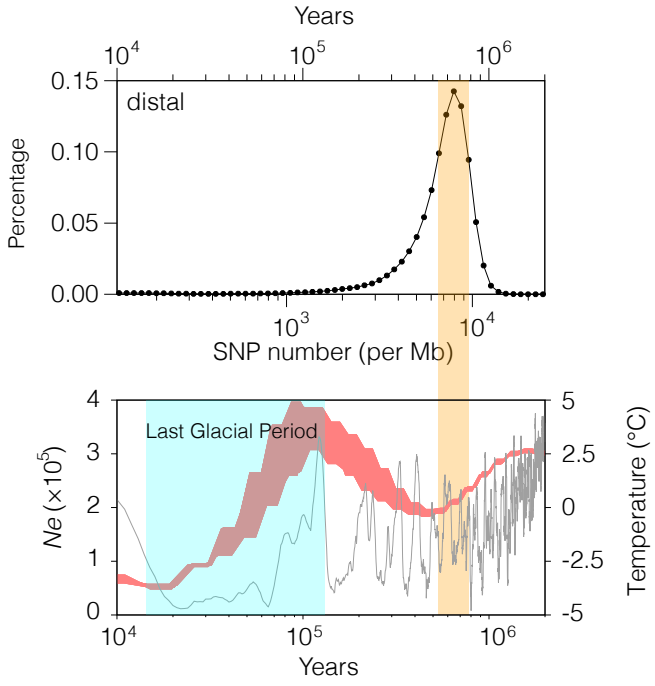


Figure 13: Distribution of sequence divergence (SNPs per Mb) in distal region and PSMC demographic history. The top panel shows the distribution of pairwise sequence divergence for all sample pairs in distal regions of the genome. The bottom panel shows the historic trajectories of effective population sizes in wild barley as inferred by PSMC (red) and global average surface temperatures (Snyder 2016) (gray). The orange shading marks a simultaneous decline of population size and temperature that corresponds to peaks in the SNP distribution. The last glacial period (100 ka BP to 12 ka BP) is marked by blue shading.

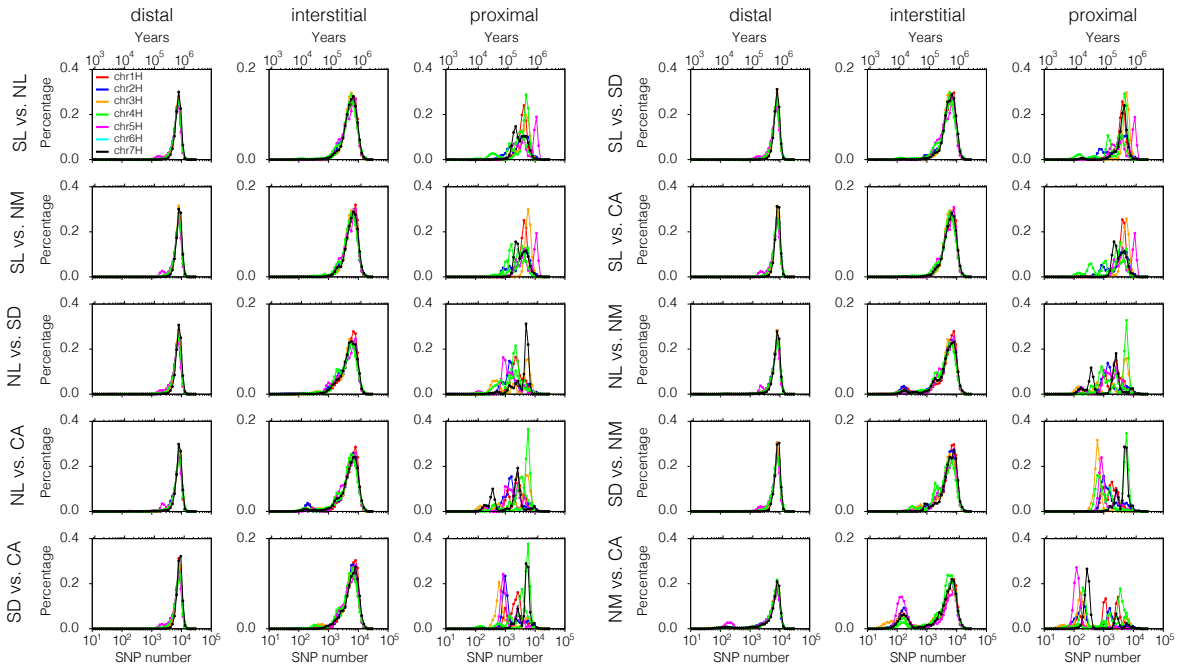
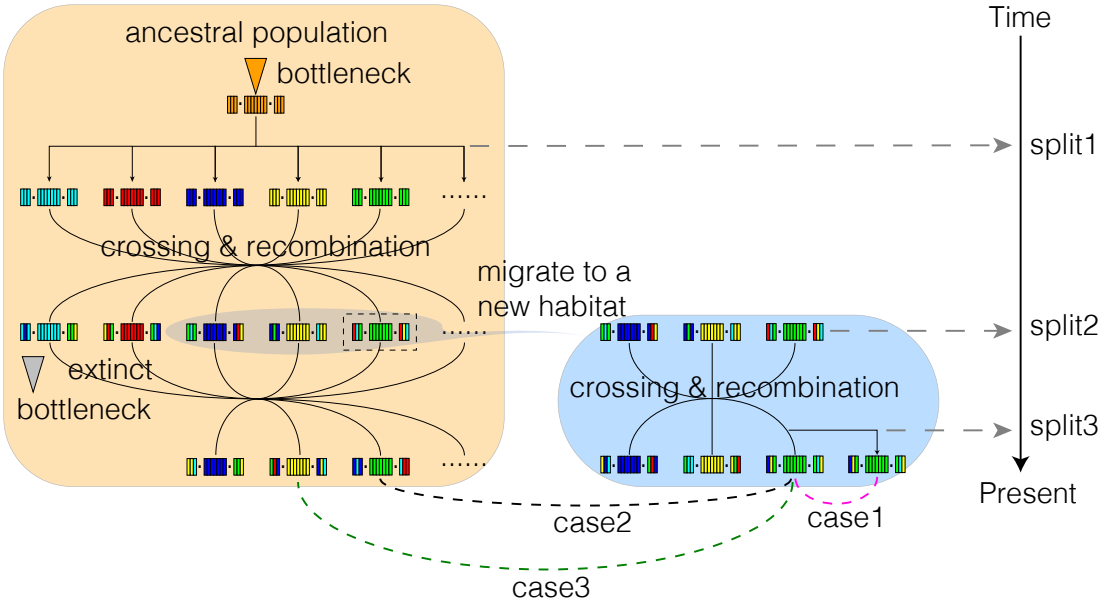


Figure 14: Distributions of sequence divergence (SNPs per Mb) in three genomic compartments in pairwise comparisons between members of different wild barley populations.



case1: SNP number distribution is unimodal. The distal and proximal region both indicated split3.
case2: SNP number distribution is bimodal. The distal region indicated split1 while proximal region indicated split2.
case3: SNP number distribution is unimodal. The distal and proximal region both indicated split1.

Figure 15: Patterns of sequence divergence differ along the genome. This figure complements **Fig. 11.** The three numbered cases shown in both figures correspond to each other. The groups of colored bars each represent a single chromosome of different wild barley individuals. The outer bars stand for genomic windows in distal regions, the inner ones for those in proximal regions. After an initial bottleneck an ancestral population of wild barley split into structured subpopulations (represented by different colors). This process may have occurred independently in different mutually isolated populations (e.g. in the Southern Levant or Central Asia). Here, one local ancestral population is shown as an example. Gene flow between subpopulations reshuffled haplotypes over time (left-hand part). As the population colonized new habitats, bottlenecks occurred ancestral haplotypes were lost (right-hand part). Since the recombination rate in proximal regions is low, these can be considered as a single recombinational unit (~haplotype block). Because proximal regions are physically extensive, long shared haplotype blocks strongly influence the distribution of sequence divergence in windows of fixed physical size. In case1, two individuals are compared that come from the

same extant wild barley population and neither has received recent gene flow from other populations. They share the same haplotype blocks in both distal and proximal regions and the distribution of sequence divergence is thus unimodal. In case2, two individuals come from a different population. Recombination and gene flow have reshuffled and broken up haplotypes in distal regions, whereas the proximal haplotypes trace back to a common ancestor at time point split2 and have remained intact in this scenario. Hence, the distribution of sequence divergence is bimodal: the peak in distal regions corresponds to early divergence (split1), whereas the peak in proximal regions reflects later divergence (split2) of the two lineages that lead to either individual. In case 3, also the proximal regions trace back to different ancestral haplotypes, which diverged early (split1). Hence, the distribution of sequence divergence along the genome is unimodal with a single early peak.

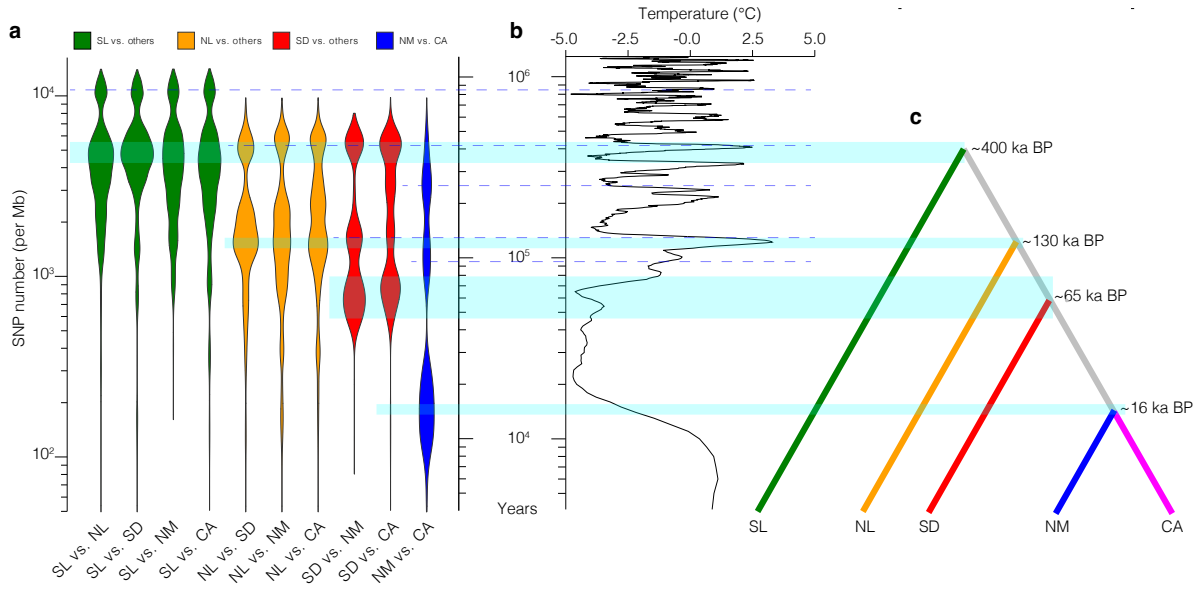


Figure 16: Evolutionary history of wild barley. (a) Violin plots showing the distributions of pairwise sequence divergence in proximal regions of five wild barley populations. Blue shading highlights the peaks in the distribution that mark the most recent divergence between pairs of populations. Earlier such events are marked by dashed lines. **(b)** Global average surface temperatures (Snyder 2016) in the past 2 million years. **(c)** The divergence of barley populations (most recent inferred split times) is represented as a tree.

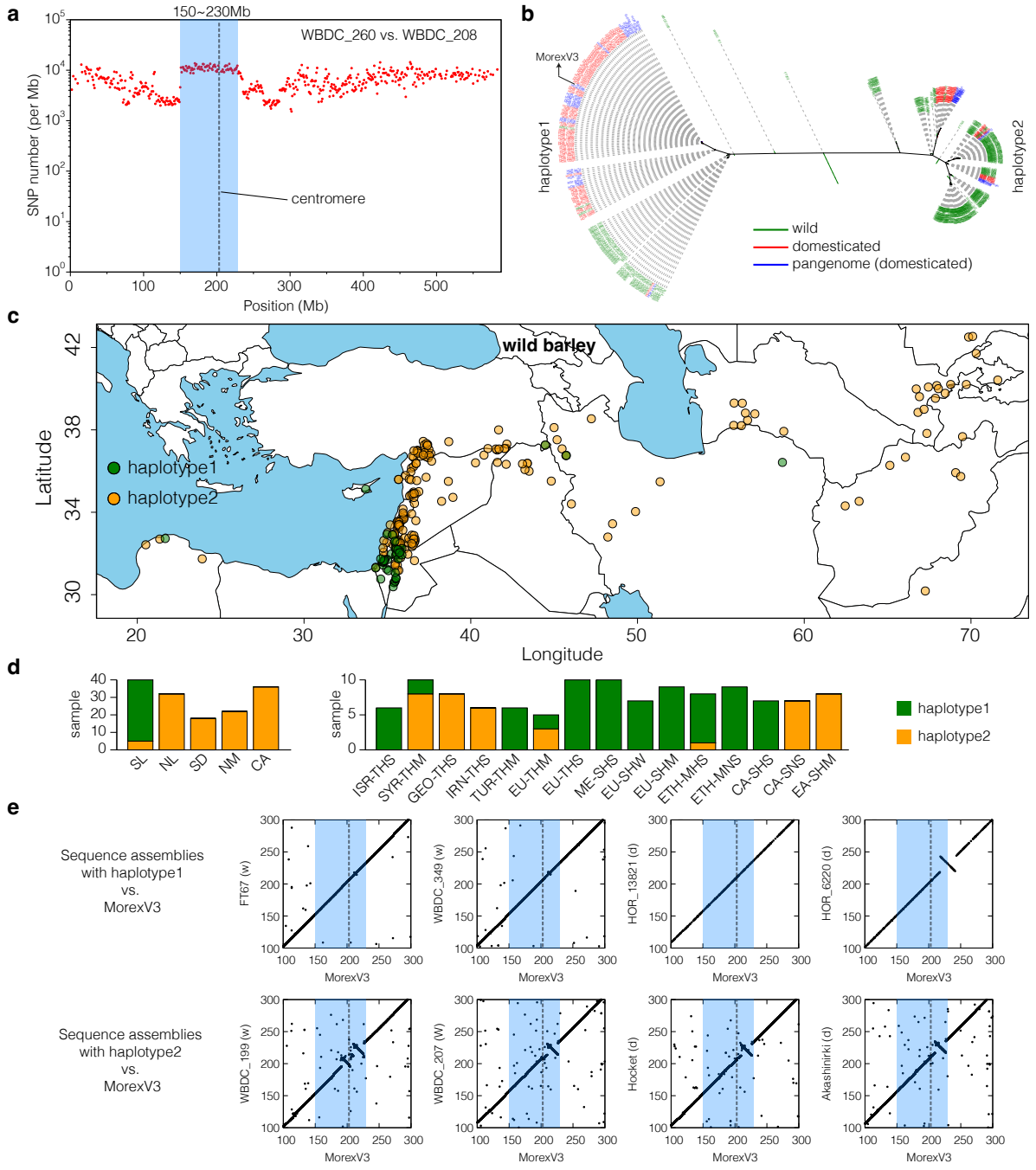


Figure 17: Deeply diverged pericentromeric haplotypes on chromosome 5H. (a) Sequence divergence (SNPs per Mb) on chromosome 5H between wild barleys WBDC260 (SL) and WBDC208 (CA). **(b)** Neighbor-joining tree of wild and domesticated barleys including 48 samples from the pangenome of Jayakodi et al. (2024) **(c)** Collection sites of wild barleys carrying either haplotype. The "green" haplotype is most common in the SL population. **(d)** Frequencies of both haplotypes in wild and domesticated barley populations. **(e)** Alignments on chromosome 5H, 100 to 300 Mb between sequence assemblies of wild (w) and domesticated (d) pangenome accession with either haplotype to the MorexV3 reference. The accession names are indicated on the y-axis. In panels **(a)** and **(e)**, the boundaries of divergent haplotypes (150

to 230 Mb) are marked by blue shading and the dashed line at 205 Mb indicates the position of the centromere in the MorexV3 reference.

3.2 A haplotype-based view of barley evolution

To incorporate domesticated barley into our analysis, we selected 302 accessions (**Fig. 18, Table S2**) from a comprehensive collection of 19,778 domesticated barley accessions (Milner et al. 2019). Of these, 116 were sequenced to approximately 10-fold whole-genome coverage, and 186 were sequenced to about 3-fold whole-genome coverage. We analyzed these data using IntroBlocker (Wang et al. 2022), revealing that, similar to wheat, sequence divergence in domesticated barley exhibited a bimodal distribution (Wang et al. 2022). This pattern held true for both distal and proximal genomic regions (**Fig. 19, 20**). The more recent divergence peak, at approximately 98 SNPs per megabase (~8,000 years of divergence), corresponds to a bottleneck associated with the coalescence of many haplotypes into shared ancestors within the hypothetical founder population(s) of domesticated barley. The older peak, at approximately 6,500 SNPs per megabase (~530,000 years), aligns with the divergence patterns observed in wild barley and arises from comparisons between haplotypes that diverged before the domestication event (**Fig. 19, 20**). To distinguish haplotypes that diverged before domestication from those that split after domestication, we established a threshold of 400 SNPs per 1-Mb window, which corresponds to a divergence time of approximately 32,000 years (**Fig. 19, 20**). While exemplary visualizations were drawn using a 5-Mb window size (**Fig. 21**), subsequent analyses were conducted using 100-kb windows. This finer scale was chosen after evaluating haplotype lengths in regions surrounding a key domestication gene (**Fig. 5**).

A notable feature of the whole-genome ancestral haplotype group (AHG) maps in barley is the presence of long centromeric haplotypes shared between wild and domesticated barley. This shared haplotype pattern provides clear visual evidence

supporting the mosaic ancestry of domesticated barley (**Fig. 21**). Due to the lower diversity of haplotypes in pericentromeric regions compared to distal regions in both wild and domesticated barley, the diversity panel used in this study captures nearly all pericentromeric haplotypes but does not reach saturation for distal regions (**Fig. 22a**). For instance, domesticated barley has a single predominant pericentromeric haplotype on chromosome 1H, which is primarily found in Central Asian wild barley (**Fig. 21**). Overall, 55.9% of domesticated barley haplotypes are shared with at least one wild barley sample. Conversely, only 7.0% of wild barley haplotypes overlap with domesticated barley (**Fig. 23a**). Saturation analysis suggests that a larger sample of wild barley genotypes could reveal additional shared haplotypes (**Fig. 23b**). However, some haplotypes unique to domesticated barley, particularly in distal regions, may lack a wild counterpart. These domesticate-specific haplotypes are often rare (**Fig. 23c**) and might have originated post-domestication through recombination of wild-derived haplotypes or due to genetic drift eliminating their wild counterparts. As expected following a domestication bottleneck, the haplotype frequency spectrum differs significantly between wild and domesticated barley. Common haplotypes (with major allele frequencies above 20%) are rare in wild barley but more prevalent in domesticated barley (**Fig. 22b**). Despite this, 79% of domesticated haplotypes that have a recognizable wild counterpart occur at low frequencies (<5%) in wild barley populations (**Fig. 23d**). Seven genomic regions were identified where haplotype diversity is markedly reduced in domesticated barley compared to wild barley (**Fig. 24**). High genetic differentiation between wild and domesticated barley is apparent at haplotype levels (**Fig. 25**). This differentiation complicates the mapping of selection sweeps through outlier scans; on average, 7.5% of the genome lacks shared haplotypes in pairwise comparisons between domesticated populations (**Table 5**). Instead of widespread adaptive evolution, we propose that this pattern likely reflects local lineage sorting.

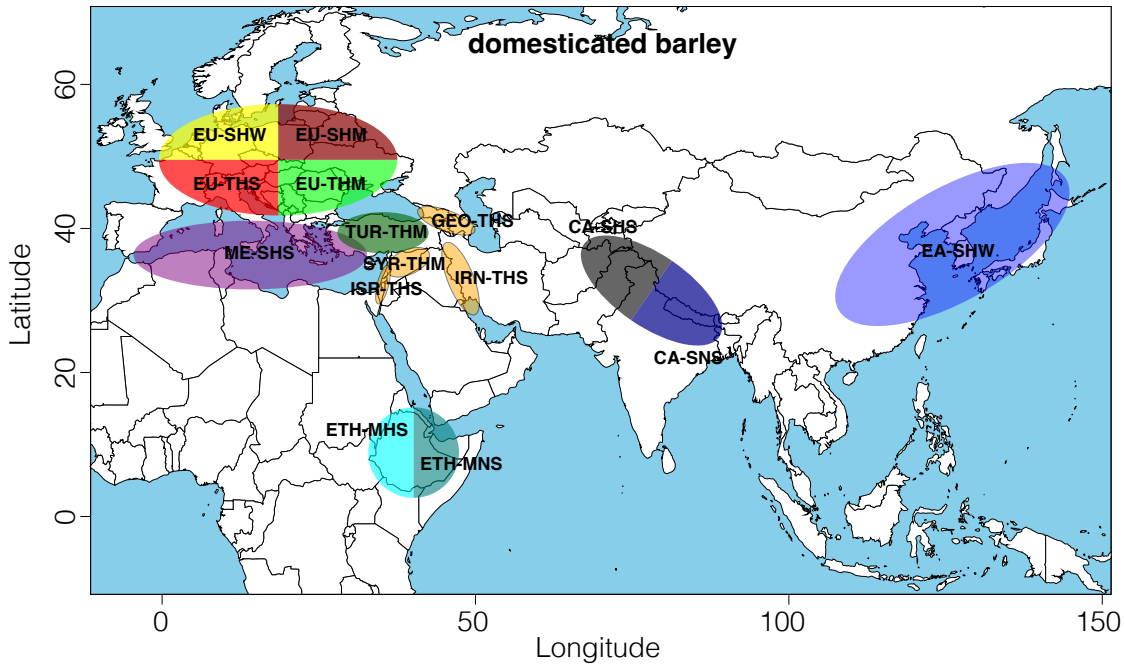


Figure 18: Diversity panel of domesticated barley. Assignment to macrogeographic regions of 15 populations inferred from GBS data of 19,778 domesticated barley (Milner et al. 2019). The population names encode the samples' most common origin and their predominant morphological and phenological characters (row type, lemma adherence, annual growth habit) as detailed in **Table 1 and Table S2**.

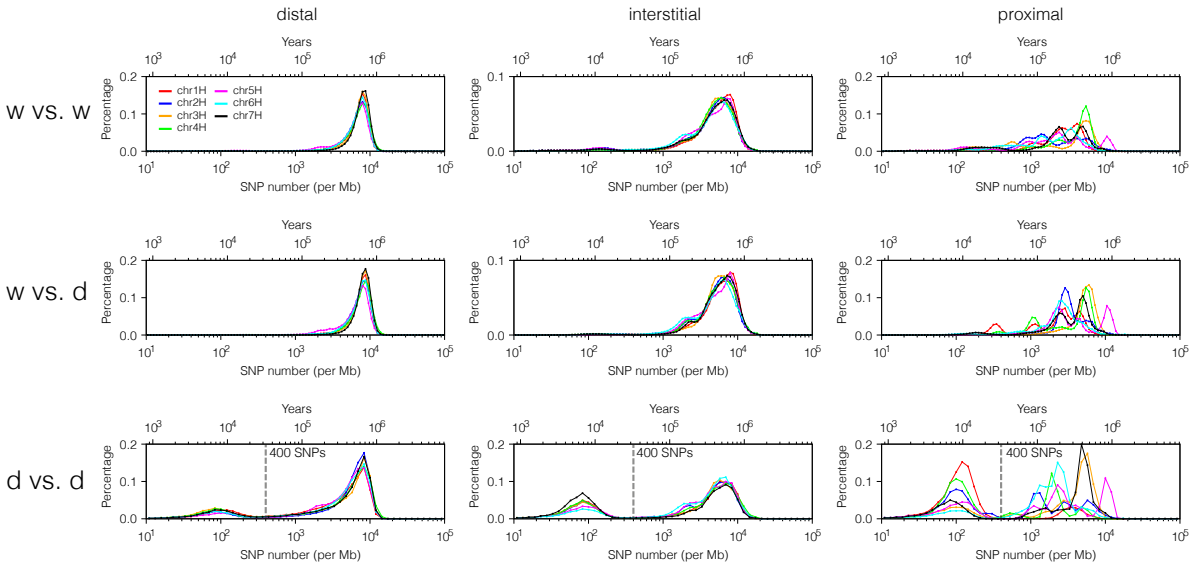


Figure 19: Finding a threshold for defining AHG. Sequence divergence (SNPs per Mb) in pairwise comparisons between wild (w) and (d) domesticated barleys. A value of 400 SNPs per Mb was chosen as the threshold to separate haplotypes in IntroBlocker.

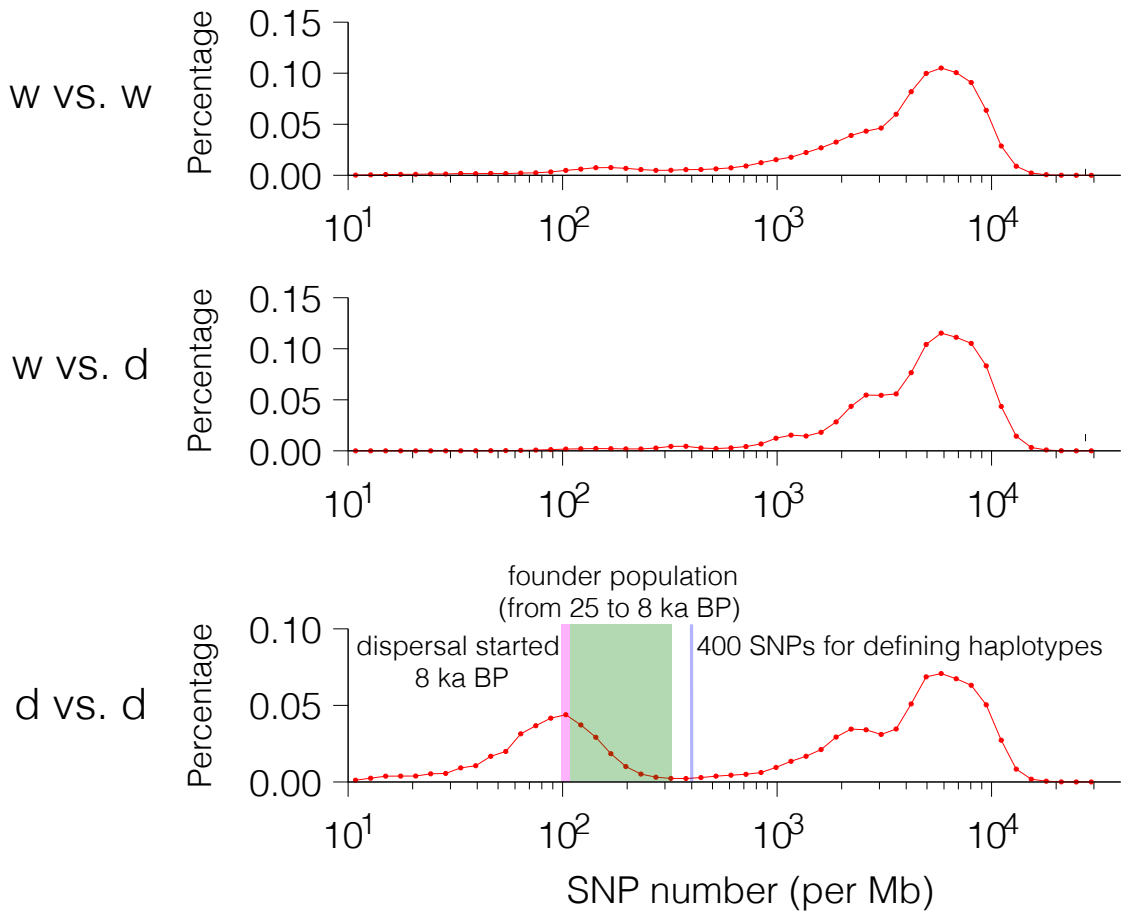


Figure 20: Distribution of sequence divergence (SNPs per Mb) between pairs of wild (w) and domesticated (d) barley. The blue line marks the threshold (400 SNPs per Mb) to delineate the post- and pre-domestication origin of haplotypes in IntroBlocker. The green shading marks the persistence of a hypothetical founder population of diverse wild ancestry that split up into geographically isolated populations from 8 ka BP onwards (purple line).

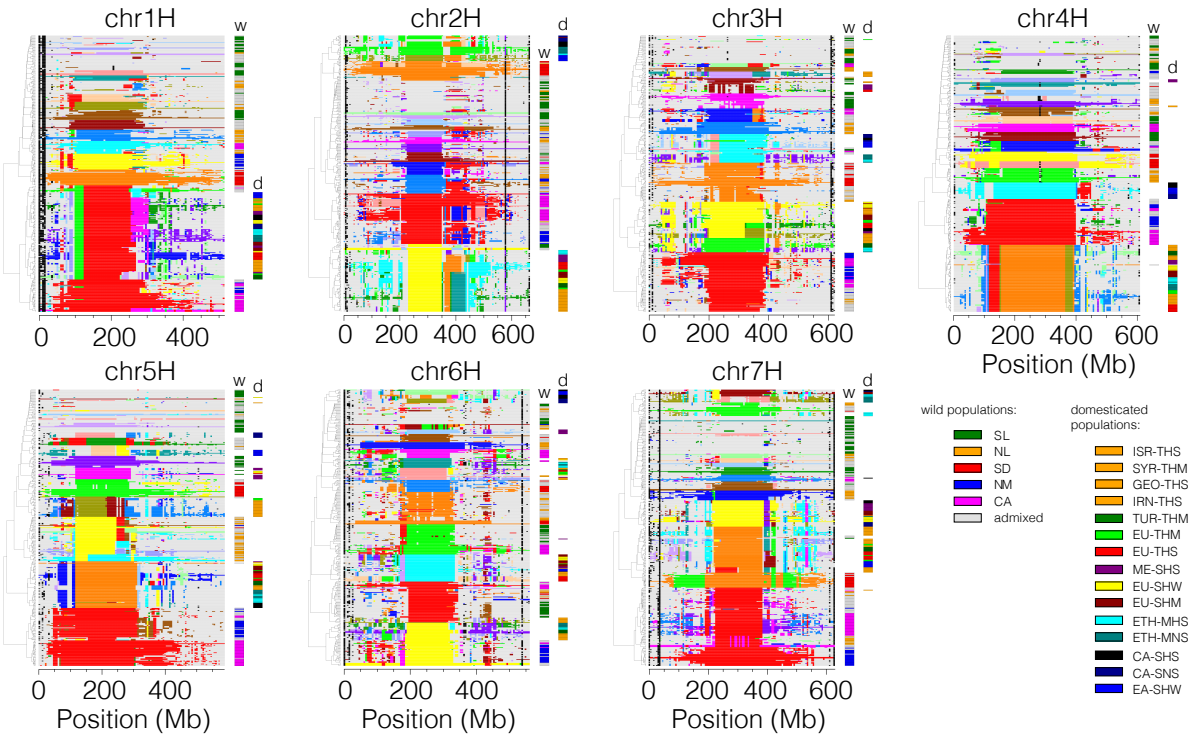


Figure 21: Ancestral haplotype groups (AHGs) in domesticated barley. Mosaic view of AHGs on the seven chromosomes of barley. The data shown are from an IntroBlocker run with a 5 Mb window size (shift: 5 Mb). Colors were assigned to 20 most frequent AHGs by IntroBlocker in semi-supervised mode giving priority to wild over domesticated samples. Black color indicates missing data, gray stands for less frequent haplotypes. Colored bars on the right-hand side of each sub-panel assign samples to wild (w) or domesticated (d) subpopulations according to the legend at the bottom right.

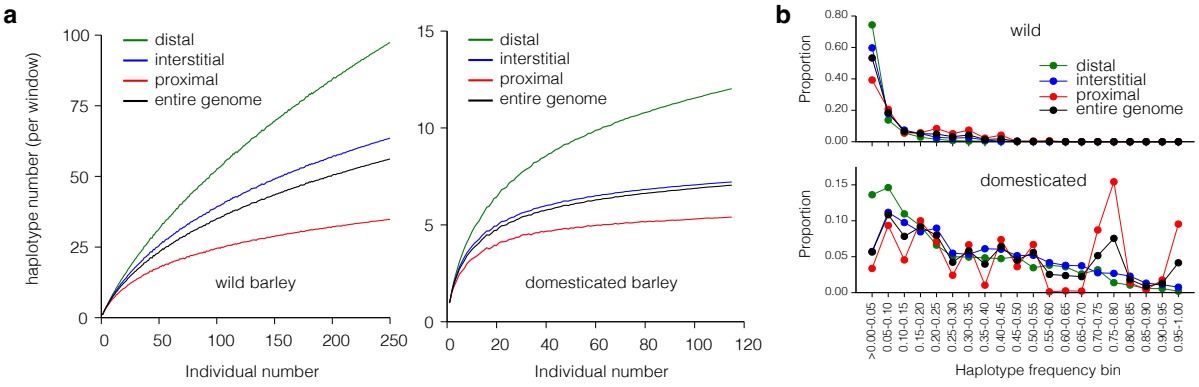


Figure 22: Haplotype-based diversity statistics. (a) Number of observed distinct haplotypes per genomic window in wild and domesticated barley as a function of sample size. The solid line and shaded area represent, respectively, the average and 95% confidence interval of 100

random sub-samples. **(b)** Haplotype frequency spectra (bin size: 0.05) in wild and domesticated barley in different genomic compartments.

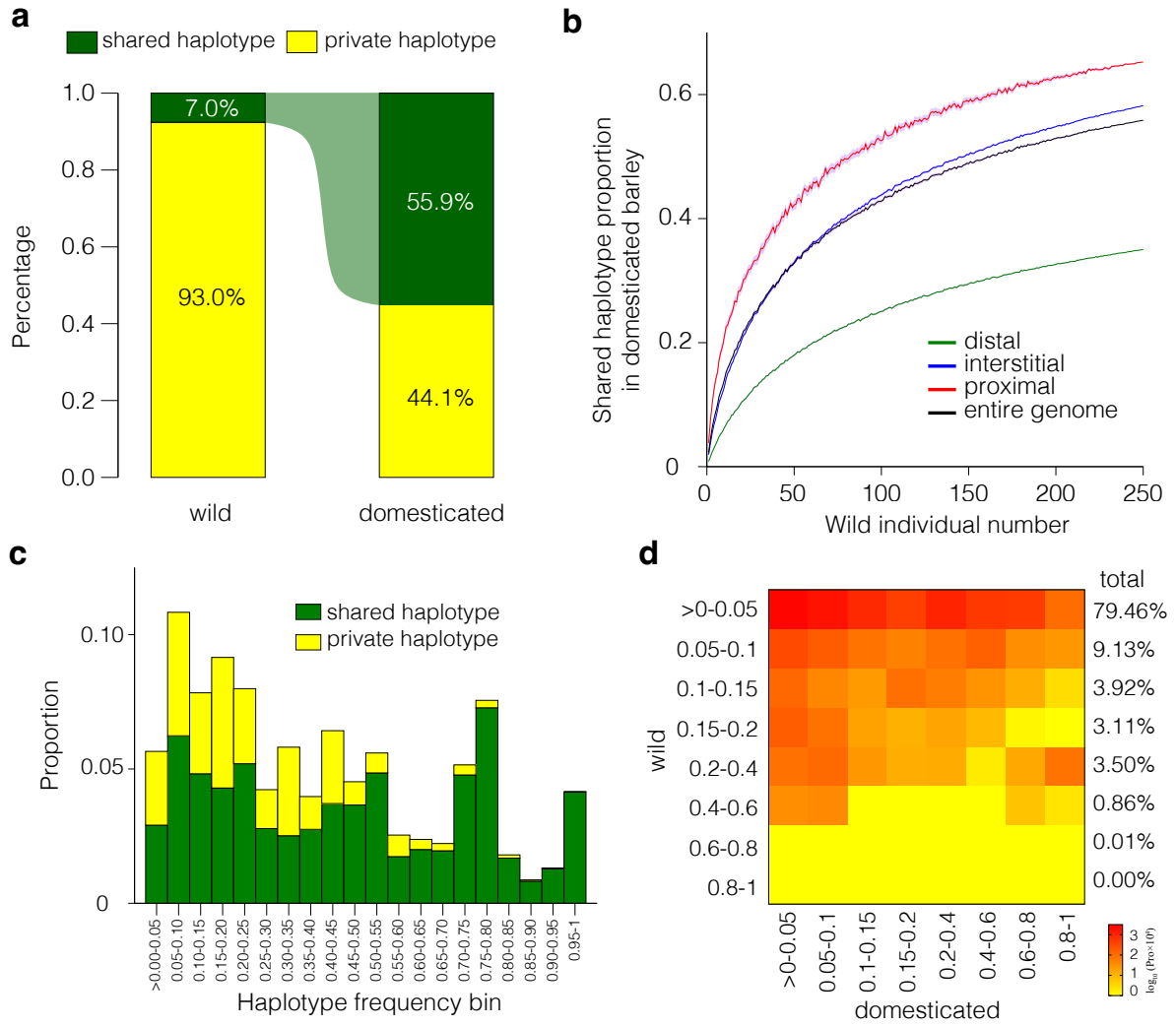


Figure 23: Haplotype diversity in wild and domesticated barley. (a) Proportions of shared and private haplotypes in wild and domesticated barley. **(b)** Saturation curves show how the proportion of shared haplotypes in different genomic compartments increases with the number of potential wild barley counterparts. Solid lines and shading denote the average and 95% confidence intervals of 100 random samples. **(c)** Proportions of haplotypes of domesticated barleys that are shared with wild barley at different frequency bins. **(d)** Normalized two-dimensional haplotype frequency spectrum in wild and domesticated barley. The value (percentage) in each cell was determined by dividing counts by the total number of shared haplotypes. Percentages at the right margins indicate the relative sizes of frequency bins in domesticated barley (row sums), e.g. 79.46% of shared haplotypes occurs at 5% frequency or less in wild barley.

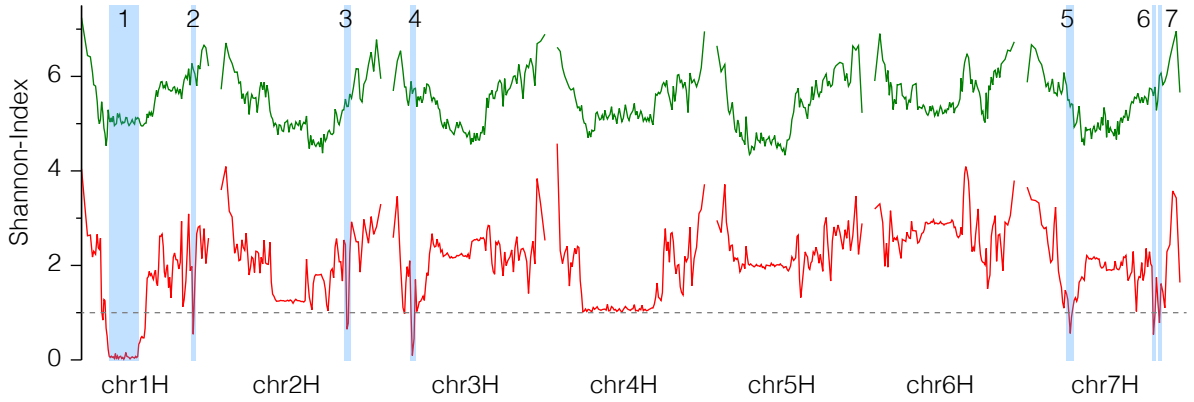


Figure 24: Haplotype-based Shannon indices in wild and domesticated barleys. Seven regions with values below 1 in domesticated samples (dashed line) were defined as putative selective sweeps. The genome sequence of B1K-04-12 was used as a reference for this analysis.

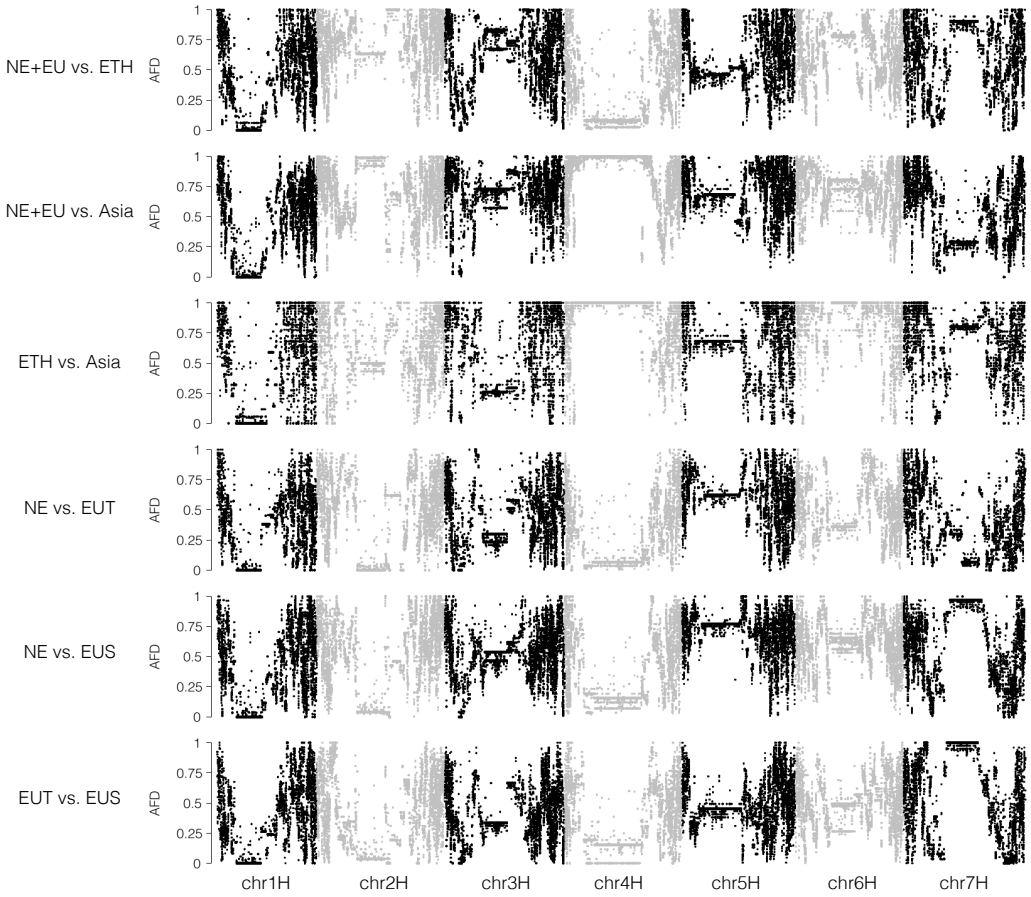


Figure 25: Absolute allele frequency difference (AFD) between different domesticated barley populations in sliding windows (size: 100 kb, shift: 20 kb) along the genome. AFD was computed on the haplotype matrix of high-coverage (~10x) samples

Table 5: AFD between pairs of domesticated barley populations.

population pairs	≥0.5	≥0.6	≥0.7	≥0.8	≥0.9	=1
NE+EU vs. ETH	56.28%	47.89%	36.69%	23.93%	13.61%	3.46%
NE+EU vs. Asia	74.25%	66.64%	50.62%	36.56%	22.41%	10.38%
ETH vs. Asia	73.34%	62.68%	49.52%	43.34%	32.07%	25.22%
NE vs. EUT	40.36%	30.48%	14.39%	8.95%	3.67%	0.41%
NE vs. EUS	54.31%	44.13%	29.50%	16.56%	8.60%	0.60%
EUT vs. EUT	31.89%	23.64%	15.09%	10.04%	7.09%	4.90%
average	55.07%	45.91%	32.64%	23.23%	14.57%	7.50%

3.3 The origins of domesticated haplotypes in time and space

We investigated the temporal and spatial origins of haplotypes in domesticated barley by using IntroBlocker (Wang et al. 2022) with various thresholds corresponding to divergence time brackets. This allowed us to identify which extant wild barley genomes most closely resemble the haplotypes present in domesticated barley. The resulting genome-wide map of spatiotemporal relationships further supports the mosaic genomic composition of domesticated barley (**Fig. 26**). This mosaic pattern emerged early in barley domestication. Approximately 91% of domesticated barley haplotypes with wild counterparts diverged between 32,000 and 8,000 years before present (ka BP), a timeframe spanning the emergence of the immediate wild progenitor and the early stages of domestication (**Fig. 27a**). Less than 9% of haplotypes reflect more recent gene flow. All five wild barley populations contributed to the genomic makeup of domesticated barley, albeit in varying proportions. Wild barley populations from the Southern and Northern Levant and Central Asia each contributed between 20% and 27% of the haplotypes, while the Syrian Desert and Western Asia populations contributed 16.3% and 12.9%, respectively (**Fig. 27a**). Domesticated barley populations also varied in the degree to which they incorporated genetic material from specific wild populations. For instance, haplotypes from Central Asian wild

barleys were more prevalent in domesticated barleys from East and Central Asia compared to those in other regions (**Fig. 27c**). This strong genetic affinity between wild and domesticated barleys from the East had been previously noted by Morrell et al. (Morrell and Clegg 2007), who interpreted it as evidence for a secondary center of domestication east of the Zagros Mountains in Iran. However, our findings suggest that this pattern likely resulted from gene flow between local wild populations and already domesticated barley populations originating from the Western Fertile Crescent. Conversely, the Northern Levantine wild barley population contributed more extensively to domesticated barley in Western Asia and Europe, while Central Asian ancestry dominated in East and Central Asian domesticated barleys. Additionally, Mediterranean barleys showed a higher proportion of Southern Levantine ancestry. These patterns suggest that early farmers may have taken different routes out of the Fertile Crescent, influencing the regional genetic makeup of domesticated barley. While our findings align qualitatively with those of Poets et al. (2015), who also observed regional differences in wild barley contributions, our analysis—leveraging whole-genome data rather than 5,000 SNP markers—assigns a more balanced contribution of wild populations. In contrast to Poets et al. (2015), who estimated that Southern Levantine wild barley contributed more than 50% to all domesticated populations, our results highlight a broader and more complex distribution of wild barley contributions.

Domesticated barley populations also differ in the extent of recent gene flow they have received from wild barley populations (**Fig. 27b, d**). Such introgressions are most frequent in cultivated barleys from regions where wild barley is abundant, such as Western and Central Asia. For instance, 12.8% of haplotypes in Syrian barley accessions (SYR-THM) can be attributed to recent wild barley introgressions (dating to within the last 8,000 years), primarily originating from Central Asian and Northern Mesopotamian wild populations. Interestingly, we identified evidence of recent wild haplotypes in Northern European barleys, a

region far removed from the natural range of wild barley (**Fig. 28**). Specifically, the cultivar ‘Kiruna’ (HOR 17134) shares a haplotype on chromosome 7H (between 100–200 Mb) with a Central Asian wild barley. This unexpected observation likely reflects the deliberate use of wild barley as a genetic resource by breeders. For example, ‘Kiruna’ has ‘Vogelsanger Gold’ in its pedigree, a variety developed in the 1960s that contains an introgression from wild barley (Dreiseitl and Nesvadba 2021). A similar case was observed in HOR 17572, which is cataloged as an Austrian landrace but contains the same wild-derived haplotype. We hypothesize that this may be due to errors in the passport records or accidental outcrossing during ex situ conservation and management. Such occurrences highlight the importance of careful curation and documentation in maintaining barley germplasm collections.

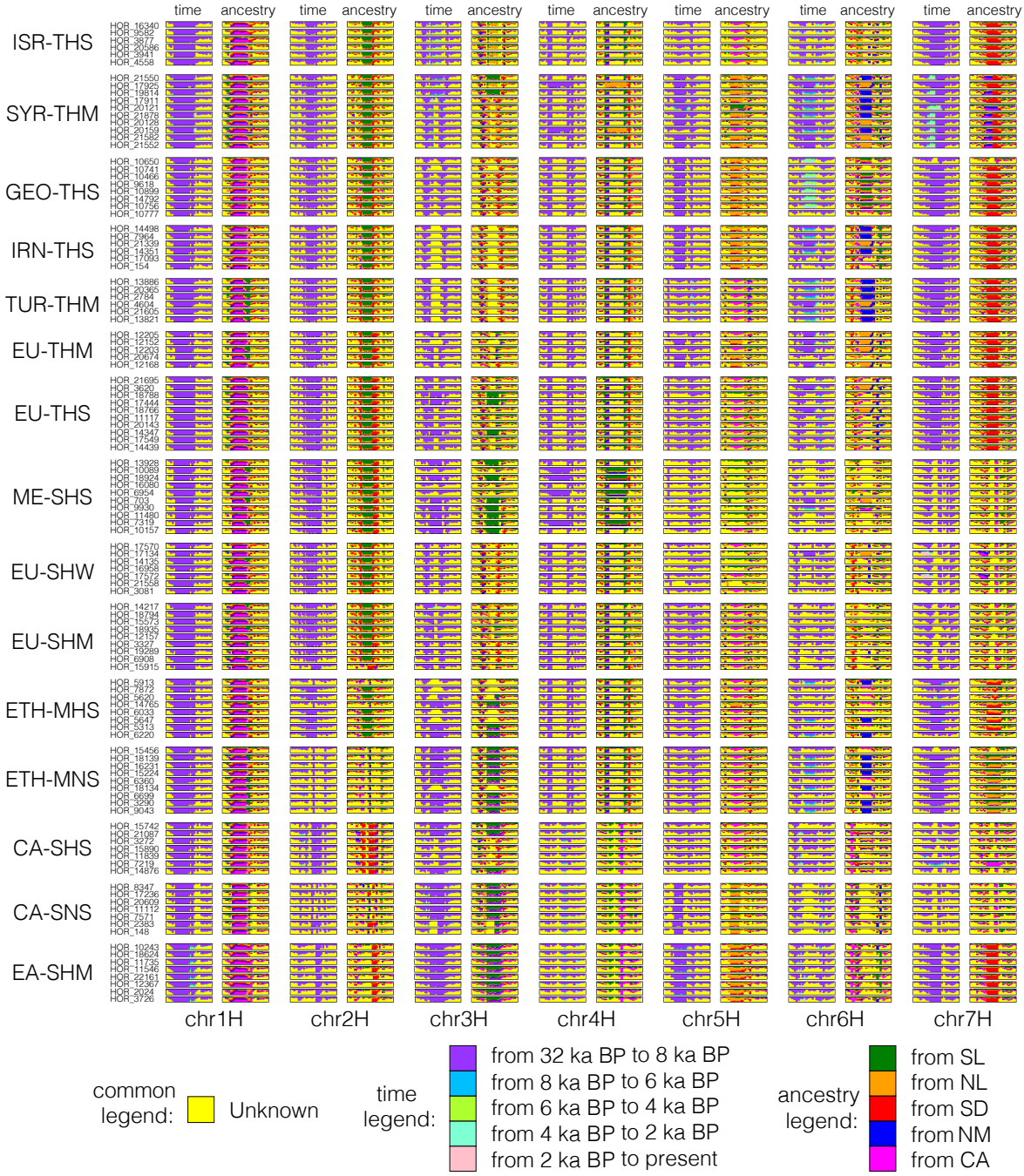


Figure 26: Spatiotemporal origins of haplotypes in domesticated barley. The inferred times at which haplotypes entered the domesticated gene pool and the most likely wild source populations are shown along the genome (20 Mb windows) for 116 domesticated barleys from 15 populations. Colors correspond to periods (time) and population (ancestry) as indicated in the legend. Yellow color indicates unknown origins.

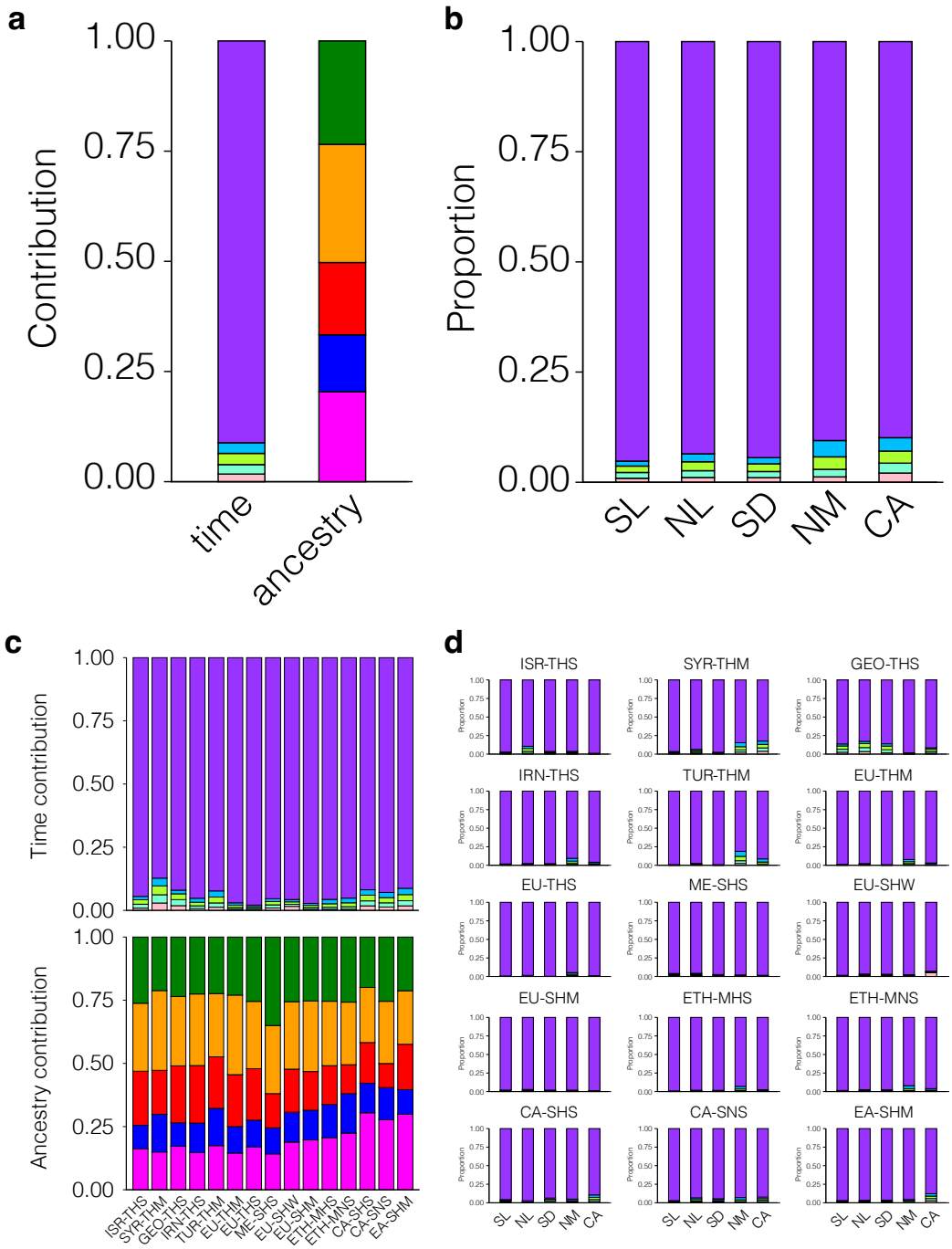


Figure 27: Statistics of spatial and temporal origins of haplotypes in domesticated barley. **(a)** Spatiotemporal origins of domesticated barley haplotypes across the entire genome. **(b)** Time periods at which haplotypes from each of the five populations entered the domesticated gene pool. **(c)** Spatiotemporal origins of haplotypes in 15 domesticated barley populations. **(d)** Time periods at which haplotypes from each of the five populations entered 15 domesticated barley populations. Haplotypes of unknown provenance were ignored when considering proportions in panels **(b)** to **(d)**.

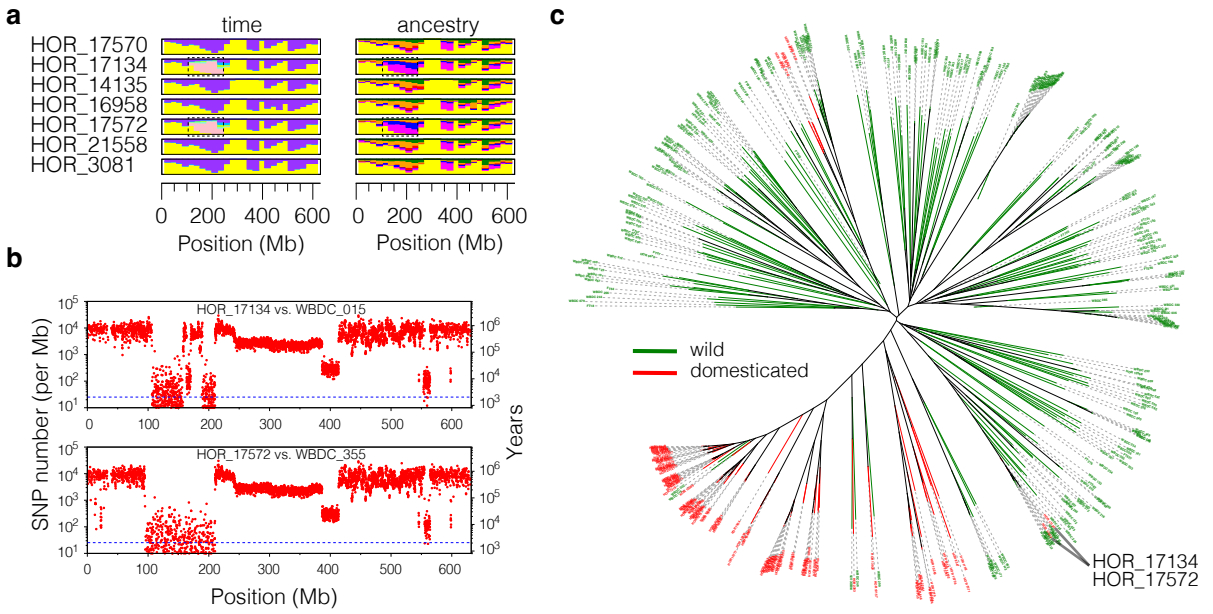


Figure 28: Recent gene flow from Central Asian wild barley to European six-rowed winter barley. (a) Spatiotemporal origins of haplotypes of the EU-SHW (European six-rowed winter barley) population on chromosome 7H. The dashed rectangle marks a haplotype that owes its presence in domesticated barley to recent gene flow. **(b)** Sequence divergence (SNPs per Mb) on chromosome 7H between two EU-SHW accessions (HOR 17134 and HOR 17572) and two Central Asian wild barleys (WBDC 055 and WBDC 355). The dashed line marks 2 ka BP of divergence (random mutation rate: 6.13 SNPs per Gb per generation). **(c)** Neighbor-joining tree computed from 1.85 M bi-allelic SNPs in the interval 120 Mb to 160 Mb on chromosome 7H. Two accessions, HOR 17134 and HOR 17572, cluster with wild barleys.

3.4 Relationships between domesticated lineages

To investigate the divergence between domesticated barley populations, we analyzed post-domestication haplotype divergence to infer split times among these groups hierarchically (**Fig. 7**). Pairwise divergence times were calculated using only SNPs within haplotypes derived from the same wild barley lineage (**Fig. 29a**). Initially, domesticated barley was grouped into three primary populations: Western barley (Western Asia + Europe), Eastern barley, and Ethiopian barley. These groups diverged around 8.5 ka BP, coinciding with the spread of agriculture from the Fertile Crescent during this period. Western

barley subsequently split into three distinct lineages—Western Asian, two-rowed European, and six-rowed European barleys—approximately 7.5 ka BP, aligning with archaeological findings (Liu et al. 2019) that barley had reached Europe, North Africa, and Central Asia by about 7 ka BP. Further subdivision occurred between 7 and 5 ka BP, as these populations diversified geographically and phenotypically. Divergence time distributions occasionally displayed multiple peaks. For European barleys, this could be attributed to gene flow between populations, which are differentiated more by morphology and phenology than by geography. For Western Asian barleys, such as those from Georgia and Iran (GEO-THS, IRN-THS), fine-scale population differentiation is plausible. These landraces may have originated from a common ancestral population but have evolved in reproductive isolation due to the distinct and isolated environments of mountainous regions. These findings are summarized graphically in **Fig. 29b**, contextualized with known agricultural dispersal routes supported by archaeological evidence (**Fig. 29c**). This framework provides insight into the genetic and geographic divergence that accompanied the spread and adaptation of barley in early agricultural societies.

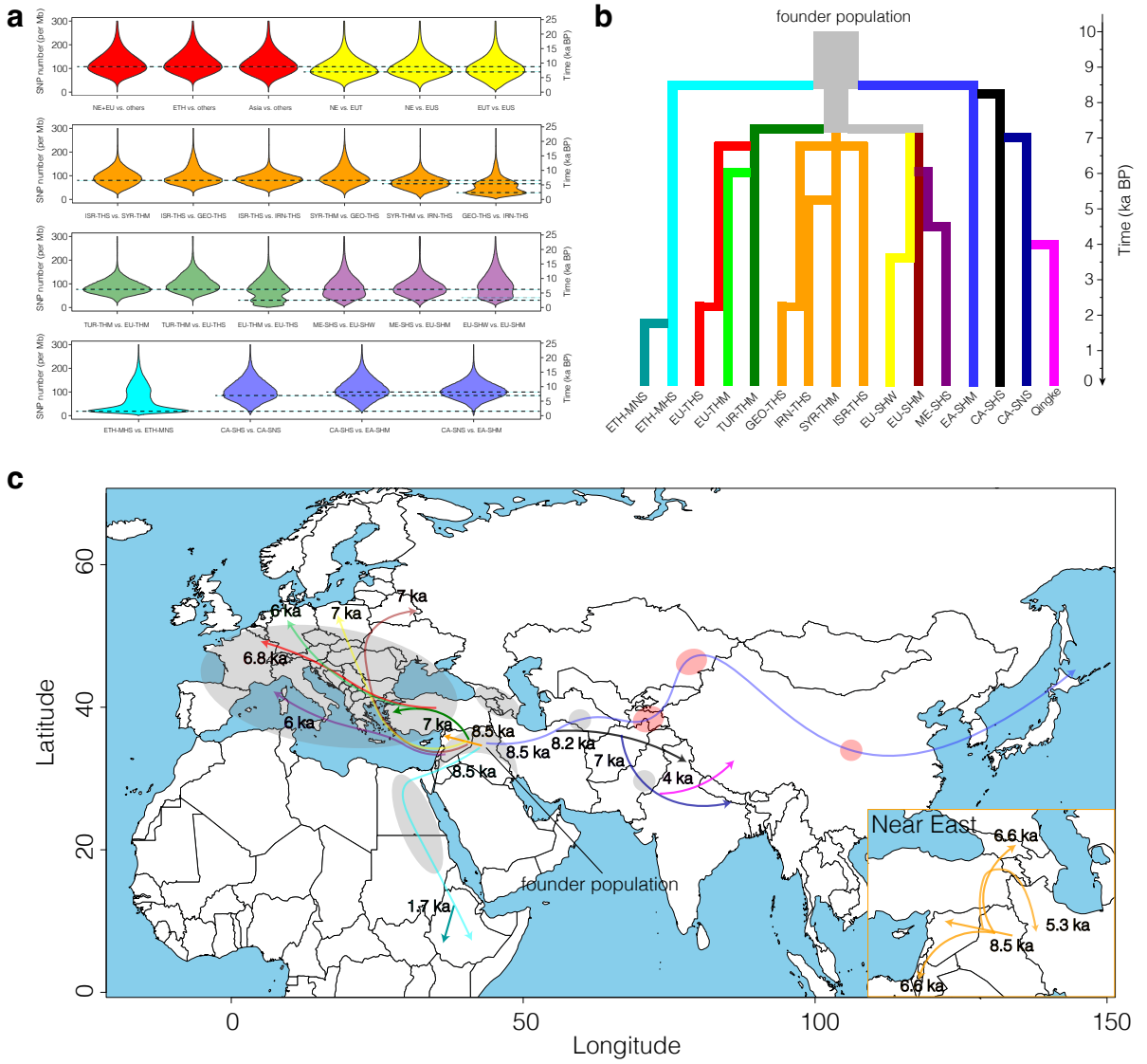


Figure 29: Divergence and dispersal of domesticated barley. (a) Violin plots showing the distribution of sequence divergence (SNPs per Mb) in pairwise comparisons between samples from different populations of domesticated barley. Dashed lines mark the peaks of the distributions (split times). Multimodal distributions may have risen from episodes of gene flow. **(b)** Schematic diagram illustrating the lineal descent and split times between 15 barley populations defined in this study and Tibetan barleys (Qingke) studied by Zeng et al. (2018). **(c)** Map showing when and along which routes domesticated barley spread from its center of origin in the Fertile Crescent. Gray shading indicates barley archaeological sites dating back about 7,000 years; red shading indicates barley archaeological sites dating back about 5,000 years (Liu et al. 2019).

3.5 A single-gene view of mosaic ancestry

Our understanding of barley crop evolution has been significantly shaped by the genetic analysis of loci where mutant alleles confer key domestication traits, such as non-shattering spikes, fertile lateral grains (six-rowed spikes), and the loss of lemma adherence (naked or hulless barley). These traits are controlled by the genes *BRITTLE RACHIS 1 and 2* (Pourkheirandish et al. 2015), *SIX ROWED SPIKE 1* (Komatsuda et al. 2007), and *NUDUM* (Taketa et al. 2008), respectively. The domesticated forms carry specific loss-of-function alleles, such as *btr1*, *btr2*, *vrs1.a1* to *vrs1.a4*, and *nud*. Interestingly, these loci did not emerge in genome-wide scans for regions with exceptionally low haplotype diversity (**Fig. 30**). This is likely because of the presence of multiple independent loss-of-function alleles at the *BTR1/2* and *VRS1* loci, as well as the limited geographic cultivation of naked barleys. Nevertheless, the persistence of long haplotypes around these genes (**Fig. 31a**), combined with the accumulation of rare variants since their most recent common ancestor (<10 ka), enabled us to estimate the origin of domesticated loss-of-function alleles. To approximate the ages of these haplotypes, we analyzed distributions of pairwise SNP numbers within these regions and translated them into divergence times (**Fig. 31a**). This approach provides insights into the evolutionary timeline of key domestication events and helps refine our understanding of how barley transitioned from wild to domesticated forms.

The estimated age of 25 ka BP for the *btr1* haplotype (**Fig. 31a**) significantly predates the earliest archaeobotanical evidence of domesticated barley by approximately 15,000 years. It is plausible that non-shattering barley and its associated haplotypes existed as rare variants in wild populations before they were selected by early cultivators. Although the precision of molecular dating is limited by uncertainties in mutation rate estimates, we propose the following relative timeline for the emergence of mutant alleles and their surrounding

haplotypes: *btr1*, *btr2*, *vrs1.a1*, *nud*, *vrs1.a3*, *vrs1.a2*, and *vrs1.a4*. The closest wild relatives of these mutant haplotypes were found in different modern wild barley populations (**Fig. 31b, c**):

- Southern Levant: *btr1*, *nud*, *vrs1.a3*
- Northern Levant: *btr2*, *vrs1.a2*
- Northern Mesopotamia and Central Asia: *vrs1.a1*, *vrs1.a4*

These findings align with earlier studies of the *btr1/btr2* locus by Pourkheirandish et al. (2015), who proposed two origins of tough-rachis barley: one in the Northern Levant and another in the Southern Levant. The early emergence of the *nud* mutation fits with the observation that hulless barley from geographically distant regions such as Tibet and Ethiopia share the same 17-kb deletion spanning the *NUD* gene **(Fig. 31a**). Despite this shared mutation, their overall genomic compositions differ markedly: Ethiopian (ETH-MNS) and Central Asian (CA-SNS) populations share no haplotypes across 44.8% of the genome. We speculate that the ancestors of Central Asian and Ethiopian barleys acquired the *nud* allele in the Fertile Crescent as it spread from a single source in the Southern Levant, integrating into barley's early gene pool before these lineages dispersed to their respective regions.

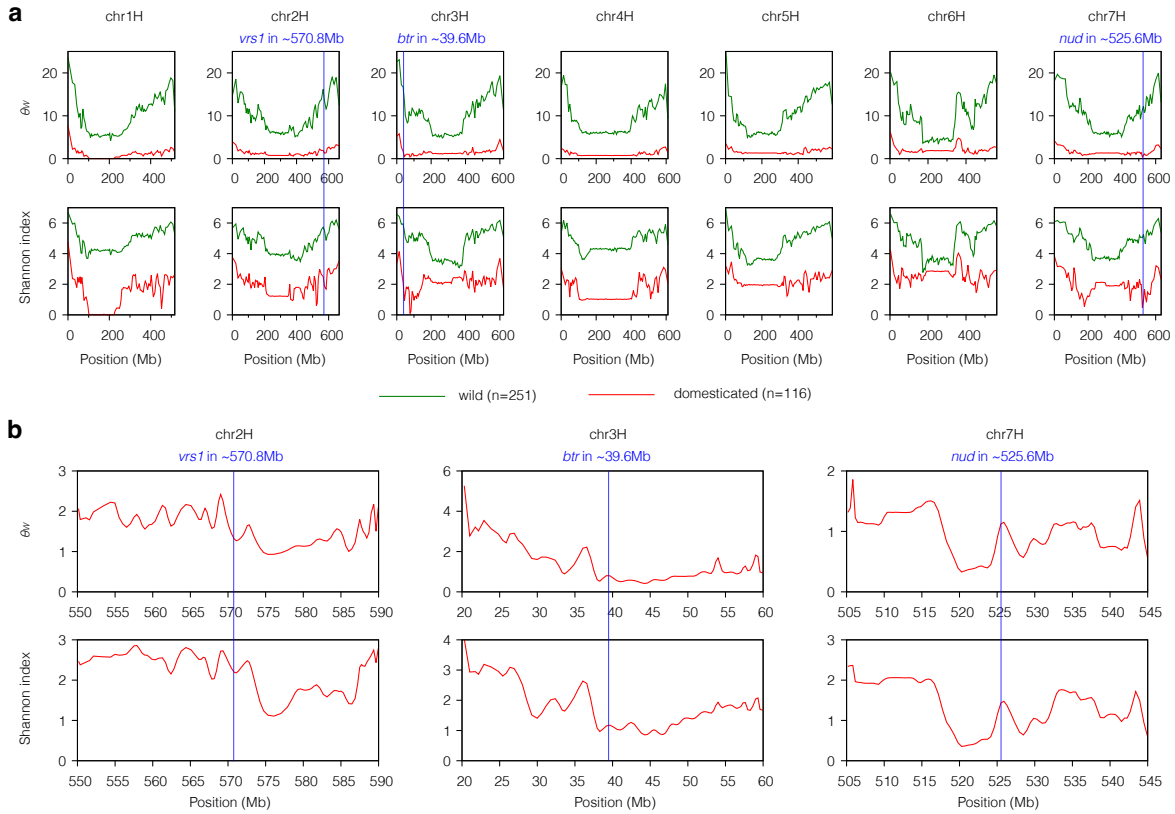


Figure 30: Haplotype-based diversity statistics. (a) Watterson's θ and the Shannon index along the seven chromosomes of barley in wild and domesticated forms. Blue lines mark the location of *Btr1/2*, *Vrs1* and *Nud* loci. **(b)** Haplotype diversity (θ_W and Shannon index) in domesticated barley around these loci.

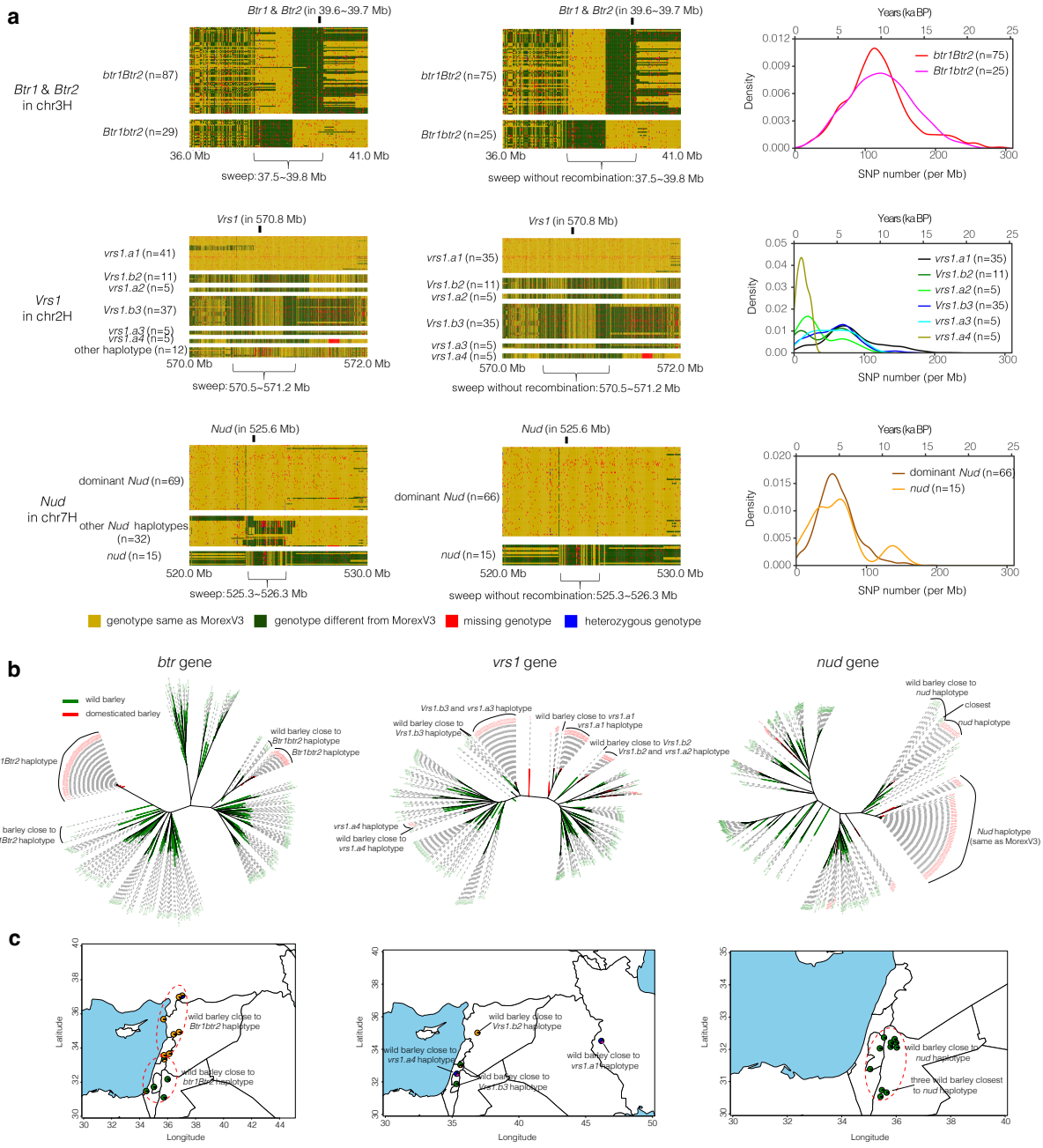


Figure 31: Origins of haplotypes of domestication genes. (a) The left-hand panels show the SNP haplotypes at the *Btr1/2*, *Vrs1* and *Nud* loci. The middle panels show the SNP haplotypes only of samples without recombination events that most likely occurred post-domestication. The right-hand panel shows the distribution of sequence divergence (SNPs per Mb) at these three loci. Only samples without recombination events were used. The window size for computing SNP numbers was 100 kb (shift: 20k kb). **(b)** Neighbor-joining tree of wild and domesticated barley computed from SNPs in the "sweep" interval denoted in panel **(a)**. **(c)** Collection sites of those wild barley samples whose haplotypes at the three loci are most closely related to domesticated haplotypes.

3.6 Persistent population structure revealed by ancient DNA

We examined ancient DNA sequences from 23 barley grains (**Table S3**), dated between 6000 and 2000 calibrated years before present (cal BP), to complement our haplotype map of extant genomes. Authenticity was confirmed by the short fragment lengths, nucleotide misincorporation profiles, and high mapping rates of ancient DNA reads to the barley reference genome (**Fig. 32**). In a principal component analysis (PCA), all ancient barleys clustered with cultivated types (**Fig. 33a**). They also carried the domesticated *btr1Btr2* haplotype commonly found in Western barleys (**Table S3**). The Yoram Cave and Timna34 samples were two-rowed barleys possessing the *Vrs1.b2* allele, also prevalent in Western types, while samples from Abi'or Cave carried the six-rowed (*vrs1.a1*) allele (**Table S3)**. To explore relationships between ancient and modern barley populations, we employed identity-by-state (IBS) analysis with genotyping-by-sequencing (GBS) data (**Fig. 33b**), ADMIXTURE (**Fig. 33c**), and D-statistics (**Fig. 33d**). Ancient two-rowed barleys were most closely related to present-day Western Asian populations, whereas six-rowed barleys showed genetic affinities with Mediterranean barleys but also carried Western Asian and European ancestry components. The six-rowed barleys from Abi'or Cave, dated to 2000 cal BP (Roman period), suggest that secondary contact between geographically distant barley populations may have been facilitated by Mediterranean sea-borne trade. Although these ancient samples represent a narrow geographic range and do not provide insights into Eastern or Ethiopian barley, they demonstrate that two- and six-rowed Western barleys were already genetically distinct thousands of years ago. This ancient diversity can be meaningfully contextualized within modern barley diversity. All ancient barley samples, except TU1120, carried the same pericentromeric haplotype on chromosome 1H as extant domesticated barley (**Fig. 33e**), which is likely of Central Asian origin (**Fig. 26**). This haplotype may have been introgressed into Southern Levantine barley through gene flow and recombination within a hypothetical founder population. These findings support

earlier observations by Mascher et al. 2016 that 6000-year-old barley more closely resembled modern domesticated barley than local wild populations.

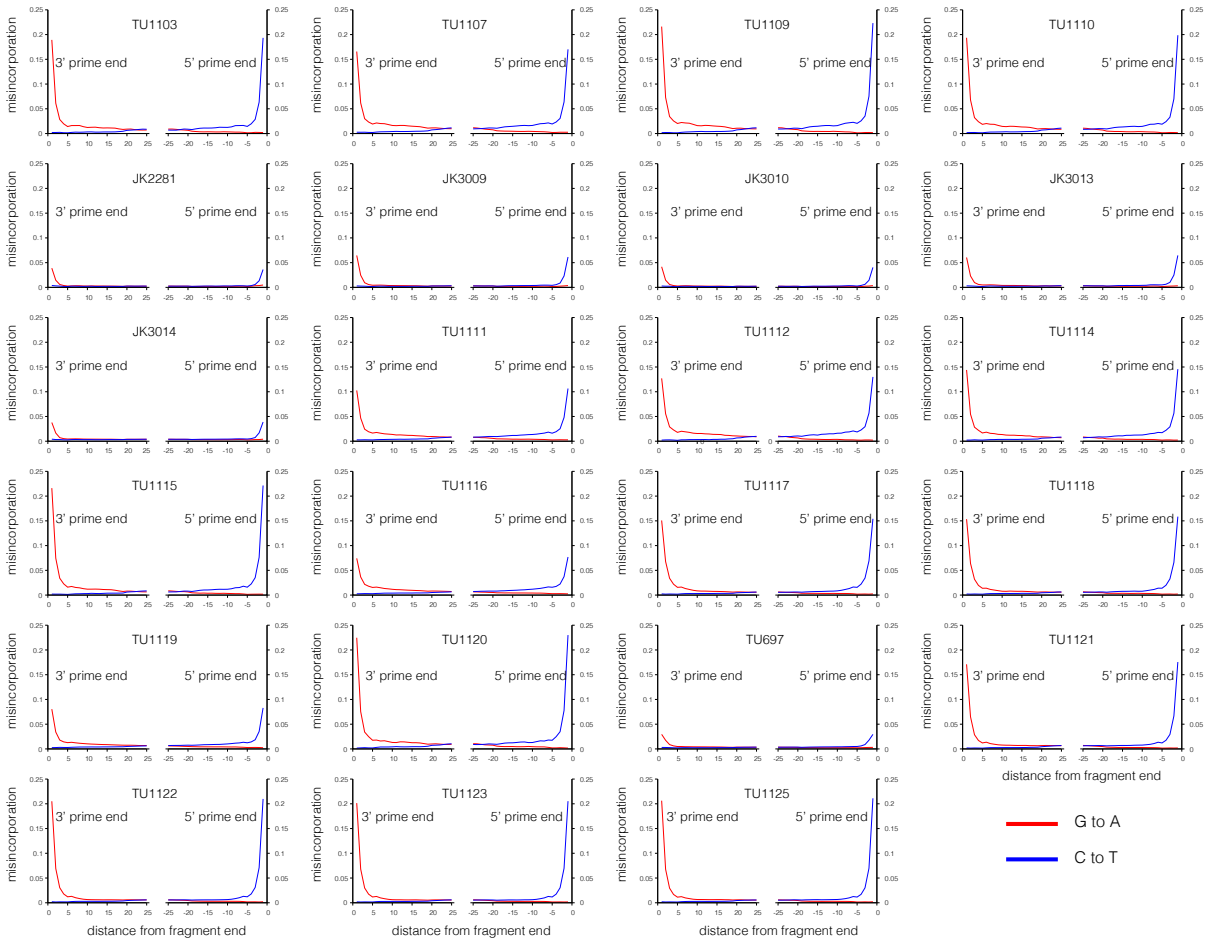


Figure 32: Nucleotide misincorporation profiles in the sequence data 23 ancient DNA samples.

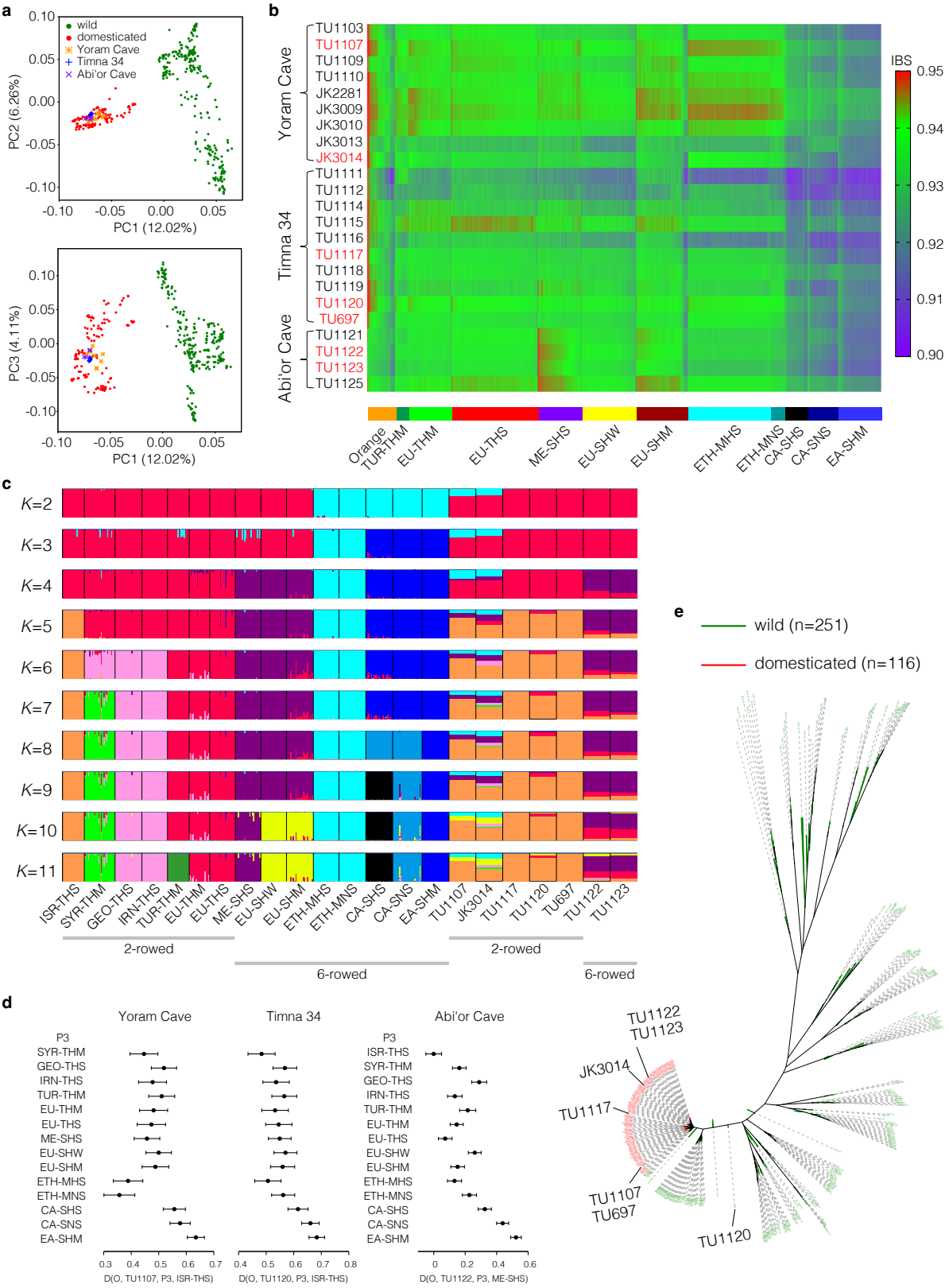


Figure 33: Ancient samples in the diversity space of extant barleys. (a) PCA on high-coverage wild and domesticated samples onto which 23 ancient barleys were projected. **(b)** Heatmap showing the identity-by-state (IBS) similarity between 23 ancient samples and 19,778

domesticated samples genotyped by Milner et al. (2019). Red font marks ancient samples sequenced at high coverage. Ancient samples with high similar to the "Orange" and ME-SHS population. The "Orange" population included Western Asian samples. **(c)** Ancestry coefficients as determined by ADMIXTURE with the number ancestral population (K) ranging from 2 to 11. The SNP matrix for the ADMIXTURE runs included 302 domesticated samples and high-coverage ancient samples. The widths of the bars corresponding to individual ancient samples have been increased by a factor of 20. (**d)** D statistics for different comparisons among ancient barleys and 15 domesticated barley populations. The outgroup (O) was *H. pubiflorum*. For samples from Yoram cave and Timna, P_4 was ISR-THS population; for that from Abi'or cave it was ME-SHS. Positive values indicate that the ancient samples shared more derived alleles with P_4 than with P_3. (**e)** Neighbor-joining tree of 251 wild barleys, 116 domesticated barley and 7 high coverage ancient barleys. The tree was computed from 1.75 M biallelic SNPs in the interval 150 to 200 Mb on chromosome 1H. No MAF filter was applied.

3.7 Six-rowed wild-growing barleys are hybrids of diverse origins

We selected 77 accessions (**Table S4**) from the German Federal ex situ genebank, housed at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, for genome-wide genotyping. These accessions included 17 classified as *H. agriocrithon var. agriocrithon*, 8 as *H. agriocrithon var. dawoense*, 11 as *H. agriocrithon var. paradoxon*, and 41 as *Hordeum* × *lagunculiforme*. Some of these accessions trace back to the seed stocks used by Åberg and Schiemann when they developed their theories on the origin of domesticated barley (Åberg 1938; Schiemann 1951). A single plant from each accession was genotyped using the genotyping-by-sequencing (GBS) method. A principal component analysis (PCA) of these 77 accessions, combined with a dataset of 389 wild and 400 domesticated barleys from diverse global sources (Milner et al. 2019; Sallam et al. 2017), revealed several patterns (**Fig. 34**). Except for *Hordeum* × *lagunculiforme*, the putative hybrid forms generally occupied intermediate positions between wild and domesticated types or clustered with domesticated barleys. The unusual placement of *Hordeum* × *lagunculiforme* in the PCA was not due to genetic divergence but resulted from a sampling artifact

(Elhaik 2021). When fewer *lagunculiforme* accessions were included, their PCA position shifted to an intermediate location between *H. spontaneum* and domesticated barley from Eastern regions (**Fig. 35**). This was further supported by identity-by-state (IBS) analysis, which showed all members of this group to be genetically near-identical (**Fig. 36**). For the other accessions, no clear clusters corresponding to their taxonomic designations (*var. agriocrithon*, *var. dawoense*, or *var. paradoxon*) were observed (**Fig. 34b**). This finding reinforces earlier criticisms by Harlan and de Wet (1971) regarding the limitations of formal infraspecific taxonomy. However, a haplotype-based classification provided clearer insights. Among the 77 accessions, 26 had been previously analyzed by Pourkheirandish et al. (2018), who categorized them as *eu-agriocrithon* or *pseudo-agriocrithon* types based on their *BTR1/2* and *VRS1* haplotypes (**Fig. 34a**). The PCA results aligned with this classification. Fifteen *pseudo-agriocrithons* clustered with domesticated barley, supporting the idea that they represent brittle rachis revertants resulting from recombination between domesticated barley lines. In contrast, all but one *eu-agriocrithon* (HOR 3900) were positioned between wild and domesticated forms in the PCA (**Fig. 34a**), consistent with their intermediate genetic profile.

Pourkheirandish et al. (2018) hypothesized that *eu-agriocrithons* are linked to six-rowed wild progenitors, from which current six-rowed domesticated barley forms are derived. The intermediate position of *eu-agriocrithons* in genetic analyses may indicate either a close relationship to certain domesticates or a more recent hybrid origin. To investigate this further, we selected 17 accessions (**Table S5**) for whole-genome shotgun sequencing (WGS), including all *eu-agriocrithons* identified by Pourkheirandish et al. (2018) and representative *lagunculiforme* accessions (**Fig. 34d**). These sequences were analyzed alongside WGS data from 99 wild barleys (Jayakodi et al. 2020), 195 domesticated barleys (Jayakodi et al. 2020) and three wild-growing Tibetan barleys classified as *pseudo-agriocrithons* (Zeng et al. 2018). A neighbor-joining (NJ) tree (**Fig. 37a**) and a principal

component analysis (PCA) (**Fig. 37b**) based on WGS variants showed patterns similar to those derived from genotyping-by-sequencing (GBS) data. Most accessions occupied positions between wild and domesticated clusters (**Fig. 37**). *Pseudo-agriocrithons* identified by Zeng et al. (2018) appeared closely related to domesticated barley from East and Central Asia, as did the purported wild barley HOR 3900. A model-based ancestry analysis using ADMIXTURE provided further insights (**Fig. 37**). When four ancestral populations (K=4) were modeled, wild barley separated into two groups: Western and Central Asian (WWA and WCA) accessions (**Fig. 38**). Domesticated barley also split geographically into Western (WD) and Eastern (ED) ancestral populations. These broad population structures align with prior population genomic studies (Morrell and Clegg 2007; Russell et al. 2016; Milner et al. 2019). Most putative crop-wild hybrids displayed mixed ancestry from wild and domesticated populations, though a few accessions, including three *pseudo-agriocrithons* and HOR 3900, had exclusively Eastern domesticated ancestry. Other hybrids showed variable proportions of wild and domesticated ancestry. Interestingly, the geographic locations where hybrids were collected corresponded to their respective ancestral populations (**Fig. 37**). Hybrids from Tibet displayed ED and WCA ancestry, while those from Israel combined WD and WWA ancestry. Hybrids from the Caucasus and Central Asia exhibited mixed Eastern and Western ancestries. For instance, HOR 4904 from Azerbaijan contained nearly equal proportions of all four components. Consistent with the PCA, taxonomic classifications did not correlate with genomic ancestry.

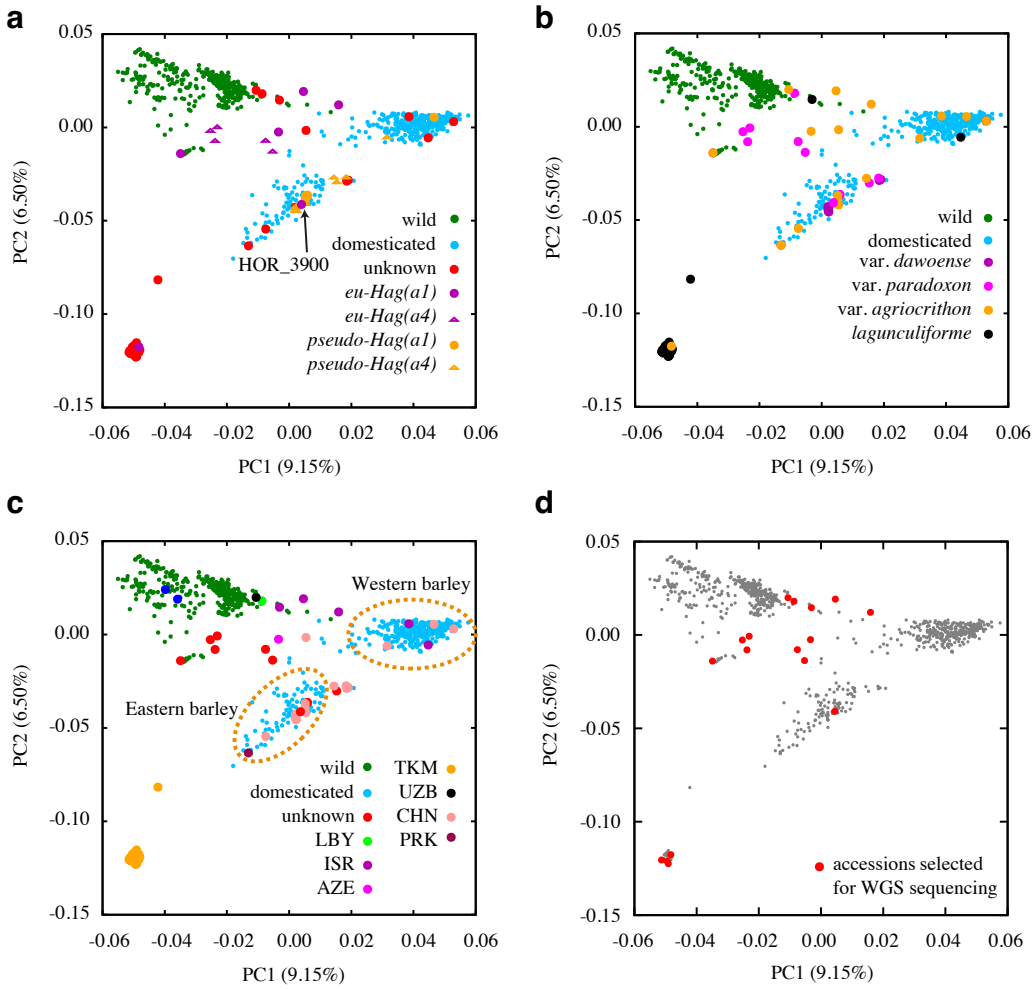


Figure 34: Positions of *H. agriocrithon* accessions in the global barley diversity space. Principal component analysis (PCA) of *H. agriocrithon* and global diversity panels of domesticated and wild (*H. spontaneum*) barley. The PCA was computed from genome-wide marker data obtained by genotyping-by-sequencing. The proportion of variance explained by the PCs is given in the axis labels. All four panels show the same PCA colored according to different factors. **(a)** *H. agriocrithon* were according to the classification into *pseudo-eu-agriocrithon* of Pourkheirandish et al. (2018) based on haplotypes at domestication genes. (**b)** *H. agriocrithon* accessions were colored according to infraspecific taxonomy. **(c)** *H. agriocrithon* accessions were colored according to their countries of origin. (**d)** *Eu-agriocrithon* selected for WGS sequencing.

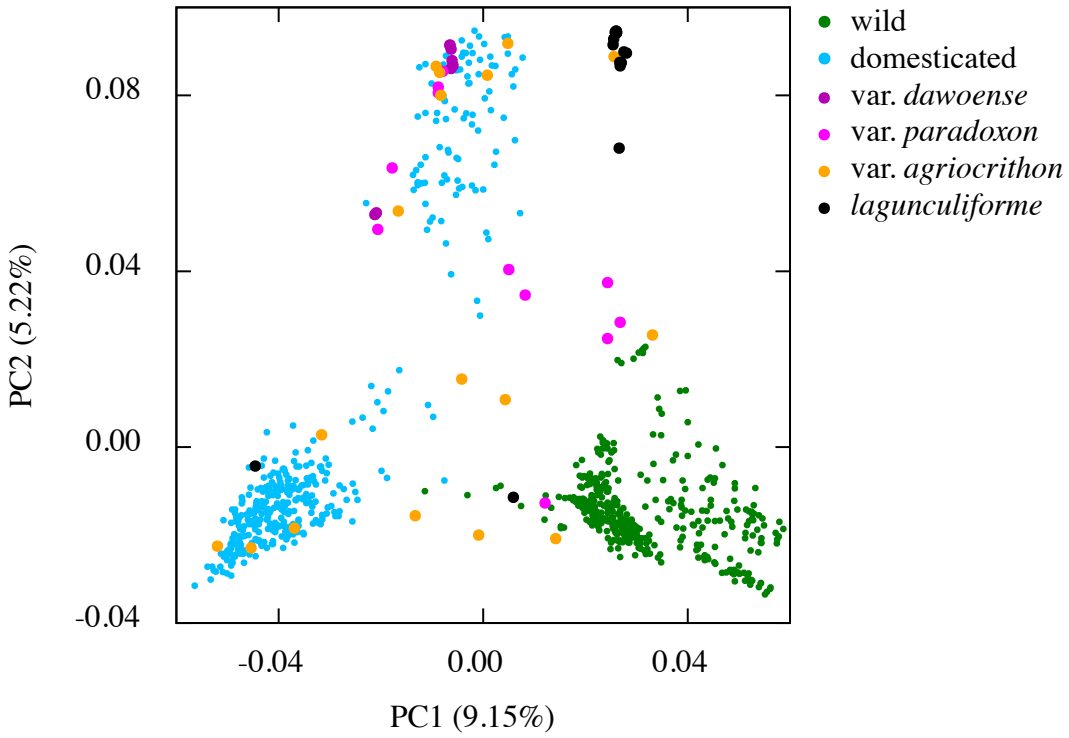


Figure 35. PCA analyses based on GBS SNPs with only 19 accessions of *lagunculiforme*. After reducing the accessions numbers of *lagunculiforme* from 41 (**Fig. 34**) to 19, *lagunculiforme* was placed at an intermediate position between *H. spontaneum* and Eastern domesticated barley.

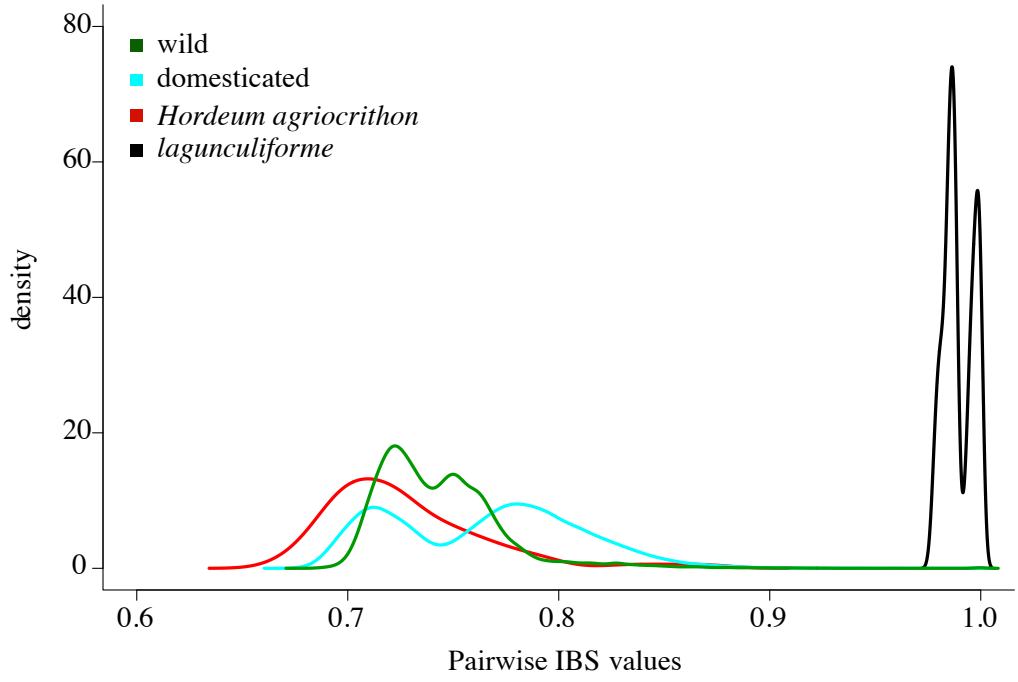


Figure 36. Distribution of pairwise IBS values based on GBS SNPs in wild barley, domesticated barley, *H. agriocrithon* and *lagunculiforme* accessions.

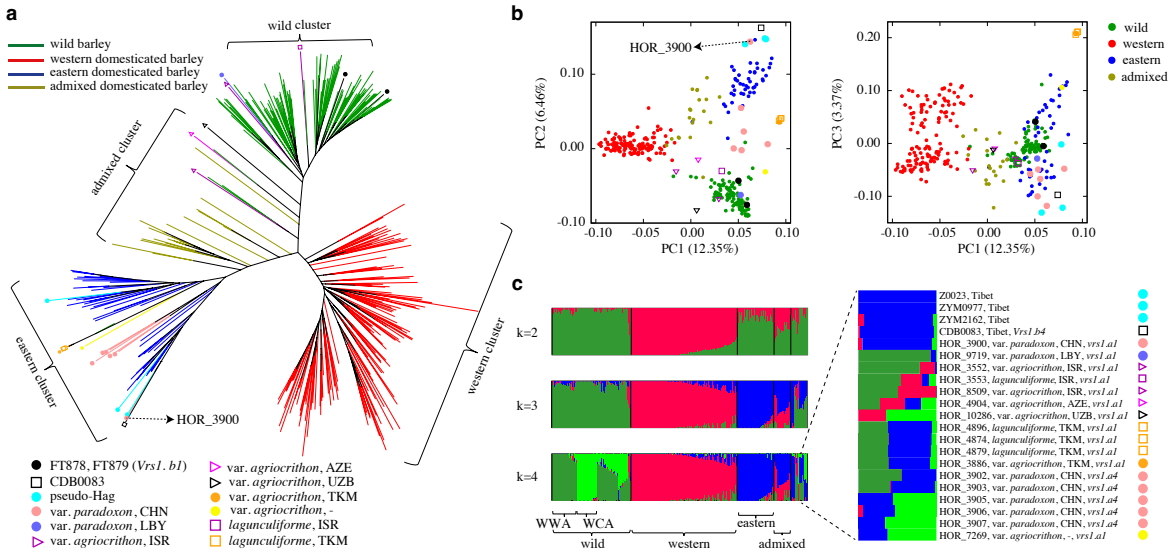


Figure 37: Genomic ancestry of *H. agriocrithon*. (a) Neighbor-joining tree based on genetic distance computed from the a whole-genome variant matrix. (**b)** Principal component analyses (PCA) showing PC1 vs. PC2 (left) and PC1 vs. PC3 right. The proportion of variance explained by the PCs is given in the axis labels. **(c)** Individual ancestry coefficients in ADMIXTURE with the the number of ancestral popultions (*K*) ranging from 2 to 4. WWA: wild barley subpopulation mainly distributed in western Asia. WCA: wild barley subpopulation mainly distributed in Central Asia. Color codes and labels for *H. agriocrithon* taxa are given **(a)** and also used in **(b)** and **(c)**.

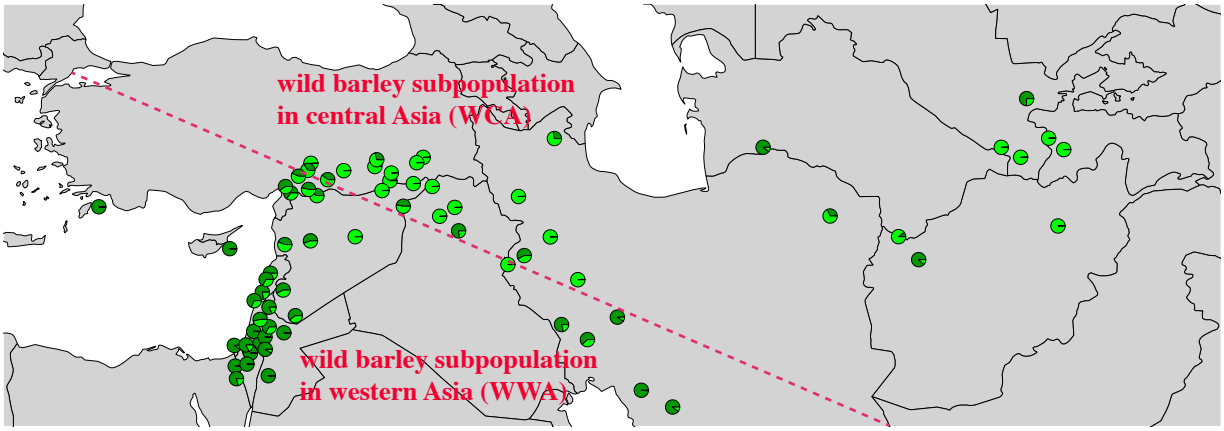


Figure 38. Population structure of wild barley in Western and Central Asia. Pie charts of each accession show the ancestry proportions in an ADMIXTURE analysis with *K*=4. The collection sites of wild barley accessions were provided by Russell et al. (2016). The picture was plotted with R package 'marmap'.

4. Discussion

4.1 Domestication history of barley

Pankin et al. (2017) proposed two potential models to explain the mosaic ancestry of domesticated barley: (i) repeated introgressions from various wild barley populations into an early domesticated "*proto-vulgare*" lineage or (ii) the influence of population structure within the wild progenitors from which the domesticated gene pool was derived. These models are not mutually exclusive, and our analyses suggest that both processes likely contributed to barley's domestication history. The geographically diverse origins of haplotypes associated with key domestication genes, such as *btr1/2*, *vrs1*, and *nud*, align with two scenarios: (i) an initial polycentric phase of domestication and (ii) a prolonged period during which barley cultivation remained localized within the Fertile Crescent. During this period, local wild barley populations and early farming communities co-evolved across the region (**Fig. 39a**). The early estimated age of the *btr1* non-shattering haplotype aligns with evidence of domestic-type abscission scars observed in wild barley rachises from the Ohalo II archaeological site in the Southern Levant, dated to 23,000 years before present (Snir et al. 2015). This suggests that barley cultivation may have begun before the fixation of hallmark domestication traits, such as non-brittle spikes (Snir et al. 2015; Zhang et al. 2024). Such a scenario supports the notion of an extended "proto-domestication" phase. Additionally, the long duration of this proto-domestication phase is reflected in demographic reconstructions, which show a sustained decline in effective population sizes between 25,000 and 10,000 years before present (**Fig. 39b**). This wide trough suggests a gradual transition from wild harvesting to active cultivation, potentially driven by the shifting practices of early human populations.

Early cultivated barley populations were interconnected enough through gene flow to enable the dissemination of beneficial genes with significant effects.

Simultaneously, hybridization between domesticated barley and local wild populations contributed to enriching the genetic diversity of cultivated forms. This interplay is reflected in current patterns of haplotype sharing. In Western Asia, this exchange of genetic material has persisted beyond the initial domestication phase and continues to shape the region's barley populations. Evidence of post-domestication haplotype exchanges indicates that present-day Western Asian barleys are unlikely to be direct linear descendants of the original founder populations. Instead, they represent a dynamic and continuously evolving genetic mosaic.

As early agriculturalists expanded beyond the Fertile Crescent, domesticated barley diverged into distinct lineages that largely evolved in isolation. This geographic separation marked the end of significant wild haplotype introgression into these lineages. However, as demonstrated by the current population structure, geography was not the sole driver of population differentiation. Farmers' cultivation practices also played a crucial role, as seen in the development of distinct gene pools, such as those underlying European two-rowed and six-rowed barleys. The mosaic ancestry arising from these processes presents challenges for identifying adaptive loci. As *proto-vulgare* haplotypes were sorted into local lineages, adaptation to new environments likely influenced regional patterns of genetic diversity. However, disentangling the effects of lineage sorting from local adaptation is complex. In many genomic regions, haplotypes are not shared between lineages, a phenomenon also seen in selective sweeps. A promising approach to differentiate ancestral haplotype structure from adaptive evolution is mutational genomics. For example, the flowering time regulator *HvCENTRORADIALIS* was initially identified as a major mutation in a classical barley mutant. Subsequent sequence analyses revealed its role in geographic range expansion (Comadran et al. 2012) and linked it to newly arisen structural variations (Jayakodi et al. 2020), offering insights into the interplay between mutation and adaptation.

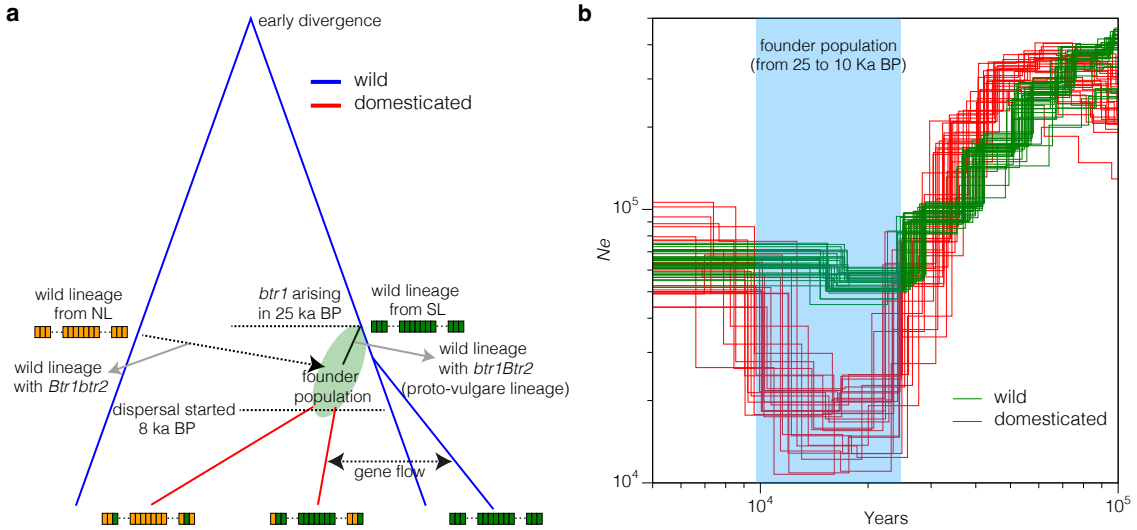


Figure 39: Origin and evolution of domesticated barley. (a) Model explaining the origin of the bimodal distribution of sequence divergence in domesticated barley. Two wild populations (SL and NL) are shown for illustration. In fact, all five wild barley populations contributed to the formation of the founder population. Blue and red lines stand for domesticated and wild lineages, respectively. The rectangles represent a chromosome subdivided into windows. Distal and proximal regions are separated by dashed lines. Green and orange genomic windows in extant barleys (bottom) are derived from progenitor haplotypes in the SL and NL populations, respectively. About 25 ka BP, the *btr1* mutation conferring a tough rachis arose in the SL population. The *btr2* mutation arose ca. 22 ka BP in the NL population. Later, non-brittle barley was taken into cultivation independently by early farmers in the respective regions. Gene flow facilitated by human migration or cultural exchange gave rise to an admixed founder population. As farmers and their crops moved out of the Fertile Crescent, the founder population split into several isolated lineages that retained genomic windows of diverse wild ancestry. Later gene flow from wild to domesticated occurs in regions of sympatry. **(b)** Trajectories of historic effective population size inferred by PSMC using wild barley (green) and domesticated barleys (red).

4.2 Origin of six-rowed wild barley

Our data challenge the hypothesis proposed by Pourkheirandish et al. (2018) that genebank accessions classified as *H. agriocrithon* represent descendants of wild barley populations from which six-rowed domesticated barleys emerged during the early Neolithic. Instead, our findings suggest that these accessions are crop-

wild hybrids of more recent origin and diverse geographic backgrounds. This conclusion aligns with previous studies employing less dense molecular markers and comparisons of genetic diversity between *H. spontaneum* and *H. agriocrithon* (Murphy et al. 1982; Tanno and Takeda 2004). While our dataset corroborates the observation of Pourkheirandish et al. (2018) regarding the coexistence of wild haplotypes at the *BTR1/2* locus and domesticated alleles at *VRS1* in *eu-agriocrithon* barleys, our genome-wide ancestry analysis provides further context. Beyond these loci, the presence of *vrs1* haplotypes in wild-growing barleys reflects broader domesticated ancestry. This pattern supports a history of hybridization between domesticated and wild barley populations that had already undergone genetic divergence.

We emphasize that our findings indicate that the extant genebank accessions are not direct descendants of the ancestral populations. However, this does not negate the broader hypothesis proposed by Pourkheirandish et al. (2018) that six-rowed barley could be as ancient as, or even predate, non-shattering barley. Many archaeologists support the notion of a protracted domestication process lasting one or two millennia, during which *btr1/2* and *vrs1* alleles may have segregated within early cultivated barleys (Purugganan et al. 2019). While our sequence data do not provide insights into allele frequencies in the distant past, archaeological evidence suggests a chronological sequence: two-rowed domesticated barleys appear earlier than six-rowed types. The earliest remains of two-rowed domesticated barley date to the first half of the 8th millennium BCE (Hopf 1983), whereas six-rowed barleys, such as those from Çatalhöyük, are not older than approximately 9,350 cal BP (Helbaek 1964).

Since Western botanists first encountered wild-growing six-rowed barley in East and Central Asia, this intriguingly named taxon (*Hordeum agriocrithon*, derived from ἄγριος – wild, and κριθή – barley) has captivated barley geneticists. Much

of the subsequent research, inspired by Åberg's work, has revolved around the question of whether *H. agriocrithon* represents an ancestral form of domesticated barley. However, genetic evidence tells a different story. Domesticated barley exhibits a mosaic ancestry derived from various wild barley populations in the Fertile Crescent (Poets et al. 2015), making it impossible to identify extant descendants of a singular source population. Consequently, *H. agriocrithon* is not an ancestral form of domesticated barley.

Nonetheless, we believe that *H. agriocrithon* still holds valuable lessons for barley geneticists. Crop-wild hybrids between *H. spontaneum* and two-rowed barleys, while less striking, may be as prevalent as *H. agriocrithon* (Helbaek 1959). Genotyping barley that grows wild near cultivated fields could reveal how frequently two-rowed crop-wild hybrids occur. Six-rowed wild-growing barleys have been observed as ruderals, weeds (Kamm 1954), or in newly colonized regions such as the Tibetan Plateau. Could their higher fertility or potential pleiotropic effects of *VRS1* on vegetative growth habit (Thirulogachandar et al. 2017) offer fitness advantages over two-rowed forms in such environments? Despite the broad mutational potential for altering row type, there are no known unadmixed six-rowed populations of *H. spontaneum*. Similarly, six-rowed brittle-rachis barleys are absent from the archaeobotanical record. This absence has been attributed to the disruption of the wedge-shaped dispersal unit caused by the presence of lateral grains (Komatsuda et al. 2007), which may lead to reduced fitness. This hypothesis could be tested in common garden experiments involving six-rowed *vrs1* mutants and their *H. spontaneum* progenitors (Pourkheirandish et al. 2015).

5. Outlook

This study systematically explored the domestication history of barley and made significant progress. However, as the research deepened, new questions have emerged that require further investigation.

Our study identified seven potential domestication regions in domesticated barley population. The most striking case is the proximal region of chromosome 1H, where a single haplotype is shared by all domesticated barley worldwide. This is almost certainly not a coincidence but rather a result of artificial selection during domestication. Within each of these regions, we selected 10 to 20 candidate genes, none of which overlap with any currently cloned barley genes. Future barley researchers interested in this area should focus on the domestication genes within these regions.

Many studies based on whole-genome resequencing attempt to identify selective sweeps by examining reductions in genetic diversity in crops compared to their progenitors. However, our findings suggest that this approach has limited utility for species like barley, which originated from multiple wild ancestry. A striking example are the *NON-BRITTLE RACHIS* genes: one domesticated haplotype, *btr1Btr2* originates from the wild barley in Southern Levant, while another haplotype, *Btr1btr2*, originates from the wild barley in Northern Levant. In a global sampling scenario, the two haplotypes are almost equally represented among the samples. As a result, the genetic diversity at this locus does not appear significantly reduced, making it difficult to detect a selective sweep. This highlights that using diversity reduction to pinpoint selective sweeps in barley may only be effective under specific conditions, such as when the analysis is limited to populations carrying a single haplotype. Another common method for identifying selective sweeps is comparing F_{ST} between different crop populations. As discussed in our study, population-specific haplotypes can arise either from

genuine selective sweeps or from the allocation of distinct ancestral haplotypes to different populations during lineage sorting. Disentangling these two possibilities is inherently challenging. We underscore that both approaches—diversity comparison and F_{ST}-based analysis—are frequently misapplied in population genetics studies, particularly in cases involving complex domestication histories like that of barley. For mosaic-ancestry crops like barley, effectively identifying selective sweeps will require future breakthroughs in genetic theories or algorithms.

6. Summary

Barley, one of the earliest domesticated crops, possesses a complex and multifaceted domestication history. Its evolutionary trajectory and the molecular mechanisms underpinning its domestication have been the focus of extensive research. These studies have refuted the hypothesis of a single origin, instead supporting a model of mosaic genomic ancestry. The increasing availability of high-resolution genome sequences now enables a more detailed examination of this model, prompting critical questions: where do the constituent haplotypes of this mosaic originate? Did all domesticated barley populations derive equal contributions from wild progenitors, or did certain wild populations disproportionately shape specific domesticated lineages? To address these questions, we employed a haplotype-based framework to elucidate patterns of genetic diversity and population structure in wild and cultivated barley. We analyzed the genomes of 628 genebank accessions alongside 23 archaeological specimens. These data enabled us to infer the spatiotemporal origins of haplotypes and characterize the contributions of distinct wild barley populations during the initial domestication phase and subsequent gene flow events. Ancient DNA evidence corroborated the genomic patterns observed in extant barley populations. Our findings suggest that an early domesticated founder population emerged in the Fertile Crescent, likely through a protracted period of pre-domestication cultivation. A key implication of this study is that the pronounced haplotype differentiation among barley populations—resulting from processes independent of, or superimposed upon, selective pressures—poses significant challenges for the identification of adaptive loci.

Gene flow between domesticated crops and their wild relatives is a well-documented phenomenon in regions where their ranges overlap. The resulting hybrid forms, often classified as semi-domesticates, are frequently interpreted as potential "missing links" between fully domesticated crops and their wild

progenitors. Wild barley populations in Central and Eastern Asia, collectively referred to as *Hordeum agriocrithon*, exhibit morphological traits intermediate between wild and domesticated forms. These include spike shattering for natural seed dispersal and lateral grain production, a characteristic absent in other wild barley populations but favored by early cultivators. Due to these intermediate traits, *H. agriocrithon* has been repeatedly proposed as a progenitor of domesticated barley. Through genome-wide marker analysis and whole-genome resequencing, we demonstrate that all *H. agriocrithon* accessions in a major germplasm collection are hybrids, originating independently through multiple admixture events between domesticated and wild barley populations. While *H. agriocrithon* does not appear to have played a unique role in barley domestication, further investigation into the adaptive dynamics of bidirectional gene flow between crops and wild populations offers a promising direction for future research.

7. Zusammenfassung

Gerste, eine der frühesten domestizierten Kulturpflanzen, besitzt eine komplexe und vielschichtige Domestikationsgeschichte. Ihre evolutionäre Entwicklung und die molekularen Mechanismen, die ihrer Domestikation zugrunde liegen, wurden intensiv erforscht. Diese Studien haben die Hypothese eines einzigen Ursprungs widerlegt und stattdessen ein Modell einer mosaikartigen genomischen Abstammung unterstützt. Mit der zunehmenden Verfügbarkeit hochauflösender Genomsequenzen ist nun eine detailliertere Untersuchung dieses Modells möglich, was zentrale Fragen aufwirft: Woher stammen die Haplotypen, die Bausteine dieses Mosaiks? Haben alle domestizierten Gerstenpopulationen gleichermaßen Beiträge von Wildvorfahren erhalten, oder haben bestimmte Wildpopulationen spezifische domestizierte Linien stärker geprägt? Um diese Fragen zu beantworten, haben wir einen haplotypbasierten Ansatz angewandt, um die Muster der genetischen Vielfalt und der Populationsstruktur in wilden und domestizierten Gersten zu entschlüsseln. Wir analysierten die Genome von 628 Genbank-Akzessionen sowie 23 archäologischen Proben. Diese Daten ermöglichten es uns, die räumlich-zeitlichen Ursprünge der Haplotypen zu rekonstruieren und die Beiträge verschiedener Wildgerstenpopulationen während der initialen Domestikationsphase und späterer Genflussereignisse zu charakterisieren. Die Analyse von DNA aus archäobotanische Proben stützen die genomischen Muster, die in heutigen Gerstenpopulationen beobachtet wurden. Unsere Ergebnisse legen nahe, dass eine frühe domestizierte Gründerpopulation im Fruchtbaren Halbmond entstand, wahrscheinlich durch eine verlängerte Phase der Vordomestikation. Eine zentrale Implikation dieser Studie ist, dass die ausgeprägte Haplotypdifferenzierung zwischen Gerstenpopulationen – resultierend aus Prozessen, die unabhängig von oder zusätzlich zu selektiven Kräften wirken – die Identifikation adaptiver Loci erheblich erschwert.

Genfluss zwischen domestizierten Kulturpflanzen und ihren wilden Verwandten ist ein gut dokumentiertes Phänomen in Regionen, in denen sich ihre Verbreitungsgebiete überschneiden. Die resultierenden Hybridformen, oft als halbdomestizierte Formen klassifiziert, werden häufig als potenzielle „fehlende Glieder“ zwischen vollständig domestizierten Kulturpflanzen und ihren wilden Vorfahren interpretiert. Wilde Gerstenpopulationen in Zentral- und Ostasien, zusammengefasst unter der Bezeichnung *Hordeum agriocrithon*, weisen morphologische Merkmale auf, die zwischen wilden und domestizierten Formen intermediär liegen. Dazu gehören das Zerbrechen der Ähren zur natürlichen Samenausbreitung sowie die Produktion seitlicher Körner, ein Merkmal, das bei anderen Wildgerstenpopulationen fehlt, aber von frühen Kultivierenden bevorzugt wurde. Aufgrund dieser intermediären Merkmale wurde *H. agriocrithon* wiederholt als potenzieller Vorfahr der domestizierten Gerste vorgeschlagen. Durch die Analyse genomweiter Marker und die vollständige Genomsequenzierung zeigen wir, dass alle *H. agriocrithon*-Akzessionen in einer bedeutenden Keimplasmasammlung Hybriden sind, die unabhängig durch mehrere Vermischungsereignisse zwischen domestizierten und wilden Gerstenpopulationen entstanden sind. Obwohl *H. agriocrithon* offenbar keine einzigartige Rolle in der Domestikation der Gerste gespielt hat, bietet die weitere Untersuchung der adaptiven Dynamiken des bidirektionalen Genflusses zwischen Kulturpflanzen und Wildpopulationen eine vielversprechende Perspektive für zukünftige Forschung.

8. Reference

Åberg E. 1938. *Hordeum agriocriihon nova sp*., a wild six-rowed barley. *Lantbrukshogskolans Annaler* 6, 159-216.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19(9), 1655–1664.

Allaby RG, Brown TA. 2003. AFLP data and the origins of domesticated crops. *Genome* 46, 448–453.

Ames N, Dreiseitl A, Steffenson BJ, Muehlbauer GJ. 2015. Mining wild barley for powdery mildew resistance. *Plant Pathol* 64, 1396–1406.

Berner D. 2019. Allele Frequency Difference AFD-An Intuitive Alternative to FST for Quantifying Genetic Population Differentiation. *Genes (Basel)* 10(4), 308.

Bhatia G, Patterson N, Sankararaman S, Price AL. 2013. Estimating and interpreting FST: the impact of rare variants. *Genome Res* 23(9), 1514-21.

Blattner FR. 2018. Taxonomy of the genus *Hordeum* and barley (*Hordeum vulgare*). *The barley genome* 11–23.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7.

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single

nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6(2), 80-92.

Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, Shaw P, Bayer M, Thomas W, Marshall D, Hedley P, Tondelli A, Pecchioni N, Francia E, Korzun V, Walther A, Waugh R. 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat Genet* 44(12), 1388-1392.

Cubry P, Tranchant-Dubreuil C, Thuillet AC, Monat C, Ndjiondjop MN, Labadie K, Cruaud C, Engelen S, Scarcelli N, Rhoné B, Burgarella C, Dupuy C, Larmande P, Wincker P, François O, Sabot F, Vigouroux Y. 2018. The Rise and Fall of African Rice Cultivation Revealed by Analysis of 246 New Genomes. *Curr Biol* 28(14), 2274-2282.e6.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10(2), giab008.

Dreiseitl A, Nesvadba Z. 2021. Powdery mildew resistance genes in single-plant progenies derived from accessions of a winter barley core collection. *Plants (Basel)* 10(10), 1988.

Ellis RP, Forster BP, Robinson D, Handley LL, Gordon DC, Russell JR, Powell W. 2000. Wild barley: a source of genes for crop improvement in the 21st century? *J Exp Bot* 51, 9–17.

Elhaik E. 2021. Why most principal component analyses (PCA) in population genetic studies are wrong. *bioRxiv*. doi: https://doi.org/10.1101/2021.04.11.439381

Ellstrand NC, Meirmans P, Rong J, Bartsch D, Ghosh A, de Jong TJ, Haccou P, Lu BR, Snow AA, Stewart CN Jr, Strasburg JL, van Tienderen PH, Vrieling K, Hooftman D. 2013. Introgression of crop alleles into wild or weedy populations. *Annu Rev Ecol Evol Syst* 44, 325-345.

Fang Z, Gonzales AM, Clegg MT, Smith KP, Muehlbauer GJ, Steffenson BJ, Morrell PL. 2014. Two genomic regions contribute disproportionately to geographic differentiation in wild barley. *G3 (Bethesda)* 4, 1193–1203.

Fuller DQ, Denham T, Allaby R. 2023. Plant domestication and agricultural ecologies. *Curr Biol* 33, R636–R649.

Gutaker RM, Burbano HA. 2017. Reinforcing plant evolutionary genomics using ancient DNA. *Curr Opin Plant Biol* 36, 38–45.

Harlan JR, de Wet JM. 1971. Toward a rational classification of cultivated plants. *Taxon* 20, 509-517.

Hasselbach-Andee R (Ed.). 2020. A companion to ancient Near Eastern languages. *John Wiley & Sons*.

Harlan JR, Zohary D. 1966. Distribution of wild wheats and barley: the present distribution of wild forms may provide clues to the regions of early cereal domestication. *Science* 153, 1074–1080.

Helbaek H. 1959. Domestication of food plants in the Old World: Joint efforts by botanists and archeologists illuminate the obscure history of plant domestication. *Science* 130(3372), 365-372.

Helbaek H. 1964. First impressions of the Çatal Hüyük plant husbandry. *Anatolian Studies* 14, 121-123.

Heun M, Schafer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F. 1997. Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278, 1312–1314.

Hopf M. 1983. The plants found at Jericho. *Excavations at Jericho VL: British School of Archaeology in Jerusalem* 580-621.

Hübner S, Höffken M, Oren E, Haseneyer G, Stein N, Graner A, Schmid K, Fridman E. 2009. Strong correlation of wild barley (*Hordeum spontaneum*) population structure with temperature and precipitation variation. *Mol Ecol* 18, 1523–1536.

International Wheat Genome Sequencing Consortium (IWGSC). 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361(6403), 7191.

Ivandic V, Hackett CA, Zhang ZJ, Staub JE, Nevo E, Thomas WT, Forster BP. 2000. Phenotypic responses of wild barley to experimentally imposed water stress. *J Exp Bot* 51, 2021–2029.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14), 1801-6.

Jakob SS, Rödder D, Engler JO, Shaaf S, Ozkan H, Blattner FR, Kilian B. 2014. Evolutionary history of wild barley (*Hordeum vulgare subsp. spontaneum*) analyzed using multilocus sequence data and paleodistribution modeling. *Genome Biol Evol* 6, 685–702.

Janzen GM, Wang L, Hufford MB. 2019. The extent of adaptive wild introgression in crops. *New Phytol* 221(3), 1279-1288.

Jayakodi M, Lu Q, Pidon H, Rabanus-Wallace MT, Bayer M, Lux T, Guo Y, Jaegle B, Badea A, Bekele W, Brar GS, Braune K, Bunk B, Chalmers KJ, Chapman B, Jørgensen ME, Feng JW, Feser M, Fiebig A, Gundlach H, Guo W, Haberer G, Hansson M, Himmelbach A, Hoffie I, Hoffie RE, Hu H, Isobe S, König P, Kale SM, Kamal N, Keeble-Gagnère G, Keller B, Knauft M, Koppolu R, Krattinger SG, Kumlehn J, Langridge P, Li C, Marone MP, Maurer A, Mayer KFX, Melzer M, Muehlbauer GJ, Murozuka E, Padmarasu S, Perovic D, Pillen K, Pin PA, Pozniak CJ, Ramsay L, Pedas PR, Rutten T, Sakuma S, Sato K, Schüler D, Schmutzer T, Scholz U, Schreiber M, Shirasawa K, Simpson C, Skadhauge B, Spannagl M, Steffenson BJ, Thomsen HC, Tibbits JF, Nielsen MTS, Trautewig C, Vequaud D, Voss C, Wang P, Waugh R, Westcott S, Rasmussen MW, Zhang R, Zhang XQ, Wicker T, Dockter C, Mascher M, Stein N. 2024. Structural variation in the pangenome of wild and domesticated barley. *Nature* 1-9.

Jayakodi M, Padmarasu S, Haberer G, Bonthala VS, Gundlach H, Monat C, Lux T, Kamal N, Lang D, Himmelbach A, Ens J, Zhang X-Q, Angessa TT, Zhou G,

Tan C, Hill C, Wang P, Schreiber M, Boston LB, Plott C, Jenkins J, Guo Y, Fiebig A, Budak H, Xu D, Zhang J, Wang C, Grimwood J, Schmutz J, Guo G, Zhang G, Mochida K, Hirayama T, Sato K, Chalmers KJ, Langridge P, Waugh R, Pozniak CJ, Scholz U, Mayer KFX, Spannagl M, Li C, Mascher M, Stein N. 2020. The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* 588, 284–289.

Jónsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29(13), 1682-4.

Kamm A. 1954. The discovery of wild six-rowed barley and wild *Hordeum* intermedium in Israel. *Kungliga Lantbrukshogskolans Annaler* 21, 287-320.

Kircher M, Sawyer S, Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* 40(1), e3.

Kistler L, Thakar HB, VanDerwarker AM, Domic A, Bergström A, George RJ, Harper TK, Allaby RG, Hirth K, Kennett DJ. 2020. Archaeological Central American maize genomes suggest ancient gene flow from South America. *Proc Natl Acad Sci* 117, 33124–33129.

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc Natl Acad Sci* 104, 1424–1429.

Lakew B, Eglinton J, Henry RJ, Baum M, Grando S, Ceccarelli S. 2011. The potential contribution of wild barley (*Hordeum vulgare ssp. spontaneum*) germplasm to drought tolerance of cultivated barley (*H. vulgare ssp. vulgare*). *Field Crops Res* 120, 161–168.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49(W1), W293-W296.

Li H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* 37(23), 4572-4574.

Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475, 493–496.

Liu M, Li Y, Ma Y, Zhao Q, Stiller J, Feng Q, Tian Q, Liu D, Han B, Liu C. 2020. The draft genome of a wild barley genotype reveals its enrichment in genes related to biotic and abiotic stresses compared to cultivated barley. *Plant Biotechnol J* 18, 443–456.

Liu X, Jones PJ, Motuzaite Matuzeviciute G, Hunt HV, Lister DL, An T, Przelomska N, Kneale CJ, Zhao Z, Jones MK. 2019. From ecological opportunism to multi-cropping: mapping food globalisation in prehistory. *Quat Sci Rev* 206, 21-28.

Lundqvist U, Abebe B, Lundqvist A. 1989. Gene interaction of induced intermedium mutations of two-row barley. *Hereditas* 111, 37–47.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol* 14(1), e1005944.

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, Bayer M, Ramsay L, Liu H, Haberer G, Zhang XQ, Zhang Q, Barrero RA, Li L, Taudien S, Groth M, Felder M, Hastie A, Šimková H, Staňková H, Vrána J, Chan S, Muñoz-Amatriaín M, Ounit R, Wanamaker S, Bolser D, Colmsee C, Schmutzer T, Aliyeva-Schnorr L, Grasso S, Tanskanen J, Chailyan A, Sampath D, Heavens D, Clissold L, Cao S, Chapman B, Dai F, Han Y, Li H, Li X, Lin C, McCooke JK, Tan C, Wang P, Wang S, Yin S, Zhou G, Poland JA, Bellgard MI, Borisjuk L, Houben A, Doležel J, Ayling S, Lonardi S, Kersey P, Langridge P, Muehlbauer GJ, Clark MD, Caccamo M, Schulman AH, Mayer KFX, Platzer M, Close TJ, Scholz U, Hansson M, Zhang G, Braumann I, Spannagl M, Li C, Waugh R, Stein N. 2017. A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544(7651), 427-433.

Mascher M, Richmond TA, Gerhardt DJ, Himmelbach A, Clissold L, Sampath D, Ayling S, Steuernagel B, Pfeifer M, D'Ascenzo M, Akhunov ED, Hedley PE, Gonzales AM, Morrell PL, Kilian B, Blattner FR, Scholz U, Mayer KF, Flavell AJ, Muehlbauer GJ, Waugh R, Jeddeloh JA, Stein N. 2013. Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. *Plant J* 76(3), 494-505.

Mascher M, Schuenemann VJ, Davidovich U, Marom N, Himmelbach A, Hübner S, Korol A, David M, Reiter E, Riehl S, Schreiber M, Vohr SH, Green RE, Dawson IK, Russell J, Kilian B, Muehlbauer GJ, Waugh R, Fahima T, Krause J,

Weiss E, Stein N. 2016. Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. *Nat Genet* 48, 1089–1093.

Mascher M, Wicker T, Jenkins J, Plott C, Lux T, Koh CS, Ens J, Gundlach H, Boston LB, Tulpová Z, Holden S, Hernández-Pinzón I, Scholz U, Mayer KFX, Spannagl M, Pozniak CJ, Sharpe AG, Šimková H, Moscou MJ, Grimwood J, Schmutz J, Stein N. 2021. Long-read sequence assembly: a technical evaluation in barley. *Plant Cell* 33, 1888-1906.

Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010(6), pdb.prot5448.

Meyer RS, Choi JY, Sanches M, Plessis A, Flowers JM, Amas J, Dorph K, Barretto A, Gross B, Fuller DQ, Bimpong IK, Ndjiondjop MN, Hazzouri KM, Gregorio GB, Purugganan MD. 2016. Domestication history and geographical adaptation inferred from a SNP map of African rice. *Nat Genet* 48(9), 1083-1088.

Milner SG, Jost M, Taketa S, Mazón ER, Himmelbach A, Oppermann M, Weise S, Knüpffer H, Basterrechea M, König P, Schüler D, Sharma R, Pasam RK, Rutten T, Guo G, Xu D, Zhang J, Herren G, Müller T, Krattinger SG, Keller B, Jiang Y, González MY, Zhao Y, Habekuß A, Färber S, Ordon F, Lange M, Börner A, Graner A, Reif JC, Scholz U, Mascher M, Stein N. 2019. Genebank genomics highlights the diversity of a global barley collection. *Nat Genet* 51(2), 319-326.

Murphy PJ, Witcombe JR, Shewry PR, Miflin BJ. 1982. The origin of six-rowed ‘wild’ barley from the western Himalaya. *Euphytica* 31, 183-192.

Morrell PL, Clegg MT. 2007. Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc Natl Acad Sci* 104(9), 3289-3294.

Morrell PL, Toleno DM, Lundy KE, Clegg MT. 2005. Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare ssp. spontaneum*) despite high rates of self-fertilization. *Proc Natl Acad Sci* 102, 2442–2447.

Naz AA, Arifuzzaman M, Muzammil S, Pillen K, Léon J. 2014. Wild barley introgression lines revealed novel QTL alleles for root and related shoot traits in the cultivated barley (*Hordeum vulgare L.*). *BMC Genet* 15, 107.

Nevo E, Chen G. 2010. Drought and salt tolerances in wild relatives for wheat and barley improvement. *Plant Cell Environ* 33, 670–685.

Orlando L, Allaby R, Skoglund P, Der Sarkissian C, Stockhammer PW, Ávila-Arcos MC, Fu Q, Krause J, Willerslev E, Stone AC, Warinner C. 2021. Ancient DNA analysis. *Nat Rev Methods Primers* 1, 14.

Pankin A, Altmüller J, Becker C, von Korff M. 2018. Targeted resequencing reveals genomic signatures of barley domestication. *New Phytol* 218, 1247-1259.

Pankin A, von Korff M. 2017. Co-evolution of methods and thoughts in cereal domestication studies: a tale of barley (*Hordeum vulgare*). *Curr Opin Plant Biol* 36, 15-21.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192(3), 1065-93.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet* 2(12), e190.

Poets AM, Fang Z, Clegg MT, Morrell PL. 2015. Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biol* 16, 173.

Posth C, Yu H, Ghalichi A, Rougier H, Crevecoeur I, Huang Y, Ringbauer H, Rohrlach AB, Nägele K, Villalba-Mouco V, Radzeviciute R, Ferraz T, Stoessel A, Tukhbatova R, Drucker DG, Lari M, Modi A, Vai S, Saupe T, Scheib CL, Catalano G, Pagani L, Talamo S, Fewlass H, Klaric L, Morala A, Rué M, Madelaine S, Crépin L, Caverne JB, Bocaege E, Ricci S, Boschin F, Bayle P, Maureille B, Le Brun-Ricalens F, Bordes JG, Oxilia G, Bortolini E, Bignon-Lau O, Debout G, Orliac M, Zazzo A, Sparacello V, Starnini E, Sineo L, van der Plicht J, Pecqueur L, Merceron G, Garcia G, Leuvrey JM, Garcia CB, Gómez-Olivencia A, Połtowicz-Bobak M, Bobak D, Le Luyer M, Storm P, Hoffmann C, Kabaciński J, Filimonova T, Shnaider S, Berezina N, González-Rabanal B, González Morales MR, Marín-Arroyo AB, López B, Alonso-Llamazares C, Ronchitelli A, Polet C, Jadin I, Cauwe N, Soler J, Coromina N, Rufí I, Cottiaux R, Clark G, Straus LG, Julien MA, Renhart S, Talaa D, Benazzi S, Romandini M, Amkreutz L, Bocherens H, Wißing C, Villotte S, de Pablo JF, Gómez-Puche M, Esquembre-Bebia MA, Bodu P, Smits L, Souffi B, Jankauskas R, Kozakaitė J, Cupillard C, Benthien H, Wehrberger K, Schmitz RW, Feine SC, Schüler T, Thevenet C, Grigorescu D, Lüth F, Kotula A, Piezonka H, Schopper F, Svoboda J, Sázelová S, Chizhevsky A, Khokhlov A, Conard NJ, Valentin F, Harvati K, Semal P, Jungklaus B, Suvorov A, Schulting R, Moiseyev V, Mannermaa K, Buzhilova A, Terberger T, Caramelli D, Altena E, Haak W, Krause J. 2023. Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* 615, 117–126.

Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, Sameri M, Azhaguvel P, Sakuma S, Dhanagond S, Sharma R, Mascher M, Himmelbach A, Gottwald S, Nair SK, Tagiri A, Yukuhiro F, Nagamura Y, Kanamori H, Matsumoto T, Willcox G, Middleton CP, Wicker T, Walther A, Waugh R, Fincher GB, Stein N, Kumlehn J, Sato K, Komatsuda T. 2015. Evolution of the grain dispersal system in barley. *Cell* 162, 527–539.

Pourkheirandish M, Kanamori H, Wu J, Sakuma S, Blattner FR, Komatsuda T. 2018. Elucidation of the origin of '*agriocrithon*' based on domestication genes questions the hypothesis that Tibet is one of the centers of barley domestication. *Plant J* 94, 525–534.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3), 559-75.

Purugganan MD. 2019. Evolutionary insights into the nature of plant domestication. *Curr Biol* 29, R705-R714.

Rabanus-Wallace MT, Hackauf B, Mascher M, Lux T, Wicker T, Gundlach H, Baez M, Houben A, Mayer KFX, Guo L, Poland J, Pozniak CJ, Walkowiak S, Melonek J, Praz CR, Schreiber M, Budak H, Heuberger M, Steuernagel B, Wulff B, Börner A, Byrns B, Čížková J, Fowler DB, Fritz A, Himmelbach A, Kaithakottil G, Keilwagen J, Keller B, Konkin D, Larsen J, Li Q, Myśków B, Padmarasu S, Rawat N, Sesiz U, Biyiklioglu-Kaya S, Sharpe A, Šimková H, Small I, Swarbreck D, Toegelová H, Tsvetkova N, Voylokov AV, Vrána J, Bauer E, Bolibok-Bragoszewska H, Doležel J, Hall A, Jia J, Korzun V, Laroche A, Ma XF, Ordon F, Özkan H, Rakoczy-Trojanowska M, Scholz U, Schulman AH,

Siekmann D, Stojałowski S, Tiwari VK, Spannagl M, Stein N. 2021. Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. *Nat Genet* 53(4), 564-573.

R Core Team. 2013. R: A language and environment for statistical computing. *Foundation for Statistical Computing, Vienna, Austria*.

Renaud G, Stenzel U, Kelso J. 2014. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res* 42(18), e141.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16(6), 276-277.

Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. 2015. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos Trans R Soc Lond B Biol Sci* 370(1660), 20130624.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4(1), 137-8.

Russell J, Mascher M, Dawson IK, Kyriakidis S, Calixto C, Freund F, Bayer M, Milne I, Marshall-Griffiths T, Heinen S, Hofstad A, Sharma R, Himmelbach A, Knauft M, van Zonneveld M, Brown JW, Schmid K, Kilian B, Muehlbauer GJ, Stein N, Waugh R. 2016. Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat Genet* 48, 1024–1030.

Sallam AH, Guo Y, Jayakodi M, Himmelbach A, Fiebig A, Simmons J, Bethke G, Lee Y, Spanner R, Badea A, Baum M, Belzile F, Ben-David R, Brueggeman

R, Case A, Cattivelli L, Davis M, Dockter C, Doležel J, Dreiseitl A, Gavin R, Glick L, Greiner S, Hamilton R, Hayes PM, Heisel S, Henson C, Kilian B, Komatsuda T, Li C, Liu C, Mahalingam R, Maruschewski M, Matny O, Maurer A, Mayer KFX, Mayrose I, Morrell P, Moscou M, Muehlbauer GJ, Oono Y, Ordon F, Özkan H, Pecinka A, Perovic D, Pillen K, Pourkheirandish M, Russell J, Šafář J, Salvi S, Sanchez-Garcia M, Sato K, Schmutzer T, Scholz U, Scott J, Singh Brar G, Smith KP, Sorrells ME, Spannagl M, Stein N, Tondelli A, Tuberosa R, Tucker J, Turkington T, Valkoun J, Verma RPS, Vinje MA, von Korff Schmising M, Walling JG, Waugh R, Wise RP, Wulff BBH, Yang S, Zhang G, Mascher M, Steffenson BJ. 2024. Whole-genome sequencing of the wild barley diversity collection: a resource for identifying and exploiting genetic variation for cultivated barley improvement. *bioRxiv*. doi: https://doi.org/10.1101/2024.11.18.624148

Sallam AH, Tyagi P, Brown-Guedira G, Muehlbauer GJ, Hulse A, Steffenson BJ. 2017. Genome-wide association mapping of stem rust resistance in *Hordeum vulgare subsp. spontaneum*. *G3 (Bethesda)* 7(10), 3491-3507.

Schiemann E. 1932. Entstehung der Kulturpflanzen. *Borntraeger (Handbuch der Vererbungswissenschaft, Bd. 3, L)*.

Schiemann E. 1951. New results on the history of cultivated cereals. *Heredity* 5, 305-320.

Schmalenbach I, Léon J, Pillen K. 2009. Identification and verification of QTLs for agronomic traits using wild barley introgression lines. *Theor Appl Genet* 118, 483–497.

Schmalenbach I, Pillen K. 2009. Detection and verification of malting quality QTLs using wild barley introgression lines. *Theor Appl Genet* 118, 1411–1427.

Schreiber M, Stein N, Mascher M. 2018. Genomic approaches for studying crop evolution. *Genome Biol* 19, 140.

Shah N, Wakabayashi T, Kawamura Y, Skovbjerg CK, Wang MZ, Mustamin Y, Isomura Y, Gupta V, Jin H, Mun T, Sandal N, Azuma F, Fukai E, Seren Ü, Kusakabe S, Kikuchi Y, Nitanda S, Kumaki T, Hashiguchi M, Tanaka H, Hayashi A, Sønderkær M, Nielsen KL, Schneeberger K, Vilhjalmsson B, Akashi R, Stougaard J, Sato S, Schierup MH, Andersen SU. 2020. Extreme genetic signatures of local adaptation during *Lotus japonicus* colonization of Japan. *Nat Commun* 11(1), 253.

Snir A, Nadel D, Groman-Yaroslavski I, Melamed Y, Sternberg M, Bar-Yosef O, Weiss E. 2015. The origin of cultivation and proto-weeds, long before Neolithic farming. *PLoS One* 10(7), e0131422.

Snyder CW. 2016. Evolution of global temperature over the past two million years. *Nature* 538(7624), 226-228.

Spellerberg IF, Fedor PJ. 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index. *Global Ecol Biogeogr* 12(3), 177-179.

Steffenson BJ, Olivera P, Roy JK, Jin Y, Smith KP, Muehlbauer GJ. 2007. A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. *Aust J Agric Res* 58, 532–544.

Suprunova T, Krugman T, Distelfeld A, Fahima T, Nevo E, Korol A. 2007. Identification of a novel gene (*Hsdr4*) involved in water-stress tolerance in wild barley. *Plant Mol Biol* 64, 17–34.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105(2), 437-60.

Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda K, Nomura M, Suzuki T, Matsumoto T, Sato K, Kanamori H, Kawasaki S, Takeda K. 2008. Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc Natl Acad Sci* 105(10), 4062-4067.

Tanno K, Takeda K. 2004. On the origin of six-rowed barley with brittle rachis, *agriocrithon* (*Hordeum vulgare ssp. vulgare f. agriocrithon* (*Aberg*) *Bowd*.), based on a DNA marker closely linked to the *vrs1* (six-row gene) locus. *Theor Appl Genet* 110, 145-150.

Thirulogachandar V, Alqudah AM, Koppolu R, Rutten T, Graner A, Hensel G, Kumlehn J, Bräutigam A, Sreenivasulu N, Schnurbusch T, Kuhlmann M. 2017. Leaf primordium size specifies leaf width and vein number among row-type classes in barley. *Plant J* 91(4), 601-612.

von Korff M, Wang H, Léon J, Pillen K. 2006. AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H. vulgare ssp. spontaneum*). *Theor Appl Genet* 112, 1221–1231.

Wang L, Ji Y, Hu Y, Hu H, Jia X, Jiang M, Zhang X, Zhao L, Zhang Y, Jia Y, Qin C, Yu L, Huang J, Yang S, Hurst LD, Tian D. 2019. The architecture of intra-organism mutation rate variation in plants. *PLoS Biol* 17(4), e3000191.

Wang Z, Wang W, Xie X, Wang Y, Yang Z, Peng H, Xin M, Yao Y, Hu Z, Liu J, Su Z, Xie C, Li B, Ni Z, Sun Q, Guo W. 2022. Dispersed emergence and protracted domestication of polyploid wheat uncovered by mosaic ancestral haploblock inference. *Nat Commun* 13, 3891.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 7(2), 256-76.

Wendler N, Mascher M, Nöh C, Himmelbach A, Scholz U, Ruge-Wehling B, Stein N. 2014. Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnol J* 12(8), 1122-31.

Zeng X, Guo Y, Xu Q, Mascher M, Guo G, Li S, Mao L, Liu Q, Xia Z, Zhou J, Yuan H, Tai S, Wang Y, Wei Z, Song L, Zha S, Li S, Tang Y, Bai L, Zhuang Z, He W, Zhao S, Fang X, Gao Q, Yin Y, Wang J, Yang H, Zhang J, Henry RJ, Stein N, Tashi N. 2018. Origin and evolution of qingke barley in Tibet. *Nat Commun* 9(1), 5433.

Zhang C, Dong SS, Xu JY, He WM, Yang TL. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35(10), 1786-1788.

Zhang J, Jiang L, Yu L, Huan X, Zhou L, Wang C, Jin J, Zuo X, Wu N, Zhao Z, Sun H, Yu Z, Zhang G, Zhu J, Wu Z, Dong Y, Fan B, Shen C, Lu H. 2024. Rice's trajectory from wild to domesticated in East Asia. *Science* 384(6698), 901-906.

Zohary D, Hopf M, Weiss E. 2012. Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. *Oxford University Press*.

9. Appendix

Table S1: Passport data and population assignment of 380 wild barley samples.

accession	latitude	longitude	subpopulation	coverage
FT67	34.89	31.54	SL	high
FT37	34.47	31.4	SL	high
WBDC_100	35.62	30.77	SL	high
WBDC_294	34.98	31.75	SL	high
WBDC_260	35.5	30.5	SL	high
WBDC_290	34.93	31.67	SL	high
WBDC_267	35.55	30.7	SL	high
WBDC_034	NA	NA	SL	high
WBDC_270	34.87	31.68	SL	high
WBDC_293	35.17	31.8	SL	high
WBDC_291	34.92	31.72	SL	high
WBDC_033	35.22	31.27	SL	high
WBDC_035	NA	NA	SL	high
WBDC_278	34.77	31.83	SL	high
WBDC_101	35.57	30.7	SL	high
WBDC_038	35.21	46.77	SL	high
WBDC_185	21.72	32.77	SL	high
WBDC_074	22.05	32.8	SL	high
WBDC_042	35.53	32.97	SL	high
WBDC_048	44.48	37.25	SL	high
WBDC_252	35.68	32.28	SL	high
WBDC_283	35.12	32.6	SL	high
WBDC_083	35.92	32.17	SL	high
WBDC_266	35.67	30.88	SL	high
WBDC_032	35.13	33	SL	high
FT627	35.9	32.02	SL	high
FT18	35.4	32.07	SL	high
WBDC_180	24.23	31.83	SL	high
WBDC_181	36.02	32.02	SL	high

HOR_9476	NA	NA	SL	high
WBDC_082	35.78	32.27	SL	high
WBDC_256	35.63	32.33	SL	high
FT64	35.12	31.35	SL	high
WBDC_242	35.87	32.12	SL	high
HOR_1647	NA	NA	SL	high
WBDC_265	35.67	30.93	SL	high
WBDC_079	35.88	32.25	SL	high
WBDC_104	35.62	31.18	SL	high
WBDC_243	35.67	32.03	SL	high
WBDC_292	35.02	31.8	SL	high
WBDC_234	33.88	35.15	SL	high
WBDC_199	36.64	34.91	NL	high
WBDC_070	35.82	35.61	NL	high
WBDC_171	35.85	33.58	NL	high
WBDC_303	36.11	33.74	NL	high
WBDC_306	36.13	33.73	NL	high
WBDC_143	35.9	33.57	NL	high
WBDC_128	36.17	33.84	NL	high
WBDC_170	35.95	33.63	NL	high
WBDC_304	35.96	33.64	NL	high
WBDC_140	36.08	34.2	NL	high
WBDC_305	36	33.66	NL	high
WBDC_061	36.58	36.16	NL	high
WBDC_064	36.24	35.8	NL	high
WBDC_062	36.84	36.38	NL	high
WBDC_202	36.55	35.6	NL	high
WBDC_004	36.54	35.97	NL	high
WBDC_055	35.82	35.61	NL	high
WBDC_001	36.7	36.22	NL	high
WBDC_009	35.75	31.53	NL	high
WBDC_203	36.45	35.63	NL	high
WBDC_299	36.71	36.66	NL	high
WBDC_063	36.64	36.72	NL	high

WBDC_106	36.74	34.94	NL	high
WBDC_314	36.55	34.88	NL	high
FT590	36.7	35.01	NL	high
WBDC_300	36.86	36.37	NL	high
WBDC_107	36.73	34.75	NL	high
WBDC_169	35.78	33.5	NL	high
WBDC_145	36.02	33.8	NL	high
WBDC_196	36.93	36.99	NL	high
WBDC_342	36.9	36.97	NL	high
WBDC_142	36.08	34.02	NL	high
WBDC_237	35.62	31.53	SD	high
WBDC_005	35.92	32.45	SD	high
WBDC_007	35.92	32.33	SD	high
WBDC_046	35.9	32.48	SD	high
WBDC_097	35.83	31.28	SD	high
WBDC_129	36.42	32.81	SD	high
WBDC_195	37.22	36.81	SD	high
WBDC_206	36.6	32.55	SD	high
WBDC_236	35.62	31.6	SD	high
WBDC_245	35.87	32.17	SD	high
WBDC_258	35.75	31.28	SD	high
WBDC_285	35.45	32.98	SD	high
WBDC_308	36.38	32.6	SD	high
WBDC_319	36.68	32.49	SD	high
WBDC_127	36.73	32.94	SD	high
WBDC_320	36.62	32.42	SD	high
WBDC_240	35.94	31.88	SD	high
WBDC_112	36.59	33.98	SD	high
WBDC_078	40.86	36.77	NM	high
FT507	43.42	36.38	NM	high
WBDC_056	37.12	36.72	NM	high
WBDC_155	42.17	36.33	NM	high
WBDC_157	43.42	36.38	NM	high
WBDC_161	37.58	36.65	NM	high

WBDC_164	41.64	37.01	NM	high
WBDC_167	40.36	36.4	NM	high
WBDC_168	35.72	33.57	NM	high
WBDC_178	43.28	36.08	NM	high
WBDC_191	37.72	36.84	NM	high
WBDC_200	36.66	34.94	NM	high
WBDC_295	42.07	37.08	NM	high
WBDC_297	41.75	37.06	NM	high
WBDC_298	41.09	37.08	NM	high
WBDC_187	37.46	36.69	NM	high
WBDC_177	43.13	36.35	NM	high
WBDC_186	37.35	36.88	NM	high
WBDC_189	37.63	36.76	NM	high
WBDC_248	35.85	32.53	NM	high
WBDC_296	41.56	37.06	NM	high
WBDC_317	38.76	35.52	NM	high
WBDC_207	71.18	40.34	CA	high
WBDC_117	56.05	38.17	CA	high
WBDC_201	36.89	35.56	CA	high
WBDC_208	69.91	41.61	CA	high
WBDC_212	67.09	40.01	CA	high
WBDC_220	69.7	42.42	CA	high
WBDC_223	NA	NA	CA	high
WBDC_330	56.29	38.77	CA	high
WBDC_332	56.42	38.43	CA	high
WBDC_346	66.37	39.92	CA	high
WBDC_347	67	37.8	CA	high
WBDC_224	NA	NA	CA	high
WBDC_026	69.1	37.65	CA	high
WBDC_209	68.4	40.13	CA	high
WBDC_125	66.47	38.8	CA	high
WBDC_216	56.85	38.73	CA	high
WBDC_036	63	34.57	CA	high
WBDC_120	67.5	39.47	CA	high

WBDC_012	64.82	36.28	CA	high
WBDC_221	69.33	40.13	CA	high
WBDC_233	68.7	35.93	CA	high
WBDC_331	56.11	39.24	CA	high
WBDC_323	58.55	37.9	CA	high
WBDC_152	51.25	35.5	CA	high
WBDC_211	68.08	39.71	CA	high
WBDC_219	69.46	42.39	CA	high
WBDC_345	66.83	38.95	CA	high
WBDC_329	55.6	38.18	CA	high
WBDC_025	66.9	30.3	CA	high
WBDC_335	55.62	39.25	CA	high
WBDC_015	62.18	34.35	CA	high
WBDC_355	NA	NA	CA	high
WBDC_210	67.68	39.93	CA	high
WBDC_014	69	35.75	CA	high
WBDC_213	66.56	39.55	CA	high
WBDC_119	67.58	40.08	CA	high
WBDC_133	35.82	33.62	admixed	high
WBDC_348	NA	NA	admixed	high
WBDC_349	NA	NA	admixed	high
FT262	37.28	36.78	admixed	high
WBDC_276	34.9	32.35	admixed	high
WBDC_052	36.02	32.32	admixed	high
WBDC_284	35.6	32.67	admixed	high
WBDC_244	35.71	32.08	admixed	high
WBDC_028	34.93	31.95	admixed	high
HOR_22109	NA	NA	admixed	high
WBDC_030	35.18	33.08	admixed	high
WBDC_089	35.8	31.83	admixed	high
WBDC_286	35.47	32.97	admixed	high
WBDC_126	35.32	33.45	admixed	high
WBDC_138	35.76	33.42	admixed	high
FT219	35.76	33.29	admixed	high

WBDC_269	35.87	33.42	admixed	high
WBDC_255	35.67	32.35	admixed	high
WBDC_274	34.77	30.87	admixed	high
WBDC_354	NA	NA	admixed	high
WBDC_253	35.82	32.35	admixed	high
WBDC_312	36.76	34.47	admixed	high
WBDC_179	20.9	32.5	admixed	high
WBDC_277	34.93	32.15	admixed	high
WBDC_022	NA	NA	admixed	high
HOR_12568	NA	NA	admixed	high
WBDC_198	36.47	34.82	admixed	high
WBDC_194	37.19	36.97	admixed	high
WBDC_254	35.7	32.62	admixed	high
WBDC_250	35.93	32.58	admixed	high
WBDC_105	35.83	32.65	admixed	high
WBDC_134	35.77	33.62	admixed	high
WBDC_246	35.75	32.68	admixed	high
WBDC_193	37.47	37.32	admixed	high
FT630	35.87	33.47	admixed	high
WBDC_340	37.33	37.25	admixed	high
WBDC_338	37.54	37.27	admixed	high
WBDC_183	35.72	32.63	admixed	high
WBDC_008	35.62	32.67	admixed	high
FT276	35.67	32.58	admixed	high
FT731	37.28	37.42	admixed	high
FT670	36.93	36.99	admixed	high
WBDC_341	36.93	36.95	admixed	high
WBDC_281	35.58	33.03	admixed	high
WBDC_002	37.44	35.71	admixed	high
WBDC_302	35.88	33.37	admixed	high
WBDC_132	35.87	33.47	admixed	high
WBDC_343	36.95	36.87	admixed	high
WBDC_197	37.77	36.48	admixed	high
WBDC_017	35.82	32.74	admixed	high

WBDC_315	36.85	34.94	admixed	high
WBDC_054	38.36	34.57	admixed	high
WBDC_043	35.58	32.98	admixed	high
WBDC_081	35.65	32.48	admixed	high
WBDC_080	35.78	32.53	admixed	high
WBDC_271	34.48	31.43	admixed	high
WBDC_309	37.01	35.03	admixed	high
WBDC_040	34.95	32.7	admixed	high
WBDC_247	35.7	32.37	admixed	high
FT581	41.72	37.97	admixed	high
FT734	38.78	37.41	admixed	high
WBDC_109	36.72	32.77	admixed	high
WBDC_006	36.17	32.3	admixed	high
WBDC_160	36.6	32.96	admixed	high
WBDC_110	36.79	32.78	admixed	high
WBDC_318	36.74	32.6	admixed	high
WBDC_130	36.18	32.83	admixed	high
WBDC_068	36.72	32.82	admixed	high
WBDC_159	36.79	32.68	admixed	high
WBDC_139	36.1	33.93	admixed	high
FT582	42.1	37.09	admixed	high
WBDC_188	37.52	36.88	admixed	high
WBDC_010	65.73	36.67	admixed	high
WBDC_146	45.7	36.75	admixed	high
WBDC_148	45.47	37.07	admixed	high
WBDC_149	45	38.08	admixed	high
WBDC_147	45.17	37.5	admixed	high
FT660	45.7	36.75	admixed	high
WBDC_023	48.45	33.5	admixed	high
WBDC_051	39.03	34.76	admixed	high
WBDC_307	36.07	33.2	admixed	high
WBDC_165	42.21	37.29	admixed	high
WBDC_057	36.84	33.65	admixed	high
WBDC_093	36.37	32.05	admixed	high

WBDC_268	36.76	31.75	admixed	high
WBDC_287	35.5	32.5	admixed	high
FT878	46.02	34.44	admixed	high
WBDC_337	37.52	37.23	admixed	high
WBDC_150	47.2	38.5	admixed	high
FT746	36.73	36.98	admixed	high
WBDC_153	49.8	34.08	admixed	high
WBDC_192	37.61	37.04	admixed	high
WBDC_122	48.17	32.87	admixed	high
WBDC_013	44.83	35.53	admixed	high
WBDC_011	43.52	36	admixed	high
WBDC_049	44.48	37.25	admixed	high
HOR_2690	NA	NA	admixed	high
WBDC_156	41.65	36.42	admixed	high
WBDC_123	58.47	36.42	admixed	high
WBDC_085	35.65	31.98	admixed	high
WBDC_019	45.72	36.75	admixed	high
WBDC_137	35.82	33.45	admixed	high
FT144	NA	NA	NA	low or admixed*
FT147	NA	NA	NA	low or admixed*
FT218	NA	NA	NA	low or admixed*
FT232	NA	NA	NA	low or admixed*
FT248	NA	NA	NA	low or admixed*
FT268	NA	NA	NA	low or admixed*
FT272	NA	NA	NA	low or admixed*
FT279	NA	NA	NA	low or admixed*
FT286	NA	NA	NA	low or admixed*
FT31	NA	NA	NA	low or admixed*
FT332	NA	NA	NA	low or admixed*
FT333	NA	NA	NA	low or admixed*
FT338	NA	NA	NA	low or admixed*
FT340	NA	NA	NA	low or admixed*
FT361	NA	NA	NA	low or admixed*
FT363	NA	NA	NA	low or admixed*

FT376	NA	NA	NA	low or admixed*
FT42	NA	NA	NA	low or admixed*
FT462	NA	NA	NA	low or admixed*
FT469	NA	NA	NA	low or admixed*
FT470	NA	NA	NA	low or admixed*
FT473	NA	NA	NA	low or admixed*
FT56	NA	NA	NA	low or admixed*
FT566	NA	NA	NA	low or admixed*
FT568	NA	NA	NA	low or admixed*
FT572	NA	NA	NA	low or admixed*
FT589	NA	NA	NA	low or admixed*
FT592	NA	NA	NA	low or admixed*
FT604	NA	NA	NA	low or admixed*
FT613	NA	NA	NA	low or admixed*
FT616	NA	NA	NA	low or admixed*
FT628	NA	NA	NA	low or admixed*
FT631	NA	NA	NA	low or admixed*
FT632	NA	NA	NA	low or admixed*
FT657	NA	NA	NA	low or admixed*
FT658	NA	NA	NA	low or admixed*
FT661	NA	NA	NA	low or admixed*
FT671	NA	NA	NA	low or admixed*
FT730	NA	NA	NA	low or admixed*
FT741	NA	NA	NA	low or admixed*
FT747	NA	NA	NA	low or admixed*
FT748	NA	NA	NA	low or admixed*
FT75	NA	NA	NA	low or admixed*
FT754	NA	NA	NA	low or admixed*
FT871	NA	NA	NA	low or admixed*
FT873	NA	NA	NA	low or admixed*
FT875	NA	NA	NA	low or admixed*
FT879	NA	NA	NA	low or admixed*
FT880	NA	NA	NA	low or admixed*
FT885	NA	NA	NA	low or admixed*

FT886	NA	NA	NA	low or admixed*
HOR_10478	NA	NA	NA	low or admixed*
HOR_11183	NA	NA	NA	low or admixed*
HOR_12482	NA	NA	NA	low or admixed*
HOR_12541	NA	NA	NA	low or admixed*
HOR_12562	NA	NA	NA	low or admixed*
HOR_12854	NA	NA	NA	low or admixed*
HOR_12856	NA	NA	NA	low or admixed*
HOR_12866	NA	NA	NA	low or admixed*
HOR_12943	NA	NA	NA	low or admixed*
HOR_12996	NA	NA	NA	low or admixed*
HOR_13009	NA	NA	NA	low or admixed*
HOR_13035	NA	NA	NA	low or admixed*
HOR_13222	NA	NA	NA	low or admixed*
HOR_13295	NA	NA	NA	low or admixed*
HOR_13762	NA	NA	NA	low or admixed*
HOR_21893	NA	NA	NA	low or admixed*
HOR_22012	NA	NA	NA	low or admixed*
HOR_22078	NA	NA	NA	low or admixed*
HOR_4857	NA	NA	NA	low or admixed*
HOR_8538	NA	NA	NA	low or admixed*
WBDC_016	NA	NA	NA	low or admixed*
WBDC_018	NA	NA	NA	low or admixed*
WBDC_020	NA	NA	NA	low or admixed*
WBDC_021	NA	NA	NA	low or admixed*
WBDC_024	NA	NA	NA	low or admixed*
WBDC_029	NA	NA	NA	low or admixed*
WBDC_031	NA	NA	NA	low or admixed*
WBDC_041	NA	NA	NA	low or admixed*
WBDC_044	NA	NA	NA	low or admixed*
WBDC_045	NA	NA	NA	low or admixed*
WBDC_053	NA	NA	NA	low or admixed*
WBDC_058	NA	NA	NA	low or admixed*
WBDC_060	NA	NA	NA	low or admixed*

WBDC_066	NA	NA	NA	low or admixed*
WBDC_067	NA	NA	NA	low or admixed*
WBDC_072	NA	NA	NA	low or admixed*
WBDC_073	NA	NA	NA	low or admixed*
WBDC_075	NA	NA	NA	low or admixed*
WBDC_092	NA	NA	NA	low or admixed*
WBDC_094	NA	NA	NA	low or admixed*
WBDC_095	NA	NA	NA	low or admixed*
WBDC_108	NA	NA	NA	low or admixed*
WBDC_111	NA	NA	NA	low or admixed*
WBDC_113	NA	NA	NA	low or admixed*
WBDC_115	NA	NA	NA	low or admixed*
WBDC_116	NA	NA	NA	low or admixed*
WBDC_121	NA	NA	NA	low or admixed*
WBDC_136	NA	NA	NA	low or admixed*
WBDC_151	NA	NA	NA	low or admixed*
WBDC_172	NA	NA	NA	low or admixed*
WBDC_173	NA	NA	NA	low or admixed*
WBDC_174	NA	NA	NA	low or admixed*
WBDC_175	NA	NA	NA	low or admixed*
WBDC_182	NA	NA	NA	low or admixed*
WBDC_184	NA	NA	NA	low or admixed*
WBDC_190	NA	NA	NA	low or admixed*
WBDC_204	NA	NA	NA	low or admixed*
WBDC_205	NA	NA	NA	low or admixed*
WBDC_214	NA	NA	NA	low or admixed*
WBDC_215	NA	NA	NA	low or admixed*
WBDC_217	NA	NA	NA	low or admixed*
WBDC_228	NA	NA	NA	low or admixed*
WBDC_231	NA	NA	NA	low or admixed*
WBDC_232	NA	NA	NA	low or admixed*
WBDC_238	NA	NA	NA	low or admixed*
WBDC_241	NA	NA	NA	low or admixed*
WBDC_257	NA	NA	NA	low or admixed*

WBDC_275	NA	NA	NA	low or admixed*
WBDC_279	NA	NA	NA	low or admixed*
WBDC_282	NA	NA	NA	low or admixed*
WBDC_289	NA	NA	NA	low or admixed*
WBDC_311	NA	NA	NA	low or admixed*
WBDC_324	NA	NA	NA	low or admixed*
WBDC_326	NA	NA	NA	low or admixed*
WBDC_333	NA	NA	NA	low or admixed*
WBDC_336	NA	NA	NA	low or admixed*
WBDC_344	NA	NA	NA	low or admixed*
WBDC_350	NA	NA	NA	low or admixed*

***Low and admixed: accessions with low coverage or admixed with domesticated barley which were not used in the subsequent analysis.**

Table S2: Passport data and population assignment of 302 domesticated barley samples.

sample ID	subpopulation	landrace/cultivar/ breeding_material	coverage	*Btr1Btr2* genotype	spike row type	*Vrs1* genotype	*Nud* genotype	spring/winter
HOR_16340	ISR-THS	breeding_material	high	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_9582	ISR-THS	breeding_material	high	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_3877	ISR-THS	cultivar	high	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_20586	ISR-THS	landrace	high	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_3941	ISR-THS	breeding_material	high	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_4558	ISR-THS	cultivar	high	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_21550	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_17925	SYR-THM	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	winter
HOR_19814	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_17911	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_20121	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_21878	SYR-THM	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_20128	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	winter
HOR_20159	SYR-THM	landrace	high	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_21582	SYR-THM	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_21552	SYR-THM	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_10650	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_10741	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_10466	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9618	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_10899	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14792	GEO-THS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_10756	GEO-THS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_10777	GEO-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14498	IRN-THS	landrace	high	*btr1Btr2*	2-rowed	NA	*Nud*	spring

HOR_7964	IRN-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_21339	IRN-THS	landrace	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14351	IRN-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_17093	IRN-THS	landrace	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_154	IRN-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_13886	TUR-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_20365	TUR-THM	landrace	high	*Btr1btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_2784	TUR-THM	landrace	high	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_4604	TUR-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_21605	TUR-THM	landrace	high	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_13821	TUR-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_12205	EU-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_12152	EU-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_12203	EU-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_20674	EU-THM	breeding_material	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_12168	EU-THM	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_21695	EU-THS	breeding_material	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_3620	EU-THS	cultivar	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_18788	EU-THS	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_17444	EU-THS	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_18766	EU-THS	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_11117	EU-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_20143	EU-THS	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14347	EU-THS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_17549	EU-THS	cultivar	high	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14439	EU-THS	cultivar	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_13928	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_10089	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_18924	ME-SHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_16080	ME-SHS	breeding_material	high	*Btr1btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_6954	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_703	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_9930	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_11480	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_7319	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_10157	ME-SHS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_17570	EU-SHW	breeding_material	high	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_17134	EU-SHW	cultivar	high	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	winter
HOR_14135	EU-SHW	cultivar	high	*btr1Btr2*	NA	*vrs1.a3*	*Nud*	winter
HOR_16958	EU-SHW	cultivar	high	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_17572	EU-SHW	landrace	high	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_21558	EU-SHW	cultivar	high	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_3081	EU-SHW	cultivar	high	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_14217	EU-SHM	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_18794	EU-SHM	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_15573	EU-SHM	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_18935	EU-SHM	cultivar	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_12157	EU-SHM	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_3327	EU-SHM	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_19289	EU-SHM	landrace	high	*btr1Btr2*	NA	*vrs1.a3*	*Nud*	spring
HOR_6908	EU-SHM	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring

HOR_15915	EU-SHM	landrace	high	*Btr1btr2*	NA	*vrs1.a3*	*Nud*	spring
HOR_5913	ETH-MHS	landrace	high	*btr1Btr2*	deficiens	*vrs1.b3*	*Nud*	spring
HOR_7872	ETH-MHS	landrace	high	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5620	ETH-MHS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_14765	ETH-MHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_6033	ETH-MHS	landrace	high	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5647	ETH-MHS	landrace	high	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5313	ETH-MHS	landrace	high	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_6220	ETH-MHS	landrace	high	*btr1Btr2*	labile	*vrs1.a1*	*Nud*	spring
HOR_15456	ETH-MNS	breeding_material	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*nud*	spring
HOR_18139	ETH-MNS	landrace	high	*btr1Btr2*	NA	*vrs1.b3*	*nud*	spring
HOR_16231	ETH-MNS	landrace	high	*btr1Btr2*	NA	*vrs1.b3*	*nud*	spring
HOR_15224	ETH-MNS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*nud*	spring
HOR_6360	ETH-MNS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_18134	ETH-MNS	landrace	high	*btr1Btr2*	NA	*vrs1.b3*	*nud*	spring
HOR_6699	ETH-MNS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a3*	*nud*	spring
HOR_3290	ETH-MNS	landrace	high	*btr1Btr2*	2-rowed	*vrs1.b3*	*nud*	spring
HOR_9043	ETH-MNS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_15742	CA-SHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_21087	CA-SHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_3272	CA-SHS	cultivar	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_15890	CA-SHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_11839	CA-SHS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_7219	CA-SHS	landrace	high	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_14876	CA-SHS	landrace	high	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_8347	CA-SNS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_17236	CA-SNS	NA	high	*Btr1btr2*	NA	*vrs1.a4*	*nud*	spring
HOR_20609	CA-SNS	breeding_material	high	*Btr1btr2*	NA	*vrs1.a4*	*nud*	spring
HOR_11112	CA-SNS	landrace	high	*Btr1btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_7571	CA-SNS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_2383	CA-SNS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_148	CA-SNS	landrace	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_10243	EA-SHM	cultivar	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_18624	EA-SHM	landrace	high	*Btr1btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_11735	EA-SHM	breeding_material	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_11546	EA-SHM	cultivar	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_22161	EA-SHM	landrace	high	*Btr1btr2*	intermedium	*vrs1.a4*	*Nud*	winter
HOR_12367	EA-SHM	cultivar	high	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_2024	EA-SHM	cultivar	high	*Btr1btr2*	6-rowed	*vrs1.a4*	*Nud*	spring
HOR_3726	EA-SHM	cultivar	high	*Btr1btr2*	intermedium	*vrs1.a4*	*Nud*	spring
HOR_9822	ISR-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_2895	ISR-THS	breeding_material	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_20678	ISR-THS	breeding_material	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_3564	ISR-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_4073	ISR-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_16343	ISR-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_21256	ISR-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_18674	ISR-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_3997	ISR-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_8209	ISR-THS	breeding_material	low	*btr1Btr2*	2-rowed	*vrs1.b2*	*Nud*	spring
HOR_18954	ISR-THS	breeding_material	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring

HOR_15171	SYR-THM	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_19818	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_12127	SYR-THM	cultivar	low	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_21570	SYR-THM	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_4767	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_14101	SYR-THM	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_20156	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_21546	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_4969	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_21566	SYR-THM	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_18455	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	winter
HOR_2830	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_12520	SYR-THM	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_10621	SYR-THM	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_10464	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9871	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9611	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9617	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_4726	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9612	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14796	GEO-THS	landrace	low	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_10902	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_10774	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_10758	GEO-THS	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_10465	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_9606	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_10649	GEO-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14875	IRN-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_2770	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_2825	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_2872	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_2803	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_15195	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14338	IRN-THS	landrace	low	*Btr1btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_15232	IRN-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14899	IRN-THS	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_15156	IRN-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14055	IRN-THS	breeding_material	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_2832	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_21322	IRN-THS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_2808	IRN-THS	landrace	low	*btr1Btr2*	2-rowed	NA	*Nud*	spring
HOR_4295	TUR-THM	landrace	low	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_14938	TUR-THM	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	intermediate
HOR_16775	TUR-THM	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_18117	TUR-THM	landrace	low	*Btr1btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14054	TUR-THM	landrace	low	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_4332	TUR-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_4316	TUR-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_2847	TUR-THM	landrace	low	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_8124	TUR-THM	landrace	low	*Btr1btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_4532	TUR-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter

HOR_8086	TUR-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_12164	EU-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_12190	EU-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_22162	EU-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_3810	EU-THM	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_15056	EU-THM	cultivar	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_4378	EU-THM	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	winter
HOR_12210	EU-THM	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_16758	EU-THM	cultivar	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_21636	EU-THM	breeding_material	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_20197	EU-THM	cultivar	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_14023	EU-THM	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_8817	EU-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_10953	EU-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_12206	EU-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14801	EU-THS	breeding_material	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_8664	EU-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_14859	EU-THS	breeding_material	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_3623	EU-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_2534	EU-THS	cultivar	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_9851	EU-THS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*Nud*	spring
HOR_17533	EU-THS	cultivar	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_10226	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	NA	*Nud*	spring
HOR_13808	ME-SHS	landrace	low	*Btr1btr2*	NA	NA	*Nud*	winter
HOR_1181	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	NA	*Nud*	spring
HOR_685	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_19981	ME-SHS	landrace	low	*Btr1btr2*	NA	NA	*Nud*	spring
HOR_20001	ME-SHS	cultivar	low	*Btr1btr2*	NA	NA	*Nud*	spring
HOR_9845	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_4105	ME-SHS	breeding_material	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_556	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	NA	*Nud*	spring
HOR_683	ME-SHS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_17372	EU-SHW	cultivar	low	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_10218	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_4672	EU-SHW	breeding_material	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_18704	EU-SHW	cultivar	low	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_13801	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_2273	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_9298	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_19472	EU-SHW	breeding_material	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_17529	EU-SHW	cultivar	low	*btr1Btr2*	NA	*vrs1.a3*	*Nud*	winter
HOR_14845	EU-SHW	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_10217	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_17739	EU-SHW	cultivar	low	*btr1Btr2*	NA	*vrs1.a2*	*Nud*	winter
HOR_9896	EU-SHW	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a2*	*Nud*	winter
HOR_11211	EU-SHM	breeding_material	low	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
BCC_1486	EU-SHM	cultivar	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_4938	EU-SHM	breeding_material	low	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_15576	EU-SHM	landrace	low	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_19066	EU-SHM	cultivar	low	*btr1Btr2*	NA	*vrs1.b3*	*Nud*	spring
HOR_2963	EU-SHM	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring

HOR_4223	EU-SHM	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_8234	EU-SHM	breeding_material	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_4366	EU-SHM	cultivar	low	*btr1Btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_3688	EU-SHM	cultivar	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_10630	EU-SHM	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
BCC_1504	EU-SHM	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_4655	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_14757	ETH-MHS	landrace	low	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_6495	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_9893	ETH-MHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_20162	ETH-MHS	landrace	low	*btr1Btr2*	NA	*vrs1.b2*	*Nud*	spring
HOR_9313	ETH-MHS	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_7922	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_9131	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5315	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5820	ETH-MHS	landrace	low	*btr1Btr2*	deficiens	*vrs1.b2*	*Nud*	spring
HOR_5033	ETH-MHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_9096	ETH-MHS	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_20357	ETH-MNS	landrace	low	*btr1Btr2*	NA	NA	*nud*	spring
HOR_5385	ETH-MNS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_20453	ETH-MNS	landrace	low	*btr1Btr2*	NA	*vrs1.a1*	*nud*	spring
HOR_3597	ETH-MNS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*nud*	spring
HOR_4288	ETH-MNS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*nud*	spring
HOR_18041	ETH-MNS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*nud*	spring
HOR_11468	ETH-MNS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_17385	ETH-MNS	landrace	low	*btr1Btr2*	NA	NA	*nud*	spring
HOR_3600	ETH-MNS	landrace	low	*btr1Btr2*	2-rowed	*vrs1.b3*	*nud*	spring
HOR_9496	ETH-MNS	landrace	low	*btr1Btr2*	6-rowed	NA	*nud*	spring
HOR_17049	ETH-MNS	landrace	low	*btr1Btr2*	NA	*vrs1.b3*	*nud*	spring
HOR_5958	ETH-MNS	landrace	low	*btr1Btr2*	6-rowed	NA	*nud*	spring
HOR_14910	CA-SHS	landrace	low	*btr1Btr2*	NA	*vrs1.a1*	*Nud*	spring
HOR_1577	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_3838	CA-SHS	breeding_material	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_9528	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_1852	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_21273	CA-SHS	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_4575	CA-SHS	breeding_material	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_19601	CA-SHS	breeding_material	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_6989	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_11824	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_19937	CA-SHS	landrace	low	*btr1Btr2*	NA	NA	*Nud*	spring
HOR_7030	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_7000	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	NA	*Nud*	spring
HOR_1977	CA-SHS	landrace	low	*btr1Btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_16560	CA-SNS	landrace	low	*Btr1btr2*	NA	NA	*nud*	spring
BCC_554	CA-SNS	cultivar	low	*Btr1btr2*	NA	*vrs1.a1*	*nud*	spring
HOR_7598	CA-SNS	landrace	low	*btr1Btr2*	6-rowed	NA	*nud*	spring
HOR_7615	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_7630	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_2485	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a4*	*nud*	spring
HOR_13718	CA-SNS	landrace	low	*Btr1btr2*	NA	*vrs1.a1*	*nud*	spring

HOR_7710	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_7650	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_8588	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_13168	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a4*	*nud*	winter
HOR_7617	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_1534	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_7616	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	spring
HOR_7623	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_7556	CA-SNS	landrace	low	*Btr1btr2*	6-rowed	NA	*nud*	spring
HOR_11068	EA-SHM	cultivar	low	*Btr1btr2*	intermedium	NA	*Nud*	winter
HOR_11055	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_2673	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	spring
HOR_11690	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*Nud*	winter
HOR_10242	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	spring
HOR_11543	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_12395	EA-SHM	cultivar	low	*Btr1btr2*	NA	*vrs1.a4*	*Nud*	winter
HOR_11547	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_7262	EA-SHM	landrace	low	*Btr1btr2*	6-rowed	*vrs1.a3*	*Nud*	winter
HOR_12010	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	*vrs1.a1*	*nud*	winter
HOR_11924	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_11882	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	NA	*Nud*	winter
HOR_12060	EA-SHM	cultivar	low	*Btr1btr2*	6-rowed	*vrs1.a3*	*Nud*	winter

Table S3: Collection sites and sequencing statistics of 23 ancient samples.

sample	site	latitude	Longitude	*btr* genotype	*vrs1* genotype	coverage
TU1103	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.3
TU1107	Yoram Cave	31°18'48.7"N	35°21'15.5"E	*btr1Btr2*	*Vrs1.b3*	6.0
TU1109	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.3
TU1110	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.3
JK2281	Yoram Cave	31°18'48.7"N	35°21'15.5"E			1.0
JK3009	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.3
JK3010	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.4
JK3013	Yoram Cave	31°18'48.7"N	35°21'15.5"E			0.4
JK3014	Yoram Cave	31°18'48.7"N	35°21'15.5"E	*btr1Btr2*	*Vrs1.b2*	20.5
TU1111	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3
TU1112	Timna 34	29°46'08.6"N	34°57'04.4"E			0.2
TU1114	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3
TU1115	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3
TU1116	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3

TU1117	Timna 34	29°46'08.6"N	34°57'04.4"E	*btr1Btr2*	*Vrs1.b2*	5.9
TU1118	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3
TU1119	Timna 34	29°46'08.6"N	34°57'04.4"E			0.3
TU1120	Timna 34	29°46'08.6"N	34°57'04.4"E	*btr1Btr2*	*Vrs1.b2*	6.3
TU697	Timna 34	29°46'08.6"N	34°57'04.4"E	*btr1Btr2*	*Vrs1.b2*	3.9
TU1121	Abi'or Cave	31°52'17.7"N	35°25'53.0"E			0.3
TU1122	Abi'or Cave	31°52'17.7"N	35°25'53.0"E	*btr1Btr2*	*vrs1.a1*	8.0
TU1123	Abi'or Cave	31°52'17.7"N	35°25'53.0"E	*btr1Btr2*	*vrs1.a1*	7.4
TU1125	Abi'or Cave	31°52'17.7"N	35°25'53.0"E			0.3

Table S4: Passport data and GBS sequencing statistics of *H. agriocrithon* accessions in the IPK genebank.

accession	taxonomic	region	*eu/pseudo-agriocrithon*
HOR2268	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet	pseudo (*vrs1.a4*)
HOR2456	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	pseudo (*vrs1.a4*)
HOR2457	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	-
HOR2458	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	-
HOR2459	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	-
HOR2460	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	pseudo (*vrs1.a4*)
HOR2461	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	pseudo (*vrs1.a4*)
HOR2465	*Hordeum agriocrithon* A. E. Åberg var. *dawoense* A. E. Åberg	Tibet / Lhasa (Freisleben)	pseudo (*vrs1.a4*)
HOR2507	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	pseudo (*vrs1.a4*)
HOR2508	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	pseudo (*vrs1.a1*)
HOR2908	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	pseudo (*vrs1.a1*)
HOR3900	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a1*)
HOR3901	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	pseudo (*vrs1.a4*)
HOR3902	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a4*)
HOR3903	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a4*)
HOR3905	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a4*)
HOR3906	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a4*)
HOR3907	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	Tibet	eu (*vrs1.a4*)
HOR9719	*Hordeum agriocrithon* A. E. Åberg var. *paradoxon* Schiem	-	-
HOR1645	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet	pseudo (*vrs1.a4*)
HOR2452	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet / Tsela Dzong	pseudo (*vrs1.a1*)
HOR2453	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet / Tsela Dzong	pseudo (*vrs1.a1*)
HOR2455	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet / Tsela Dzong	-
HOR2462	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet / Tsela Dzong	-
HOR2463	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet	pseudo (*vrs1.a4*)
HOR2464	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Tibet / Tsela Dzong	pseudo (*vrs1.a1*)
HOR3552	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Kedma	eu (*vrs1.a1*)
HOR4414	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	N Negev / Beit Guvrin	-
HOR4904	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	Baku	eu (*vrs1.a1*)
HOR8509	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	-	eu (*vrs1.a1*)

HOR9517	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	-	pseudo (*vrs1.a4*)
HOR10286	*Hordeum agriocrithon* A. E. Åberg var. *agriocrithon* A. E. Åberg	-	-
HOR3886	*Hordeum agriocrithon* A. E. Åberg	-	eu (*vrs1.a1*)
HOR11029	*Hordeum agriocrithon* A. E. Åberg	-	-
HOR7269	*Hordeum agriocrithon* A. E. Åberg	-	eu (*vrs1.a1*)
HOR2653	*Hordeum agriocrithon* A. E. Åberg	Tibet	-
HOR3553	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	Beit Shonan Valley	-
HOR3887	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3888	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3889	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3890	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3891	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3892	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3893	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3894	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3895	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3896	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR3898	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4864	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	Manysh	-
HOR4865	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	Manysh	-
HOR4874	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4875	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4876	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4878	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4879	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4880	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4881	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4882	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4883	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4884	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4885	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4886	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4887	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4888	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4889	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4890	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4891	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4892	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4893	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4895	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4896	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4898	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4899	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4900	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4901	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR4902	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-
HOR9600	*Hordeum* x *lagunculiforme* (Bachteev) Bachteev ex Nikif.	-	-

Table S5: Passport data and WGS sequencing statistics of 17 *H. agriocrithon* accessions in the IPK genebank.

accessions	type	source	coverage
HOR_3900	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	14.9
HOR_3902	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	12.9
HOR_3903	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	14.7
HOR_3905	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	15.2
HOR_3906	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	14.8
HOR_3907	*eu-agriocrithon*	IPK genebank, var. *paradoxon*, Tibet	12.8
HOR_9719	mix-up accession	IPK genebank, var. *paradoxon*, LBY	12.2
HOR_3552	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, ISR	13.9
HOR_4904	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, AZE	13.9
HOR_8509	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, ISR	9.9
HOR_10286	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, UZB	12.5
HOR_7269	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, -	15.9
HOR_3553	*eu-agriocrithon*	IPK genebank, *lagunculiforme*, ISR	10.8
HOR_4874	*eu-agriocrithon*	IPK genebank, *lagunculiforme*, TKM	10.5
HOR_4879	*eu-agriocrithon*	IPK genebank, *lagunculiforme*, TKM	11.6
HOR_4896	*eu-agriocrithon*	IPK genebank, *lagunculiforme*, TKM	9.7
HOR_3886	*eu-agriocrithon*	IPK genebank, var. *agriocrithon*, TKM	11.9

10. Curriculum Vitae

Personal information

Name: Yu Guo
ORCID: https://orcid.org/0000-0002-5424-5797

Education

09.2007-06.2011 Master degree of Crop Breeding and Genetics

Master's thesis title: genetic diversity of hulless barley in Shangri-la region in China

Supervisor: Chunlin Long

College of Agriculture, Yunnan Agricultural University, Yunnan Province, China

09.2003-06.2007 Bachelor degree of Biology Science

Southwest Forestry University, Yunnan Province, China

Work experience

06.2019-present

PhD student, Domestication Genomics, IPK Gatersleben, Germany

09.2011-06.2019

Bioinformatics Engineer, BGI Genomics Co., Ltd. (Shenzhen, China)

Publication

1. **Guo Y**, Jayakodi M, Himmelbach A, Ben-Yosef E, Davidovich U, David M, Hartmann-Shenkman A, Kislev M, Fahima T, Schuenemann VJ, Reiter E, Krause

J, Steffenson BJ, Stein N, Weiss E, Mascher M. 2024. A haplotype-based evolutionary history of barley domestication. (**under review**)

2. Jayakodi M[#], Lu Q[#], Pidon H[#], Rabanus-Wallace MT[#], Bayer M, Lux T, **Guo Y**, Jaegle B, Badea A, Bekele W, Brar GS, Braune K, Bunk B, Chalmers KJ, Chapman B, Jørgensen ME, Feng JW, Feser M, Fiebig A, Gundlach H, Guo W, Haberer G, Hansson M, Himmelbach A, Hoffie I, Hoffie RE, Hu H, Isobe S, König P, Kale SM, Kamal N, Keeble-Gagnère G, Keller B, Knauft M, Koppolu R, Krattinger SG, Kumlehn J, Langridge P, Li C, Marone MP, Maurer A, Mayer KFX, Melzer M, Muehlbauer GJ, Murozuka E, Padmarasu S, Perovic D, Pillen K, Pin PA, Pozniak CJ, Ramsay L, Pedas PR, Rutten T, Sakuma S, Sato K, Schüler D, Schmutzer T, Scholz U, Schreiber M, Shirasawa K, Simpson C, Skadhauge B, Spannagl M, Steffenson BJ, Thomsen HC, Tibbits JF, Nielsen MTS, Trautewig C, Vequaud D, Voss C, Wang P, Waugh R, Westcott S, Rasmussen MW, Zhang R, Zhang XQ, Wicker T, Dockter C, Mascher M, Stein N. 2024. Structural variation in the pangenome of wild and domesticated barley. *Nature* 1-9.

3. Gao G, Yan L, Cai Y, **Guo Y**, Jiang C, He Q, Tasnim S, Feng Z, Liu J, Zhang J, Komatsuda T, Mascher M, Yang P. 2024. Most Tibetan weedy barleys originated via recombination between *Btr1* and *Btr2* in domesticated barley. *Plant Communications* 5(5), 100828.

4. **Guo Y**, Himmelbach A, Weiss E, Stein N, Mascher M. 2022. Six-rowed wild-growing barleys are hybrids of diverse origins. *The Plant Journal*, 111(3), 849-858.

5. Jiang C[#], Lei M[#], **Guo Y**[#], Gao G[#], Shi L, Jin Y, Cai Y, Himmelbach A, Zhou S, He Q, Yao X, Kan J, Haberer G, Duan F, Li L, Liu J, Zhang J, Spannagl M, Liu C, Stein N, Feng Z, Mascher M, Yang P. 2022. A reference-guided TILLING

by amplicon-sequencing platform supports forward and reverse genetics in barley. *Plant Communications* 3(4), 100317.

6. Koppolu R, Jiang G, Milner SG, Muqaddasi QH, Rutten T, Himmelbach A, **Guo Y**, Stein N, Mascher M, Schnurbusch T. 2022. The barley mutant multiflorus2.b reveals quantitative genetic variation for new spikelet architecture. *Theoretical and Applied Genetics* 135(2), 571-590.

7. Jayakodi M, Padmarasu S, Haberer G, Bonthala VS, Gundlach H, Monat C, Lux T, Kamal N, Lang D, Himmelbach A, Ens J, Zhang X-Q, Angessa TT, Zhou G, Tan C, Hill C, Wang P, Schreiber M, Boston LB, Plott C, Jenkins J, **Guo Y**, Fiebig A, Budak H, Xu D, Zhang J, Wang C, Grimwood J, Schmutz J, Guo G, Zhang G, Mochida K, Hirayama T, Sato K, Chalmers KJ, Langridge P, Waugh R, Pozniak CJ, Scholz U, Mayer KFX, Spannagl M, Li C, Mascher M, Stein N. 2020. The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* 588, 284–289.

8. Zeng X[#], **Guo Y**[#], Xu Q, Mascher M, Guo G, Li S, Mao L, Liu Q, Xia Z, Zhou J, Yuan H, Tai S, Wang Y, Wei Z, Song L, Zha S, Li S, Tang Y, Bai L, Zhuang Z, He W, Zhao S, Fang X, Gao Q, Yin Y, Wang J, Yang H, Zhang J, Henry RJ, Stein N, Tashi N. 2018. Origin and evolution of qingke barley in Tibet. *Nature Communications* 9(1), 5433.

9. **Guo Y**, Li Y, Huang Y, Jarvis D, Sato K, Kato K, Tsuyuzaki H, Chen L, Long C. 2012. Genetic diversity analysis of hulless barley from Shangri-la region revealed by SSR and AFLP markers. *Genetic Resources and Crop Evolution* 59, 1543-1552.

11. Eidesstattliche Erklärung/Declaration under Oath

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

_____ _____

Datum / Date

Unterschrift des Antragstellers / Signature of the applicant

12. Erklärung über bestehende Vorstrafen und anhängige Ermittlungsverfahren / Declaration concerning Criminal Record and Pending Investigations

Hiermit erkläre ich, dass ich weder vorbestraft bin noch dass gegen mich Ermittlungsverfahren anhängig sind.

I hereby declare that I have no criminal record and that no preliminary investigations are pending against me.

_____	_____
Datum / Date	Unterschrift des Antragstellers / Signature of the applicant