the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2022) **111,** 849–858 doi: 10.1111/tpj.15861

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

Yu Guo[1] , Axel Himmelbach[1] , Ehud Weiss[2] , Nils Stein[1,3] and Martin Mascher[1,4,*]

[1]*Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany,*

[2]*The Martin (Szusz) Department of Land of Israel Studies and Archaeology, Bar-Ilan University, Ramat-Gan, Israel,*

[3]*Center for Integrated Breeding Research (CiBreed), Georg-August-University Göttingen, Göttingen, Germany, and*

[4]*German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany*

Received 19 April 2022; revised 25 May 2022; accepted 30 May 2022.; published online 9 June 2022.

*For correspondence (e-mail mascher@ipk-gatersleben.de).

SUMMARY

Crop–wild gene flow is common when domesticated plants and their wild relatives grow close to each other. The resultant hybrid forms appear as semi-domesticates and were sometimes considered as missing links between crops and their wild progenitors. Wild-growing barleys in Central and Eastern Asia, named *Hordeum agriocrithon*, show hallmark characters of both wild and domesticated forms. Their spikes disintegrate at maturity to disperse without human intervention, but bear lateral grains, which were favored by early farmers and are absent from other wild barleys. As an intermediate form, *H. agriocrithon* has been proposed several times as a progenitor of domesticated barley. Here, we used genome-wide marker data and whole-genome resequencing to show that all *H. agriocrithon* accessions of a major germplasm collection are hybrid forms that arose multiple times by admixture of diverse domesticated and wild populations. Although *H. agriocrithon* barleys have not played a special role in barley domestication, future analysis of the adaptative potential of bi-directional crop–wild gene flow in extant barleys may prove a fertile research field.

Keywords: barley (*Hordeum vulgare* L.), hybridization, gene flow, domestication, population genomics, plant genetic resources.

INTRODUCTION

All crop species were domesticated from wild plants. Many domesticates are interfertile with the extant wild relatives (Harlan & de Wet, 1971) and genetic and ecological studies have found ample evidence for natural hybridization between conspecific wild and domesticated plants in regions of sympatry (Ellstrand et al., 2013; Janzen et al., 2019). Barley (*Hordeum vulgare)* is an important crop that is grown globally in temperate regions. Its conspecific wild progenitor is *H. vulgare* subsp. s*pontaneum* (henceforth *H. spontaneum*), a grass native to Western Asia (Harlan & Zohary, 1966; Zohary et al., 2012). Archeological evidence has firmly established barley cultivation as co-eval with the onset of agriculture (Asouti & Fuller, 2012; Nadel et al., 2012). The two most important characters distinguishing domesticated from wild barleys are non-shattering (or 'non-brittle') and six-rowed spikes. In wild barley, the central axis of the spikes (the rachis) disintegrates at maturity into individual dispersal units, the spikelet triplets. These triplets are composed of one central and two lateral spikelets. In wild barley, only the central spikelets bear grains, while the lateral ones grow into infertile rudiments. In six-rowed barleys, the lateral spikelets are fertile, increasing yield potential. Hence, six-rowed types spread around the world and are found in geographically and genetically diverse barley populations (Milner et al., 2019).

Hordeum spontaneum is two-rowed, but there are reports about six-rowed barleys that grow 'wild', meaning that they thrive without human support. It has been a matter of debate whether these can be considered 'wild' in the sense of being descended from *H. spontaneum* populations, rather than being crop–wild hybrids. Six-rowed wild barleys were noticed by European collection missions to East Asia, who named them *H. agriocrithon* (Åberg, 1938) and considered them the forebears of all domesticated barleys. The argument in favor of this hypothesis relied on the seemingly implausible acquisition of grain-bearing lateral spikelets by two-rowed barleys. Schiemann (1932) deemed it much more likely that two-rowed types were derived from six-rowed progenitors by loss of lateral spikelet fertility.

Later research disproved this hypothesis. Mutation breeding in the 1960s lifted the conceptual hurdles of

© 2022 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License,
which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

conceiving how six-rowed barleys could have emerged from two-rowed ones. Many complementation groups of six-rowed mutants were found after irradiation of two-rowed barleys (Lundqvist, 2014; Lundqvist et al., 1989), suggesting the mutational target for converting two-rowed into six-rowed types is large. The molecular basis of the non-brittle and six-rowed traits has been unraveled through positional cloning. Komatsuda et al. (2007) showed that all six-rowed barleys carry loss-of-function mutations in the *VRS1* homeobox domain transcription factor. Non-brittle barleys arose at least two times by independent knock-out of two closely linked genes, *NON-BRITTLE RACHIS 1* (*BTR1*) and *NON-BRITTLE RACHIS 2* (*BTR2*) (Pourkheirandish et al., 2015). With this knowledge we can conceive three possible ways of how six-rowed wild-growing barleys may have come about. The first is mutation: six-rowed types can spring up by spontaneous mutation of *SIX-ROWED SPIKE 1* (*VRS1*) and possibly other genes in *H. spontaneum* populations without any contact to domesticated barley. The second is introgression by which a domesticated *vrs1* knock-out allele is introduced into a wild genomic background. The third route is what we term de-domestication by recombination. In crosses between domesticated barleys with allelic combinations *btr1Btr2* and *Btr1btr2*, rare cross-overs yield recombinant progeny with functional alleles *Btr1Btr2*, giving rise to spontaneously 'de-domesticated' barley. If this happens in a six-rowed background, a barley with brittle spikes and fertile lateral florets emerges.

By resequencing of *BTR1/2* and *VRS1*, Pourkheirandish et al. (2018) found evidence that all three scenarios may have played a role in the origin of *H. agriocrithon* and defined the classes *eu-agriocrithon* and *pseudo-agriocrithon*. The latter group, mainly composed of six-rowed wild-growing barleys from Tibet, had alleles of *BTR1* and *BTR2* otherwise found only in domesticated barley. Hence, these populations must have arisen either by recombination or by introgression. They were considered 'falsely' wild and were named '*pseudo-agriocrithon*'. By contrast, *eu-agriocrithon* barleys from Central Asia had *Btr1Btr2* haplotypes also found in *H. spontaneum* and were considered as descendants of a six-rowed wild barley population from which extant six-rowed barleys had been domesticated. The latter hypothesis' implication that six-rowed barley predates non-shattering barley has piqued our interest. Here, we investigated it further by population genomic analyses.

RESULTS

We selected 77 accessions from the collection of the German Federal *ex situ* genebank hosted at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben (Table S1) for genome-wide genotyping. Of these, 17 were designated as *H. agriocrithon* var. *agriocrithon* or only *H. agriocrithon* (both henceforth referred to as var. *agriocrithon*), eight as *H. agriocrithon* var. *dawoense* (var. *dawoense*), 11 as *H. agriocrithon* var. *paradoxon* (var. *paradoxon*), and 41 as *Hordeum* × *lagunculiforme* (*lagunculiforme*, 'bottle-shaped' barley, a potential crop–wild hybrid; Hammer, 1984). Some of these accessions trace back to the very seed stocks upon which Åberg and Schiemann had based their theories concerning the origin of domesticated barley (Åberg, 1938; Schiemann, 1951). We refer to our set of 77 accessions collectively as 'six-rowed wild-growing barleys'. A single plant from each accession was analyzed by genotyping-by-sequencing (GBS; Wendler et al., 2014). A principal component analysis (PCA) on a marker matrix of these 77 plants combined with a set of globally diverse barleys (Table S2) composed of 389 wild and 400 domesticated forms (Milner et al., 2019; Pourkheirandish et al., 2015; Sallam et al., 2017) revealed several things (Figure 1). First, apart from *lagunculiforme*, all six-rowed wild-growing barleys occupied an intermediate position between two-rowed wild and domesticated barleys or clustered with domesticated forms. The odd placement of *lagunculiforme* is not attributable to a large divergence of that group from all other barleys, but to a PCA artifact (Elhaik, 2021) from over-sampling. An identity-by-state (IBS) showed the *lagunculiforme*s to be genetically near-identical (Figure S1). A PCA with fewer accessions of this taxon placed them in an intermediate position between *H. spontaneum* and Eastern domesticated barley (Figure S2; Table S3). Other than *lagunculiforme*, none of the accessions sharing the same taxonomic designations formed obvious clusters (Figure 1b), reinforcing the criticism by Harlan and de Wet (1971) of formal taxonomy at the infraspecific level. A 'haplotype taxonomy' was easier to interpret. Twenty-six of the six-rowed wild-growing barleys had been analyzed by Pourkheirandish et al. (2018) (Table S1). Our PCA placed 15 *pseudo-agriocrithons* into the domesticated cluster, consistent with the idea that *pseudo-agriocrithons* are brittle rachis revertants that owe their existence to recombination of two domesticates. By contrast, all but one *eu-agriocrithon* (HOR 3900) mapped in between wild and domesticated types or on the fringes of the wild cluster in the PCA (Figure 1a).

The hypothesis of Pourkheirandish et al. (2018) posits that *eu-agriocrithons* are related to the six-rowed wild progenitors to which extant six-rowed domesticated forms trace back. The intermediate position of the *eu-agriocrithons* may thus reflect their close relatedness to certain domesticates or, alternatively, a more recent hybrid origin. To elucidate the matter, we selected 17 accessions (Figure 1d) for whole-genome shotgun sequencing (WGS), among them all *eu-agriocrithons* as per Pourkheirandish et al. (2018) and selected *lagunculiforme* accessions (Figure S3). We analyzed these sequences together with

© 2022 The Authors.

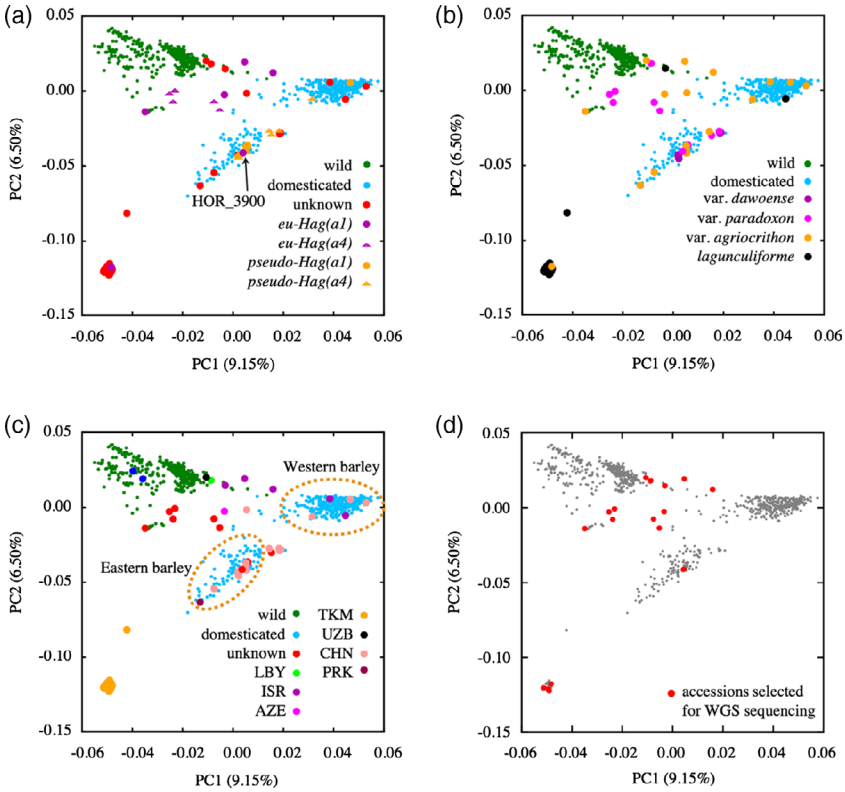


Figure 1. Positions of *H. agriocrithon* accessions in the global barley diversity space.
Principal component analysis (PCA) of *H. agriocrithon* and global diversity panels of domesticated and wild (*H. spontaneum*) barley. The PCA was computed from genome-wide marker data obtained by genotyping-by-sequencing. 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. (15) based on haplotypes at domestication genes. (b) *Hordeum agriocrithon* accessions were colored according to infraspecific taxonomy. (c) *Hordeum agriocrithon* accessions were colored according to their countries of origin. (d) *Eu-agriocrithon* selected for WGS sequencing.

WGS data (Table S4) for a set of 99 wild barleys, 195 domesticated barleys (Jayakodi et al., 2020); three wild-growing barleys from Tibet (Zeng et al., 2018) classified as *pseudo-agriocrithons* in the sense of Pourkheirandish et al. (2018); and a wild barley accession carrying a haplotype closely related to *vrs1.a4* of certain six-rowed barleys (Wang et al., 2021). A neighbor-joining (NJ) tree and a PCA computed from the WGS variant matrix revealed similar patterns as the PCA from GBS data: most accessions sat in between wild and domesticated clusters (Figure 2a,b). The *pseudo-agriocrithons* of Zeng et al. (2018) appeared to be closely related to domesticated barley from East and Central Asia, as did the purported wild barleys CDB0083 and HOR 3900. A model-based ancestry analysis with ADMIXTURE (Alexander et al., 2009) confirmed and refined these observations (Figure 2c). When four ancestral populations ($k = 4$) were used, wild barley split into groups broadly representing accessions from Western Asia (WWA, wild barley subpopulation mainly distributed in Western Asia) and Central Asia (WCA, wild barley subpopulation mainly distributed in Central Asia, Figure S4). Domesticated barley likewise split according to geography into ancestral populations of Western and Eastern origin (WD and ED). This large-scale population structure agrees with previous population genomic studies (Milner et al., 2019; Morrell & Clegg, 2007; Russell et al., 2016). Most six-rowed wild-growing barleys had indeed both wild and domesticated ancestry (Figure 2c). Those that did not – three *pseudo-agriocrithons*, HOR 3900, and CDB0083 – were of Eastern domesticated ancestry. The others had variable proportions of ancestry from wild and domesticated populations. Intriguingly, present-day collection sites of inferred hybrids were broadly co-extensive with their respective ancestral populations (Figure 2c): crop–wild hybrids from Tibet had ED and WCA ancestry, while Israeli accessions combined WD and WWA ancestry. Hybrids from the Caucasus and Central Asia had mixtures of Eastern and Western ancestries. For example, HOR 4904 from Azerbaijan had all four components in more or less equal proportions. As with the PCA, taxonomic names did not correlate with genomic ancestry.

Our admixture analysis has established the hybrid nature not only of the *pseudo-agriocrithon* but also of the *eu-agriocrithon* types among our panel of six-rowed

© 2022 The Authors.

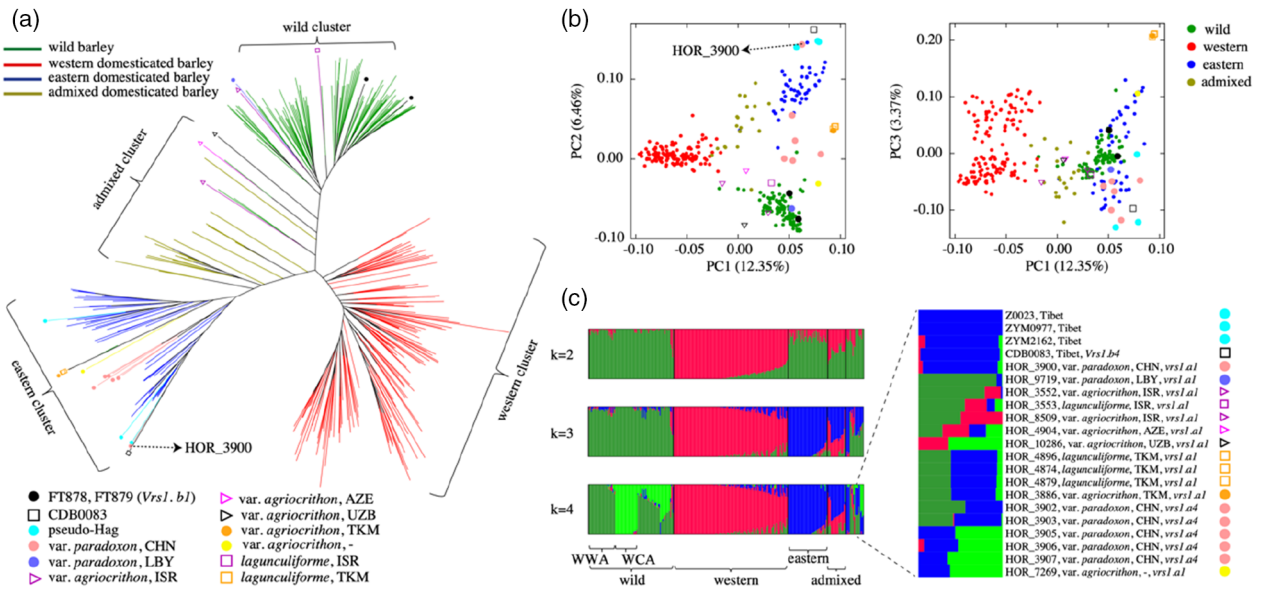


Figure 2. Genomic ancestry of *H. agriocrithon*.
(a) Neighbor-joining tree based on genetic distance computed from the whole-genome variant matrix. (b) Principal component analysis (PCA) showing PC1 versus PC2 (left) and PC1 versus PC3 (right). The proportion of variance explained by the PCs is given in the axis labels. (c) Individual ancestry coefficients in ADMIXTURE with the number of ancestral populations (*k*) ranging from two to four. 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 in (a) and also used in (b) and (c).

wild-growing barleys. But it did not tell us which parts exactly of each hybrid's genome trace back to a wild and which to a domesticated forebear. To find this out we selected ancestry-informative markers distinguishing, by means of strong allele frequency differentials, between pairs of ancestral populations and plotted graphical genotypes for these markers in our hybrids (Figure 3). The plots were concordant with ADMIXTURE in that the genomes of accessions predicted to carry two ancestry components were indeed painted in their respective colors. In addition, individual genomes showed differently patterned mosaics, reminiscent of recombinant inbred lines, each with their private cross-over sites, although some accessions shared breakpoints. To re-create such a pattern in the absence of crop–wild gene flow, six-rowed wild barleys must have mimicked by random forces and throughout the genome the allele frequency trajectories of a domesticated, i.e., strongly selected, population – an unreasonable assumption. Their mosaic genomes support the idea that six-rowed crop–wild hybrids trace back to multiple hybridization events or at least represent different lineages issuing from a past hybrid ancestor they had in common. Multiple origins are also consistent with the topology of the NJ tree and indicate that *eu-agriocrithons* are not a discrete population, which makes the case for their being an ancestral population problematic.

We wondered whether these observations can be brought in line with those of Pourkheirandish et al. (2018). The comparison of *VRS1* and *BTR1/2* haplotypes from their Sanger sequence and our WGS data were in keeping with each other (Tables S5 and S6). All accessions classified as *eu-agriocrithon* by Pourkheirandish et al. (2018) had wild *Btr1Btr2* haplotypes and domesticated *vrs1* (i.e., six-rowed) haplotypes. One accession (HOR 9719) collected from Libya designed as var. *agriocrithon* carried a cultivated *Btr1btr2* haplotype. This accession may be heterogeneous as other plants shattered and were two-rowed. Our genome-wide data made it possible to extend our analysis to flanking regions of the genes (Figure 4a,b). In agreement with previous studies we found in domesticated barleys two major haplotypes distinguishing *btr1* or *btr2* carriers and roughly correlating with Eastern or Western origin. Haplotypes in wild barleys were more diverse, but also sorted into two main clades somewhat congruous with the globally defined WWA and WCA populations (Figure 4b). Consistent with their placement in the PCA, the three *pseudo-agriocrithons* of Zeng et al. (2018), CDB0083, and HOR 3900 carried domesticated haplotypes, but had shed their recessive *btr1* and *btr2* genes owing to recombination between them. Interestingly, we also found among the 195 domesticated barleys of Jayakodi et al. (2020) a double recessive *btr1btr2* recombinant (HOR 15401 and HOR 2598), which we are, to our knowledge, the first to report. This shows that natural recombinant populations have been large enough to reshuffle alleles of *BTR1* and *BTR2* by recombination despite the two genes' close linkage, supporting the de-domestication by recombination origin of *pseudo-agriocrithon* put forward by Pourkheirandish et al. (2018).

© 2022 The Authors.

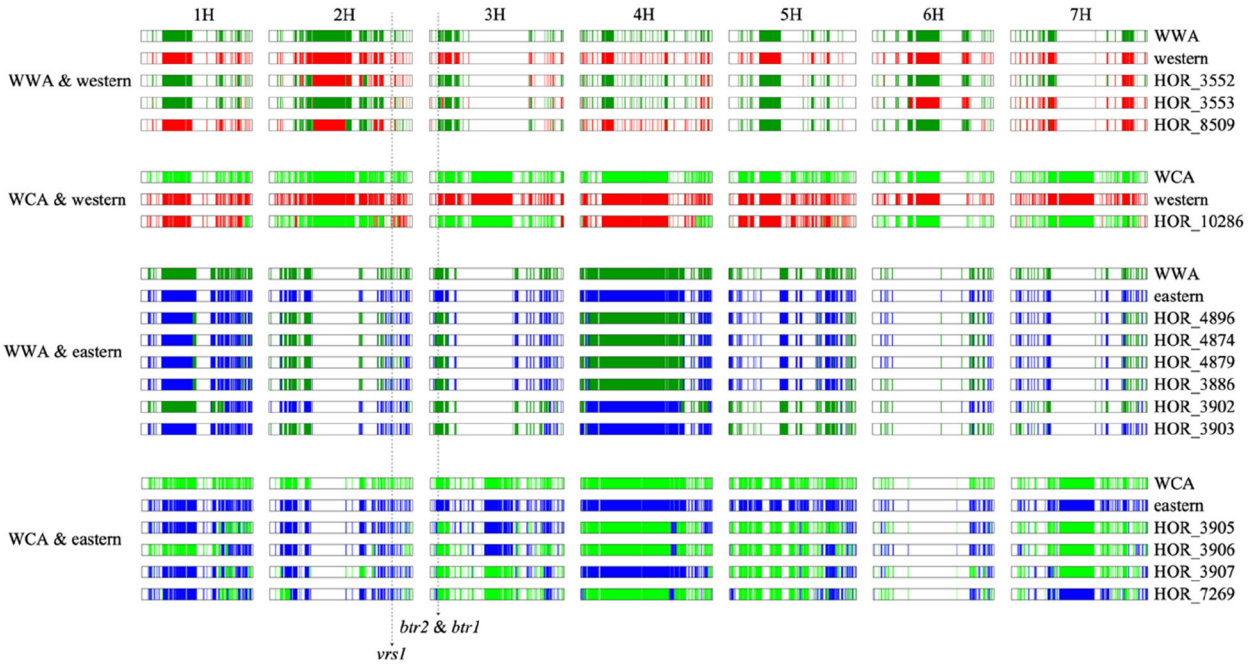


Figure 3. Graphical genotypes of ancestry-informative markers in crop–wild hybrids.
Color codes and labels for *H. agriocrithon* taxa are as in Figure 2. *Hordeum agriocrithon* accessions were grouped according to their two major ancestral populations as determined by ADMIXTURE with $k = 4$ (Figure 2c): (a) WWA and Western domesticated barley; (b) WCA and Western domesticated barley; (c) WWA and Eastern domesticated barley; (d) WCA and Eastern domesticated barley. The first two tracks in each panel show the positions of ancestry-informative markers (i.e., those with a high allele frequency differential between populations) and define the color codes. The position of the *BTR1/BTR2* and *VRS1* loci is marked by a dashed line. The accession HOR 10286 was excluded because it had substantial ancestry from more than two populations.

In our haplotype analysis at *VRS1*, we focused on *vrs1.a1* alleles and closely related *Vrs1.b1* in two-rowed types. Pourkheirandish et al. (2018) speculated that *eu-agriocrithon* barleys with *vrs1.a1* and *vrs1.a4* alleles represent an intermediate form between wild barley and six-rowed domesticated barleys. We did not focus on *vrs1.a4* because there were few carriers in our panel and no closely related wild haplotype ('*Vrs1.b4*'). The haplotype of CDB0083 had been designated as *Vrs1.b4*, but that accession came out as a crop–wild hybrid in our analysis. In total, we had 95 *vrs1.a1* carriers in domesticated barley and 10 among the *eu-agriocrithons*; *Vrs1.b1* was found in two wild barleys (Figure S5). Among the *vrs1.a1* genotypes, 74 domesticated barleys and nine *eu-agriocrithon* accessions shared a 690-kb-long region spanning *VRS1* and unbroken by recombination (Figure 5a). An NJ tree constructed from 71 773 single-nucleotide polymorphisms (SNPs) in this region placed crop–wild hybrids in derived positions nested within domesticated branches (Figure 5b), arguing against their ancestral role in barley domestication. If our *eu-agriocrithon* accessions had not crossed with domesticated barley after the latter arose, they should have split earlier from domesticated forms (Figure S6).

A similar argument derives from the distribution of genetic distances calculated from SNPs in the vicinity of *VRS1* in carriers of *Vrs1.b1* and *vrs1.a1* (Figure 5c; Table S7). The *Vrs.b1* haplotype occurs in present-day wild barleys. As such, these are not the ancestors of six-rowed domesticated barley, which must have existed 8000–10 000 years ago. Subsequently and in the absence of crop–wild recombination, the lineage leading to today's *Vrs.b1* carriers accumulated mutations independently of domesticated barley and the *Vrs.b1* haplotype we found in our sample is not ancestral, but only more closely related to *vrs1.a1* than are other wild *Vrs1* haplotypes. If *eu-agriocrithon* barleys were pristinely wild and had not come in contact with cultivated forms, they should have amassed, like *Vrs1.b1* carriers, mutations absent from domesticated barley. These would have increased the genetic distance between the latter taxon and *eu-agriocrithon*. But we see nothing of the kind. *Eu-agriocrithons* are no further away from domesticated six-rowed types than these are to each other (Figure 5c).

Finally, we computed f_4-statistics to test for admixture by comparing shared derived alleles. We considered the following four populations: the outgroup *Hordeum pubiflorum*, wild barley (unadmixed *H. spontaneum*), six-rowed domesticated carriers of *vrs1.a1*, and *eu-agriocrithons* with *vrs1.a1*. The quartets (outgroup, wild; *eu-agriocrithon*, domesticated) and (outgroup, domesticated; *eu-agriocrithon*, wild) both had significantly non-zero f_4 values (Table S8),

© 2022 The Authors.

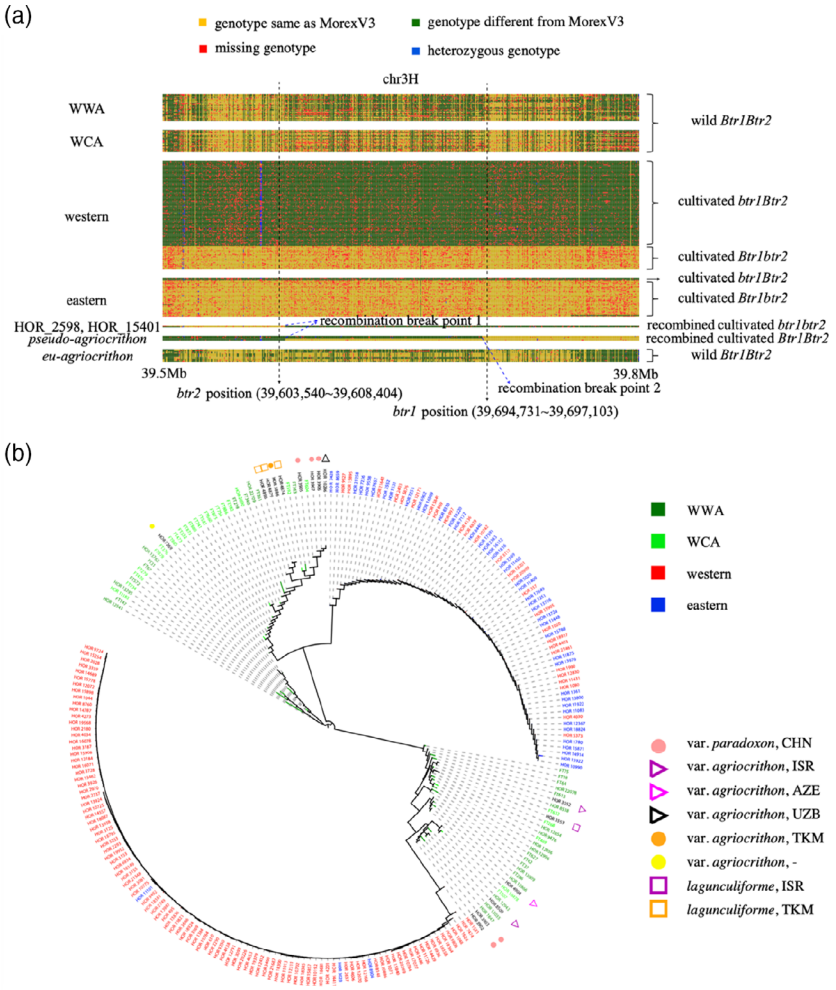


Figure 4. Genetic relationship analysis based on sequence variation in and around the *BTR1* and *BTR2* genes.
(a) SNP haplotypes around the *BTR1* and *BTR2* genes (39.5–39.8 Mb in MorexV3 3H). Genotype calls were color-coded as specified. The *pseudo-agriocrithon* accessions are (from top to bottom) Z0023, ZYM0977, ZYM2162, CDB0083, and HOR 3900. The admixed wild and domesticated barley accessions were not used. We identified the following recombinants: HOR 2598 and HOR 15401 carried *btr1btr2* and CDB0083 and HOR 3900 carried *Btr1Btr2*. (b) Neighbor-joining tree based on the SNPs between the *btr1* and *btr2* genes (10 149 SNPs in the interval 39 603 540 bp to 39 697 103 bp in MorexV3 3H were used). *Pseudo-agriocrithons* and recombinants were excluded.

consistent with a post-domestication hybrid origin of *eu-agriocrithon* (Figure S7).

DISCUSSION

Our data argue against the hypothesis of Pourkheirandish et al. (2018) that genebank accessions classified as *H. agriocrithon* are descendants of wild barley populations from which six-rowed domesticated barleys originated in the early Neolithic. Rather, these accessions are crop–wild hybrids of more recent stock and diverse provenance. Our findings agree with those of many other studies using sparser molecular marker data and comparisons of genetic diversity within *H. spontaneum* and *H. agriocrithon* (Murphy et al., 1982; Tanno & Takeda, 2004). We arrive at our conclusions from a dataset that is consistent with that of Pourkheirandish et al. (2018), who had taken the presence of wild haplotypes at the *BTR1/2* locus and domesticated alleles at *VRS1* in *eu-agriocrithon* barleys as evidence for the ancestral status of that taxon to six-rowed domesticated barleys. But beyond these two loci, our ancestry analysis with genome-wide markers showed that the presence of *vrs1* haplotypes in wild-growing barleys is only an instance of more pervasive domesticated ancestry, which is in keeping with past hybridization between members of domesticated and wild populations that had already diverged from each other.

In another important aspect, our data support a key finding of Pourkheirandish et al. (2018): domesticated barley can re-acquire its ancestral grain dispersal system by recombination-mediated elimination of recessive

© 2022 The Authors.

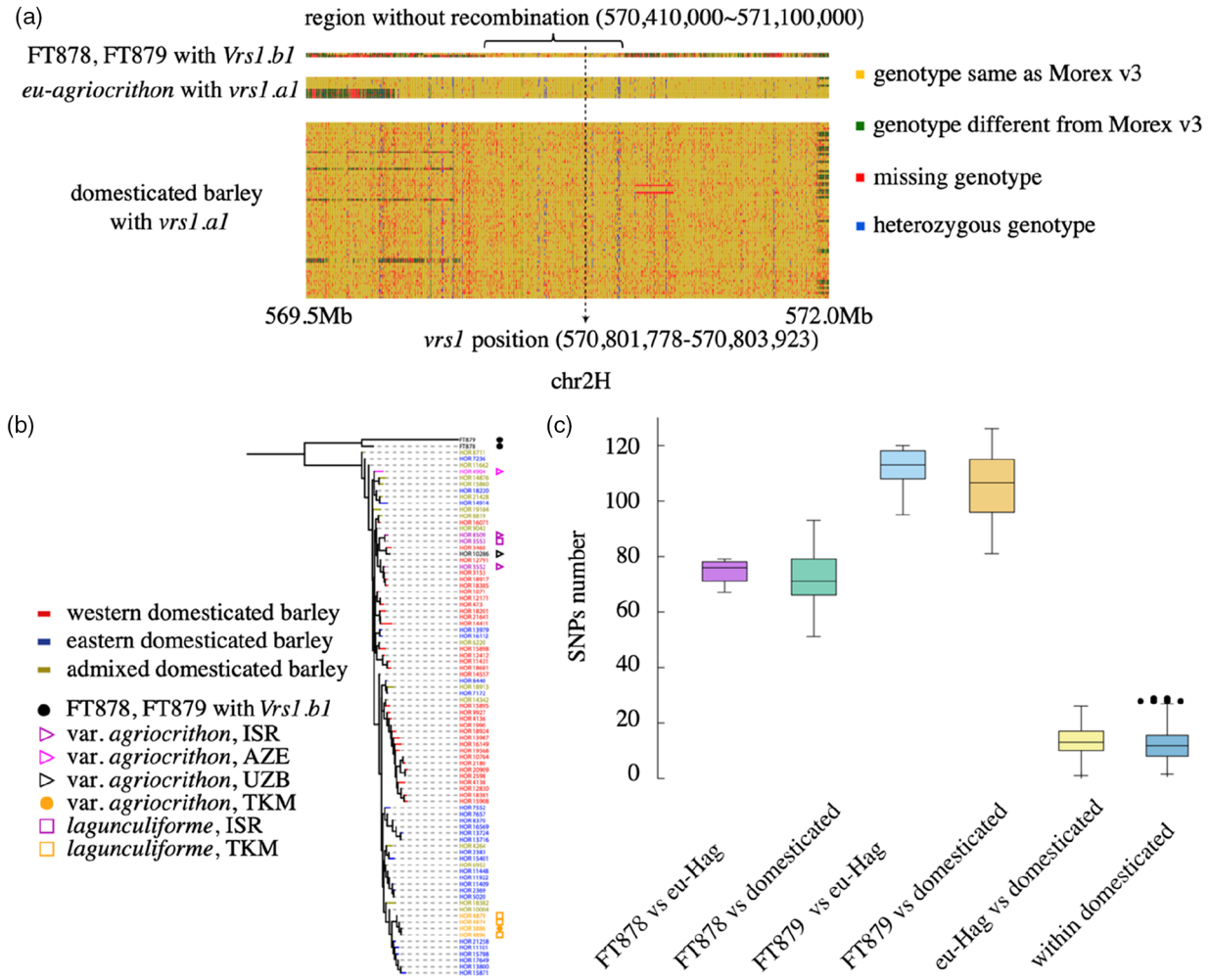


Figure 5. Genetic relationship of *vrs1.a1* carriers.
(a) Haplotypes of wild barley FT878 and FT879 (carriers of *Vrs1.b1*) and *vrs1.a1* carriers without nearby recombination events among *eu-agriocrithon* and domesticated barley in the vicinity of *VRS1* (569.5–572.0 Mb in MorexV3 2H). (b) Neighbor-joining tree based on the 71 014 SNPs in the interval 570 410 000 bp to 571 100 000 bp in MorexV3 2H (marked in (a)). (c) Number of genome-wide pairwise comparisons between *Vrs1.b1* carriers (FT878, FT879), *eu-agriocrithon*, and domesticated barley. The sequence data of *eu-agriocrithon* was down-sampled to 3× coverage to match the read depth in other samples for the analyses shown in (b) and (c).

loss-of-function alleles after hybridization between *btr1* and *btr2* carriers. This de-domestication by recombination scenario is supported by genome-wide ancestry in that the *pseudo-agriocrithon* have no wild ancestry. The concepts of *eu-agriocrithon* and *pseudo-agriocrithon* may provide a more robust framework for classifying wild-growing six-rowed barleys than traditional infraspecific taxonomy, which did not align with genomic ancestry.

We want to stress that what we found is that *extant* genebank accessions are not derivatives of the ancestral populations. This does not rule out the more general hypothesis of Pourkheirandish et al. (2018) that six-rowed barley may be as old or even older than non-shattering barley. Many archeologists argue for a protracted domestication process spanning one or two millennia during which *btr1/2* and *vrs1* alleles could conceivably both have been segregating in early cultigens (Purugganan, 2019). Our sequence data are not informative about allele frequencies in the distant past. However, we note that six-rowed barleys appear only after non-brittle types in the archeological record: the earliest remains of two-rowed domesticated barleys date back to the first half of the 8th millennium BCE (Hopf, 1983), while those of six-rowed types from Çatalhöyük are no older than 9350 cal BP (Helbaek, 1964).

Ever since Western botanists came across wild-growing six-rowed barley in East and Central Asia, this suggestively named taxon (ἄγριος – wild, κριθή,– barley) has beguiled barley geneticists. Most of the ensuing research endeavors were variations on the theme set by Åberg and focused on the question whether *H. agriocrithon* is ancestral to

© 2022 The Authors.

domesticated barley. The genetic evidence points in a different direction: domesticated barley has mosaic ancestry from different wild barleys in the Fertile Crescent (Poets et al., 2015); singling out extant descendants of a hypothetical source population is not possible; *H. agriocrithon* is not an ancestral form.

Still we believe barley geneticists can learn a lot from *H. agriocrithon*. Crop–wild hybrids between *H. spontaneum* and two-rowed barleys, though less conspicuous, may be no less abundant than *H. agriocrithon* (Helback, 1959). Genotyping barley growing wild next to farmers' fields may tell us how common two-rowed crop–wilds are. Six-rowed wild-growing barleys have been reported as ruderals and weeds (Kamm, 1954) or in recently colonized habitats such as the Tibetan Plateau. Do higher fertility and possible pleiotropic changes to vegetative growth habit (Thirulogachandar et al., 2017) confer fitness advantages over two-rowed types under these scenarios? Despite the large mutational target for altering row type, no unadmixed six-rowed populations of *H. spontaneum* are known. Neither are six-rowed, brittle-rachis barleys found on the archeobotanical record. This absence has been attributed to the disruption of the wedge-shaped dispersal unit in the presence of lateral grains (Komatsuda et al., 2007) and the attendant loss of fitness, a hypothesis that may be put to the test in common gardens of induced six-rowed *vrs1* mutants and their *H. spontaneum* parent (Pourkheirandish et al., 2015).

METHODS

Genotyping-by-sequencing

A total of 77 accessions from the IPK genebank (Oppermann et al., 2015) were genotyped by GBS. DNA was extracted from seedling leaves essentially according to the protocol of Milner et al. (2019), 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. We also used public GBS data for 389 wild barleys and 400 domesticated (Table S2) barleys reported by Sallam et al. (2017) and Milner et al. (2019). The 400 domesticated barley accessions were selected from approximately 20 000 domesticated barley (8) to represent the domesticated barley diversity (Figure S8), keeping a balance between the numbers of wild and domesticated accessions. GBS reads were mapped to the MorexV3 pseudomolecules (Mascher et al., 2021) using BWA-MEM (Li, 2013). Variant calling was done with bcftools (Danecek et al., 2021) using the command 'mpileup -q 20 -Q20'. The resultant variant matrix was filtered as follows: (i) only bi-allelic SNP sites were kept; (ii) homozygous genotype calls were retained if their read depth was ≥2 and ≤50 and set to missing otherwise; (iii) heterozygous genotype calls were retained if the read depth of both alleles was ≥2 and set to missing otherwise. SNP sites with fewer than 20% missing calls and fewer than 20% heterozygous calls were used in the subsequent analysis. PCA was done with smartpca version 7.2.1 (Patterson et al., 2006). Only variants with a minor frequency of ≥5% were used in the PCA.

Whole-genome shotgun sequencing

A subset of 17 of the 77 accessions genotyped by GBS was used for WGS. Illumina Nextera libraries were prepared from the same DNA used for GBS and sequenced on an Illumina NovaSeq 6000 at IPK Gatersleben. Public WGS data of 298 wild and domesticated accessions were included in the analysis (Table S4). Read mapping, variant calling, and filtration were done as described for the GBS data with the following modifications: (i) '--excl-flags 3332' was added to parameter for 'bcftools mpileup' to exclude duplicates and secondary alignments; (ii) a minimum read depth of 1× was required to accept homozygous genotype calls. A matrix of pairwise genetic distances based on IBS was computed with Plink V1.90 (Purcell et al., 2007). The matrix was used to construct an NJ tree with Fneighbor (http://emboss.toulouse.inra.fr/cgi-bin/emboss/fneighbor) in the EMBOSS package (Rice et al., 2000). The tree was visualized with Interactive Tree Of Life (iTOL) (Letunic & Bork, 2021). PCA was done with Eigensoft version 7.2.1 (Patterson et al., 2006). Population structure was analyzed with ADMIXTURE version 1.23 (Alexander et al., 2009). Prior to ADMIXTURE runs, linkage disequilibrium pruning was done with Plink using the parameters '--indep-pairwise 50 10 0.1'. For each value of from 2 to 4, 50 replicate runs of ADMIXTURE were combined with CLUMPP version 1.1.2 (Jakobsson & Rosenberg, 2007) and plotted with Distruct version 1.1 (Rosenberg, 2004). Individuals with >80% of WWA or WCA ancestry were considered as unadmixed wild barleys.

Graphical genotypes for ancestry-informative markers

We selected ancestry-informative markers that differentiate between pairs of ancestral populations defined in our ADMIXTURE analysis with $k = 4$. For the purpose of calculating allele frequencies in unadmixed individuals, a genotype was assigned to an ancestral population if it had ≥80% ancestry from that population. For each population, we selected SNPs sites with a major allele frequency of ≥90% among its members. For pairwise comparison, we selected SNP sites shared between two populations but with opposite major alleles. Crop–wild hybrids were assigned to pairs of ancestral populations according to their two largest ancestry components. For visualization, a sliding genomic window (size: 100 kb, shift: 20 kb) was assigned to an ancestral population if ≥80% of SNPs in it supported that population. Only windows with at least 10 SNPs were considered.

© 2022 The Authors.

Haplotype analysis at *BTR1/2* and *VRS1*

MorexV3 gene models (Mascher et al., 2021) were found by BLAST alignment (McGinnis & Madden, 2004) of the sequences of *BTR1/2* (Pourkheirandish et al., 2015) and *VRS1* (Komatsuda et al., 2007) to the MorexV3 pseudomolecules (Table S9). Only variants with a minor allele frequency of ≥10% were used for visualization (Figure 4a) using Gnuplot (http://www.gnuplot.info/). The complete WGS variant matrix filtered as described above was used to construct an NJ tree with MEGA11 (Tamura et al., 2021) using 100 bootstrap replicates. iTOL (Letunic & Bork, 2021) was used for visualization. Pairwise genetic distances were computed from the complete WGS matrix. We had sequenced *eu-agriocrithon* accessions to an approximately threefold higher coverage than other wild and domesticated accessions (20). The latter consequently had more missing data, which could affect phylogenetic inferences (Figure S9). To correct for this, we repeated mapping and SNP calling with down-sampled sequence data *eu-agriocrithon* (to approximately 3× coverage). We also added public data for one *H. pubiflorum* genotype (Mascher et al., 2013) to the SNP matrix for use as outgroup in the computation of the f_4-statistics. The NJ trees in Figures 4b and 5b and the SNP distance in Figure 5c were based on the down-sampled SNP matrix.

f_4-Statistics

Four populations (*H. pubiflorum*, wild barley, six-rowed domesticated barley with *vrs1.a1*, and *eu-agriocrithon* barleys with *vrs1.a1*) were used. For wild barley, only unadmixed accessions (WWA and WCA) were used. We calculated f_4-statistics for two quartets: f_4(*H. pubiflorum*, wild; *eu-agriocrithon*, domesticated) and f_4(*H. pubiflorum*, domesticated; *eu-agriocrithon*, wild) using the qpDstat program of the ADMIXTOOLS (Patterson et al., 2012) package with the 'f4mode: YES' option.

ACKNOWLEDGMENTS

We thank Susanne König, Mary Ziems, and Marius Dölling for technical assistance and Anne Fiebig for sequence data submission. Barley genomic research in the labs of NS and MM is supported by a grant of the German Federal Ministry of Education and Research (SHAPE II, FKZ 031B0884). Research on crop–wild hybrids is supported by a grant of the German Research Foundation (DFG, MA 6611/7-1) to MM. Funding for open-access projects was in part covered by DFG grant HE 9114/1-1. Open Access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

NS, MM, and YG designed the research. AH performed sequencing. YG analyzed data. EW helped interpret data. YG and MM wrote the manuscript with input from all co-authors.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT

GBS and WGS raw data are available from the European Nucleotide Archive (Harrison et al., 2020) under project IDs PRJEB51488 and PRJEB51487, respectively. Accession numbers for individual genotypes are listed in Tables S1 and S4. The variant matrices are available from the Plant Genomics & Phenomics Research Data Repository (Arend et al., 2016) under DOI 10.5447/ipk/2022/10. The DOI was registered with e!DAL (Arend et al., 2014).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1. Passport data and GBS sequencing statistics of *H. agriocrithon* accessions in the IPK genebank.

Table S2. Accession codes of published GBS data.

Table S3. Geographic information of Eastern and Western barley in Figure 1c.

Table S4. WGS sequencing statistics.

Table S5. Haplotypes at *BTR1* and *BTR2*.

Table S6. Haplotypes at *VRS1*.

Table S7. Genetic distance of *vrs1.a1* carriers.

Table S8. f_4-statistics.

Table S9. Positions of the *BTR1*, *BTR2*, and *VRS1* genes in the MorexV3 reference genome sequence assembly.

Figure S1. Distribution of pairwise IBS values based on GBS SNPs in wild barley, domesticated barley, *H. agriocrithon*, and *lagunculiforme* accessions.

Figure S2. PCAs based on GBS SNPs with only 19 accessions of *lagunculiforme*.

Figure S3. Sample selection for WGS in the *lagunculiforme* cluster from Turkmenistan.

Figure S4. Population structure of wild barley in Western and Central Asia.

Figure S5. Genetic relationship analysis based on sequence variation in and around the *VRS1* gene.

Figure S6. Expected tree topologies in different evolutionary scenarios.

Figure S7. Graphical model of the *eu-agriocrithon* origin consistent with f_4-statistics.

Figure S8. Selection of 400 domesticated barleys from the global barley diversity space.

Figure S9. Assessment of the impact of sequence depth.

REFERENCES

Åberg, E. (1938) Hordeum agriocrithon nova sp., a wild six-rowed barley. *Annals of the Agricultural College of Sweden*, **6**, 159–212.

Alexander, D.H., Novembre, J. & Lange, K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**, 1655–1664.

Arend, D., Junker, A., Scholz, U., Schüler, D., Wylie, J. & Lange, M. (2016) PGP repository: a plant phenomics and genomics data publication infrastructure. *Database*, **2016**, baw033. Available from: https://doi.org/10.1093/database/baw033

© 2022 The Authors.

Arend, D., Lange, M., Chen, J., Colmsee, C., Flemming, S., Hecht, D. et al. (2014) E! DAL-a framework to store, share and publish research data. *BMC Bioinformatics*, **15**, 214.

Asouti, E. & Fuller, D.Q. (2012) From foraging to farming in the southern levant: the development of Epipalaeolithic and pre-pottery Neolithic plant management strategies. *Vegetation History and Archaeobotany*, **21**, 149–162.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O. et al. (2021) Twelve years of SAMtools and BCFtools. *GigaScience*, **10**, giab008. Available from: https://doi.org/10.1093/gigascience/giab008

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

Ellstrand, N.C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A., Jong, T.J.d. et al. (2013) Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 325–345.

Hammer, K. (1984) Blütenbiologische Untersuchungen an der Gaterslebener Wildgersten-Kollektion (Hordeum L. subgenusHordeum). *Die Kulturpflanze*, **32**, 79–88.

Harlan, J.R. & de Wet, J.M.J. (1971) Toward A rational classification of cultivated plants. *Taxon*, **20**, 509–517.

Harlan, J.R. & Zohary, D. (1966) Distribution of wild wheats and barley. *Science*, **153**, 1074–1080.

Harrison, P.W., Ahamed, A., Aslam, R., Alako, B.T.F., Burgin, J., Buso, N. et al. (2020) The European nucleotide archive in 2020. *Nucleic Acids Research*, **49**, D82–D85.

Helback, 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**, 365–372.

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

Hopf, M. (1983) Appendix B: Jericho plant remains. In: Kenyon, K.M. & Holland, T.A. (Eds.) *Excavations at Jericho, vol. 5. The pottery phases of the tell and other finds*. London: British School of Archaeology in Jerusalem, pp. 576–621.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Janzen, G.M., Wang, L. & Hufford, M.B. (2019) The extent of adaptive wild introgression in crops. *New Phytologist*, **221**, 1279–1288.

Jayakodi, M., Padmarasu, S., Haberer, G., Bonthala, V.S., Gundlach, H., Monat, C. et al. (2020) The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature*, **588**, 284–289.

Kamm, A. (1954) The discovery of wild six-rowed barley and wild Hordeum intermedium in Israel. *The Annals of the Royal Agricultural College of Sweden*, **21**, 287–320.

Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D. et al. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences*, **104**, 1424–1429.

Letunic, I. & Bork, P. (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, **49**, W293–W296.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Lundqvist, U. (2014) Scandinavian mutation research in barley – a historical review. *Hereditas*, **151**, 123–131.

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

Mascher, M., Richmond, T.A., Gerhardt, D.J., Himmelbach, A., Clissold, L., Sampath, D. et al. (2013) Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. *The Plant Journal*, **76**, 494–505.

Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C.S. et al. (2021) Long-read sequence assembly: a technical evaluation in barley. *The Plant Cell*, **33**, 1888–1906.

McGinnis, S. & Madden, T.L. (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*, **32**, W20–W25.

Milner, S.G., Jost, M., Taketa, S., Mazon, E.R., Himmelbach, A., Oppermann, M. et al. (2019) Genebank genomics highlights the diversity of a global barley collection. *Nature Genetics*, **51**, 319–326.

Morrell, P.L. & Clegg, M.T. (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proceedings of the National Academy of Sciences*, **104**, 3289–3294.

Murphy, P.J., Witcombe, J.R., Shewry, P.R. & Miflin, B.J. (1982) The origin of six-rowed 'wild' barley from the western Himalaya. *Euphytica*, **31**, 183–192.

Nadel, D., Piperno, D.R., Holst, I., Snir, A. & Weiss, E. (2012) New evidence for the processing of wild cereal grains at Ohalo II, a 23 000-year-old campsite on the shore of the sea of Galilee, Israel. *Antiquity*, **86**, 990–1003.

Oppermann, M., Weise, S., Dittmann, C. & Knüpffer, H. (2015) GBIS: the information system of the German Genebank. *Database*, **2015**, bav021.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y. et al. (2012) Ancient admixture in human history. *Genetics*, **192**, 1065–1093.

Patterson, N., Price, A.L. & Reich, D. (2006) Population structure and Eigenanalysis. *PLOS. Genetics*, **2**, e190.

Poets, A.M., Fang, Z., Clegg, M.T. & Morrell, P.L. (2015) Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biology*, **16**, 173.

Pourkheirandish, M., Hensel, G., Kilian, B., Senthil, N., Chen, G., Sameri, M. et al. (2015) Evolution of the grain dispersal system in barley. *Cell*, **162**, 527–539.

Pourkheirandish, M., Kanamori, H., Wu, J., Sakuma, S., Blattner, F.R. & 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. *The Plant Journal*, **94**, 525–534.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.

Purugganan, M.D. (2019) Evolutionary insights into the nature of plant domestication. *Current Biology*, **29**, R705–R714.

Rice, P., Longden, I. & Bleasby, A. (2000) EMBOSS: the European molecular biology open software suite. *Trends in Genetics*, **16**, 276–277.

Rosenberg, N.A. (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Russell, J., Mascher, M., Dawson, I.K., Kyriakidis, S., Calixto, C., Freund, F. et al. (2016) Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nature Genetics*, **48**, 1024–1030.

Sallam, A.H., Tyagi, P., Brown-Guedira, G., Muehlbauer, G.J., Hulse, A. & Steffenson, B.J. (2017) Genome-wide association mapping of stem rust resistance in *Hordeum vulgare* subsp. spontaneum. *G3 (Bethesda, Md.)*, **7**, 3491–3507.

Schiemann, E. (1932) *Entstehung der Kulturpflanzen*. Berlin: Gebrüder Borntraeger.

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

Tamura, K., Stecher, G. & Kumar, S. (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, **38**, 3022–3027.

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. *Theoretical and Applied Genetics*, **110**, 145–150.

Thirulogachandar, V., Alqudah, A.M., Koppolu, R., Rutten, T., Graner, A., Hensel, G. et al. (2017) Leaf primordium size specifies leaf width and vein number among row-type classes in barley. *The Plant Journal*, **91**, 601–612.

Wang, J.H., Xu, Z.M., Qiu, X.B., Li, L.L., Yu, S.Y., Li, T. et al. (2021) Genetic and molecular characterization of determinant of six-rowed spike of barley carrying vrs1.a4. *Theoretical and Applied Genetics*, **134**, 3225–3236.

Wendler, N., Mascher, M., Noh, C., Himmelbach, A., Scholz, U., Ruge-Wehling, B. et al. (2014) Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnology Journal*, **12**, 1122–1131.

Zeng, X., Guo, Y., Xu, Q., Mascher, M., Guo, G., Li, S. et al. (2018) Origin and evolution of qingke barley in Tibet. *Nature Communications*, **9**, 5433.

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: Oxford University Press on Demand.

© 2022 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.,