On the origin of photoperiod non-responsiveness in barley

Rajiv Sharma[1,2*], Salar Shaaf[1,3], Kerstin Neumann[1], Yu Go[1], Martin Mascher[1,4], Michal David[5], Adnan Al-Yassin[6,7], Hakan Özkan[8], Tom Blake[9], Sariel Hübner[10], Nora P. Castañeda-Álvarez[11], Stefania Grando[6,12], Salvatore Ceccarelli[6,12], Michael Baum[6], Andreas Graner[1], George Coupland[13], Klaus Pillen[14], Ehud Weiss[5], Ian J Mackay[2], Wayne Powell[2], Benjamin Kilian[1,11*]

[1]Genebank Department, Leibniz Institute of Plant Genetics and Crop Research (IPK), 06466 Gatersleben, **Germany**
[2]Scotland's Rural College (SRUC), Edinburgh, EH93JG, **UK**
[3]Department of Agronomy and Plant Breeding, College of Agriculture and Natural Resources, Islamic Azad University, Sanandaj branch, 66169 Sanandaj, **Iran**
[4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, **Germany**
[5]The Martin (Szusz) Department of Land of Israel Studies and Archaeology, Bar Ilan University, Ramat-Gan, 5290002, **Israel**
[6]The International Center for Agricultural Research in the Dry Areas - ICARDA, **Lebanon**
[7]National Agricultural Research Center (NARC), Amman, **Lebanon**
[8]Department of Field Crops, Faculty of Agriculture, University of Çukurova, 01330, Adana, **Turkey**
[9]Department of Plant Sciences & Plant Pathology, Montana State University, Montana, **USA**
[10]Galilee Research Institute (MIGAL), Tel-Hai College, 12210 Upper Galilee, **Israel**
[11]Global Crop Diversity Trust, 53113 Bonn, **Germany**
[12]Freelance consultant, **Italy**
[13]Department of Plant Developmental Biology, Max Planck Institute for Plant Breeding Research, 50829 Cologne, **Germany**
[14]Institute of Agricultural and Nutritional Sciences, Martin-Luther-University Halle-Wittenberg, 06120 Halle/Saale, **Germany**

*__Correspondence__:
Rajiv Sharma: Rajiv.Sharma@sruc.ac.uk
Benjamin Kilian: benjamin.kilian@croptrust.org

Abstract

In barley, the transition from the vegetative to reproductive phase is complex and under the control of photoperiodic and temperature conditions. One major gene involved is *PPD-H1*, a *PSEUDO-RESPONSE REGULATOR 7* (*PRR7*) that encodes a component of the circadian clock. Mutation at *PPD-H1* resulted in the photoperiod non-responsive *ppd-H1* alleles that are beneficial under high latitudinal environments as they allow vegetative growth during the long-day summer conditions whereby higher yields are harvested by farmers. Utilizing a diverse GWAS panel of world-wide origin and a genome-wide gene-based set of 50K SNP markers, a strong association of days to heading with the *PPD-H1* gene was detected in multi-location field trials. Re-sequencing of the gene spanning putative causative SNPs, SNP22 (Turner et al. 2005) and SNP48 (Jones et al. 2008), detected recombination between the two, previously reported to be in complete LD. Phenotyping of the recombinants and phylogenetic relationships among haplotypes supported the original conclusion of Turner et al. (2005) that SNP22, present in the CCT domain, is the most likely causative SNP. To infer the origin of non-responsiveness, the *PPD-H1* gene was re-sequenced in a geo-referenced collection of 2057 wild and domesticated barleys and compared with the allelic status of the 6000-year-old barley sample from the Yoram cave in the Masada cliff. A monophyletic and post-domestication origin in the Fertile Crescent was found in contrast to the pre-domestication origin proposed by Jones et al. (2008). We show that the photoperiod non-responsiveness originated from Desert type wild barley in the Southern Levant.

Key words: allele mining, ancient DNA, crop wild relatives, domestication history, Fertile Crescent, genetic diversity, growth habit, haplotype, *Hordeum*, Israel, plant genetic resources, Southern Levant

Introduction

Variation in days to heading affects vegetative growth, the source-sink relation and ultimately determines the yield of the crop. Thus, understanding its genetic architecture is important. Heading date is influenced by various environmental cues. Among them, temperature and photoperiod are the two most important in temperate cereals (McMaster and Moragues 2019). Barley (*Hordeum vulgare* L.), the fourth most important crop, has been traditionally classified

as a long-day plant. The species flowers in (early) spring in the Fertile Crescent and during long summer days in north Europe.

Barley is one of the oldest crops and was domesticated in the Fertile Crescent (Harlan and de Wet 1971; Zohary et al. 2013). Wild barley grains have been found in large amounts at the Ohalo II archaeological site on the shore of the Sea of Galilee and were dated to 23,000 years BP (Piperno et al. 2004; Weiss et al. 2004; Snir et al. 2015). This suggests that wild barley was collected from nature long before its domestication. Archaeological data at this site also suggest pre-domestication cultivation and provide evidence that domestication was sometimes not successful (Piperno et al. 2014; Snir et al. 2015). More recent evidence showed that the non-brittle rachis phenotype of domesticated barley originated at least twice, spatially and temporally independent, in the Southern (*btr1*-type) and Northern (*btr2*-type) Levant (Pourkheirandish et al. 2015). The ancient DNA (aDNA) sequencing of 6000-year-old barley grains from the Yoram Cave in the Masada Cliff, is consistent with the origin of barley domestication in the Southern Levant (Mascher et al. 2016).

Migration of the crop outside the Fertile Crescent required adaptation to different environmental conditions (von Bothmer et al. 2003). Present-day barleys are grown in diverse climatic conditions from Scandinavian countries in Northern Europe to the Sub-Saharan desert in Africa and in temperate regions of the Americas and Asia (Dawson et al. 2015), which makes it an important crop for adaptation under a changing climate in the 21st century.

In barley, the transition from the vegetative to reproductive phase is complex (Comadran et al. 2012). Recent knowledge for key flowering genes pathways involved has been reviewed for cereals by Steffan et al. (2014) and Monteagudo et al. (2019). Two major genes involved in photoperiod response have been identified and characterized in barley: *PHOTOPERIOD-H1*, *PPD-H1* on chromosome 2H (Turner et al. 2005) and *PHOTOPERIOD-H2*, *PPD-H2* on chromosome 1H (Kikuchi et al. 2011).

PPD-H1, is a *PSEUDO-RESPONSE REGULATOR 7* (*PRR7*) gene that encodes a component

of the circadian clock (Turner et al. 2005) promoting flowering in both winter- and spring-sown conditions. Photoperiod non-responsive allele(s) (*ppd-H1*) are beneficial under high latitudinal environments as they allow vegetative growth during the long-day spring and summer growth conditions, whereby higher yields are harvested by farmers.

The gene was cloned by Turner et al. (2005) using a winter x spring mapping population. *PPD-H1* consists of 676 amino acids and two major conserved domains described as pseudo-receiver and a CCT (CONSTANS, CONSTANS-like and TOC1).

The wild-type, dominant and photoperiod responsive allele(s) (*Ppd-H1*) at *PPD-H1* accelerate flowering by upregulating *VRN-H3* (*HvFT1*), which is mediated by the activity of CONSTANS (Turner et al. 2005; Campoli et al. 2012). Mutation of *Ppd-H1* resulted in the recessive and photoperiod non-responsive *ppd-H1* allele (Takahashi et al. 1963; Turner et al. 2005).

Two diagnostic single nucleotide polymorphisms (SNPs) differentiating between sensitivity (photoperiod responsive) and insensitivity (photoperiod non-responsive) to long days have been published: Turner et al. (2005) provided strong evidence for a causative SNP in the CCT domain (SNP22), while Jones et al. (2008) described a SNP in exon 6 (SNP48). Re-sequencing different sets of wild and cultivated barleys gave contradictory results. In Turner et al. (2005) and Cockram et al. (2007), the photoperiod non-responsive allele (*ppd-H1*) was not found in wild barley. Thus, the authors concluded that a natural mutation occurred *post-domestication* during the spread of barley cultivation in Europe. Subsequently, Jones et al. (2008) found three non-responsive haplotypes within *wild* barleys from Israel and four non-responsive haplotypes within *wild* barleys from Iran concluding that (I) the non-responsive phenotype of European landraces originated in wild barley from Iran; and (II) that the impact of wild barley from Iran was significant for the domestication history of European barley. However, these results were based on a relatively small sample of potentially admixed wild barleys of genebank origins (Jakob et al. 2014). These non-responsive *wild* barleys were from *east of the Fertile Crescent,* and earlier studies had reported they contributed little to present day European barleys (Kilian

et al. 2006; Morrell et al 2007).

In this paper we study the origin and domestication history of photoperiod non-responsiveness (*ppd-H1*) in barley. We first performed Genome-Wide Association Studies (GWAS) of days to heading (Hd) scored in multi-location field trials in a diverse panel of genotypes of worldwide origin. We then re-sequenced the potentially causative genomic region at *PPD-H1* in a comprehensive geo-referenced collection of wild and domesticated barley. Additionally, we used relationships among haplotypes, bioclimatic and phenotypic data, and comparison with the allelic status of the ancient 6000 years old barley sample from the Yoram cave.

Materials and methods

The Genome-wide association study panel (GWAS panel)

The diverse spring barley association panel consisted of 127 two-rowed and 97 six-rowed barley genotypes of world-wide origin (Haseneyer et al. 2010; Pasam et al. 2012). One-hundred-and nine genotypes originated from Europe, 45 from West Asia and North Africa (WANA), 40 from East Asia and 30 from the Americas (Table S1). The panel has been successfully utilized in other GWAS for agronomic traits, salt tolerance, drought-stress and candidate gene-based re-sequencing studies (e.g. Stracke et al. 2009; Long et al. 2013; Alqudah et al. 2014, Comadran et al. 2012; Abdel-Ghani et al. 2019).

Multi-location field trials (GWAS panel)

Days to heading (Hd) was scored in multi-location field trials; four locations in Germany, and one each in USA, Turkey and Syria (Table S2). As the locations were diverse, we treated the four Germany trials as a single location in the GWAS. Hd was scored as days from the date of sowing until 50 percent of the plants had reached growth stage @GS53 (Lancashire et al. 1991). Experimental design and further details are provided in Table S2 and in Supplementary Information online.

Markers for GWAS analysis

The GWAS panel was genotyped at TraitGenetics GmbH Gatersleben, Germany using the high throughput 50k iSelect SNP chip consisting of 43,461 SNPs (Bayer et al. 2017). All allele calls were manually inspected using GENOMESTUDIO Genotyping Module v2.0.2 (Illumina, San Diego, California). SNPs with more than 10 percent of missing values and > 10% heterozygous calls were excluded. A set of 37,387 SNPs (≥0.05 minor allele frequency) were used for GWAS. SNPs were anchored to the barley reference genome "Morex" assembly (Mascher et al. 2017). Analysis details on the genome-wide association and Site X SNP interaction are provided in Supplementary Information online.

Geo-referenced Diversity panel for allele mining and phylogenetic analysis

A comprehensive geo-referenced collection (Diversity panel) of 2195 wild and domesticated barley genotypes, from more than one hundred countries, was established to investigate the origin of photoperiod non-responsiveness (*ppd-H1*) in barley. This collection comprises a targeted selection of genotypes described in several publications (Badr et al. 2000; Kilian et al. 2006; Morrell et al. 2007; Jones et al. 2008; Comadran et al. 2012; Pasam et al. 2012, 2014; Tondelli et al. 2013; Jakob et al. 2014; Pourkheirandish et al. 2015; Russell et al. 2016; Xu et al. 2018; Bustos-Korts et al. 2019), extended by newly collected wild barleys i.e. from Israel and Turkey. All germplasm materials were single seed descended (SSD) for at least two generations and spikes were carefully isolated. Based on morphological and taxonomical characterization under field conditions in Germany, 138 samples were not considered for allele mining (Table S3).

For re-sequencing, we considered in total 2057 genotypes (Fig. S4) comprising (i) 942 wild barleys (*Hordeum vulgare* L. ssp. *spontaneum* (C. Koch) Thell., *H. spontaneum*) representing the immediate progenitor of domesticated barleys; (ii) 1110 domesticated genotypes (*Hordeum vulgare* L. ssp. *vulgare*) including 433 landraces from 58 countries, 673 cultivars from 54

161 countries and 4 others); and (iii) five feral *H. agriocrithon*. (Table S4); Supplementary
162 Information online).

163 **DNA Amplification and re-sequencing at *PPD-H1***

164 Genomic DNA was isolated from single leaves of 2057 SSD-derived genotypes with the Qiagen
165 DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's
166 instructions. The GWAS panel was re-sequenced using two primer combinations. A 1367 bp
167 fragment was considered for analysis. The Diversity panel was re-sequenced using one primer
168 combination only. After trimming, a fragment of 898 bp was considered for multiple sequence
169 alignments. All details are provided in Supplementary Information online.

170 **Sequence analysis at *PPD-H1***

171 PCR products were purified by NucleoFast 96 PCR plates (Macherey-Nagel, Germany) and
172 were sequenced directly on both DNA strands on Applied Biosystems (Weiterstadt, Germany)
173 ABI Prism 3730xL sequencer using BigDye terminators. DNA sequences were processed with
174 ABI DNA SEQUENCING ANALYSIS SOFTWARE 5.2 and later manually edited by BIOEDIT 7.2.5
175 (Hall 1999). Multiple sequence alignments were generated using the *PPD-H1* sequence of
176 cultivar Igri as reference (AY970701, Turner et al. 2005).
177 Haplotypes were defined using DNASP v. 5.10.01 (Librado and Rozas 2009). Singleton
178 haplotypes were confirmed by three independent amplifications and re-sequencing.
179 Sequence diversity statistics was calculated using DNASP. Diversity-loss, for the total number
180 of sites ($L\pi_{Total}$), and for silent sites ($L\pi_{silent}$) was calculated by: $L\pi = 1-(\pi_{domesticated}/\pi_{wild})$
181 (Tenaillon et al 2004).
182 Median-Joining (MJ) networks (Bandelt et al. 1999) were constructed using DNA ALIGNMENT
183 1.3.3.2 and NETWORK 5.0.0.1 (Fluxus Technology Ltd., Clare, Suffolk, UK). SPLITSTREE4
184 version 4.15.1 was used to generate a NeighborNet planar graph of haplotypes based on the un-
185 corrected P distances (Hudson and Bryant 2006).

Geographical distribution maps

All maps were prepared using QGIS (https://qgis.org), a free and open source Geographic Information System (GIS) software. All details are provided in Supplementary Information online.

Experiments to characterize photoperiod responsive and non-responsive genotypes

Hd of the GWAS panel under controlled long and short-day conditions

To confirm that haplotype H10 is truly photoperiod responsive, the phenotypic response to daylength (in days to heading) was tested for 41 selected genotypes including (i) 33 photoperiod responsive and non-responsive genotypes of the GWAS panel, and (ii) eight wild and one domesticated barley from the Diversity panel harboring mainly haplotype H10. The tested panel comprised a total of 10 different haplotypes of *PPD-H1* (Table S4; Supplementary Information online.

Contrasting the development of photoperiod responsive and non-responsive genotypes

Data published by Alqudah et al. (2014) was explored for this analysis wherein the panel was phenotyped at four developmental stages under inductive long day conditions in the greenhouse. Thermal time was measured as Growing Degree Days (GDD) from sowing to awn-primordium, tipping, heading and anther extrusion stages. The 218 genotypes and their haplotypes considered for this study are shown in Table S4.

Heading date under vernalized and non-vernalized long-day field conditions

This experiment was designed to characterize the growth habit and photoperiod responsiveness of genotypes under contrasting vernalization treatments and long-term conditions. In total, 843 wild and domesticated barley genotypes from the Diversity panel were studied in the field at IPK under vernalized and non-vernalized conditions (Table S4; Supplementary Information online.

Inferring the allelic states of *PPD-H1* and *HvCEN* in the 6000 years old ancient barley sample JK3014 from the Yoram Cave

Ancient barley sequences (Mascher et al. 2016) were retrieved from the short-read archive accession (PRJEB12197) (https://www.ebi.ac.uk/ena/data/view/PRJEB12197). After adaptor removal and merging of overlapping paired-end sequences with LEEHOM (Renaud et al. 2014), merged reads with a minimum length of 30 bp were mapped to the barley reference genome (version: MorexV2, Monat et al. 2019) using BWA-MEM. Variants (SNP and short indels) were called from uniquely mapped reads (MAPQ >= 20) using BCFTOOLS (Li et al. 2011). The genotype calls were filtered using the following criteria: (i) the minimum reads depth was 1 for homozygous calls, and (ii) the minimum reads depth was 2 in both alleles for heterozygous calls. *PPD-H1* and *HvCEN* sequences (Comadran et al. 2012) were aligned to the MorexV2 assembly BLAST (Altschul et al. 1990). Based on the alignment result, we obtained the relative position in MorexV2 for the previously reported variants of *PPD-H1* (Turner et al. 2005) and *HvCEN*, and retrieved the genotype calls at these sites from our variant calling file.

Analysis of Environmental Data

In order to investigate and compare the sampling sites of wild, landrace and the ancient barley, 19 GIS-derived historical bioclimatic variables (mean value from 1950-2000) with a spatial resolution of 2.5-min (5 km) were extracted from www.worldclim.org using the R package raster v.3.0-7 (Hijmans et al. 2005). Data was analysed in R (R core team 2018). See Supplementary Information online.

Results

Phenotypic data evaluation for the GWAS panel

Large variation was observed for Hd in the diverse GWAS panel across the multi-location field trials (Fig. 1). The trials in Germany showed the largest, and in the USA the smallest variation

(Fig. 1; Table S5). The coefficient of variation was high (>4.93) for all trials, reflecting the diversity of the phenotype at each location. Heritability values were moderate to high (0.66 – 0.99; Table S5). We observed significant ($P < 0.001$, except USA vs. Turkey) and positive correlations between all pairwise Hd phenotypes (Table S6). The geographically closest locations in Syria and Turkey showed the highest correlations followed by Syria and Germany; while correlations with the USA were the smallest.

Multi-location heading date associations

A total of 296 SNPs across all sites displayed significant associations ($-\log_{10}P \geq 4.0$) for Hd (Table S7). Most significant regions showed associations with more than one SNP, indicative of the high SNP density and linkage disequilibrium (LD) of the genomic regions detected in the analysis (Fig. 2; Table S7). Despite diverse locations of our trials, major peaks were colocalized near genes like *PPD-H1*, *HvCEN*, *VrnH2* and *FT1*. Interestingly, the *PPD-H1* region was not significantly associated for the Montana/USA experiment, which is in agreement with its low pair-wise correlations with the other field sites (Table S6).

Five SNPs at *PPD-H1* showed the highest association for the trials in Germany with significant values up to $-\log_{10}P = 6.39$ (Table S7). Association at this region reduces days to heading (Hd) by up to 4.15 days. Further, additional SNPs located on chromosome 2H close to *PPD-H1* were significantly associated indicating the extended LD of the *PpdH1* region (Table S7). SNP48 and SNP22 were analyzed and both were significant with $-\log_{10}P = 6.21$ which is similar as the highest associated SNPs on the 50K SNP chip (BOPA2_12_30872 and JHI-Hv50k-2016-73417) (Table S7 and Fig. 3a). Highest associated SNPs were from the intronic region at *PpdH1* (intron 3) and thus less likely to be causative compared to SNP22 and SNP48. Within the 177 accessions for which complete data sets across all locations were available (Table S1), SNP22 and SNP48 were in complete LD ($r = 1$). However, in the full GWAS panel (224 accessions) these SNPs remained in nearly perfect LD with $r = 0.9817$ as also observed by Turner et al.

(2005) and Jones et al. (2008). SNP22 and SNP48 showed highly significant differences in Hd in Germany compared to Turkey and Syria (Fig. S1). Non-responsive genotypes flowered late compared to responsive types. This could be explained by longer day lengths at higher latitudes in Germany compared to other sites (Table S2).

Multi-location heading date - SNP interaction

From the significantly associated 258 non-redundant SNPs, 59% (N=153) showed a significant interaction with sites ($P \leq 0.05$ "p (Site x SNP)"), and 19% (N=51) had a significant interaction with Bonferroni corrected value of 0.05 (Table S8). For 35% (N=54) of significantly interacting SNPs, the relative magnitudes of the interaction and main effects caused the net effect of the SNP on Hd to reverse between at least one pair of sites. For example, the effect on Hd of SNP JHI-Hv50k-2016-73562 near the *PPD-H1* region was predicted as -3.075 days in Germany but +0.739 days in Turkey suggesting contrasting effects of alleles in varying environments (Table 9).

Haplotype diversity at PPD-H1 in the GWAS panel

The GWAS panel harbored 14 haplotypes (H) (Fig. 3b). Except for genotypes of European origin, which comprised mostly non-responsive haplotypes, no geographical cline in haplotype frequencies was found, suggesting non-responsiveness was selected in European barleys to allow vegetative growth under long-day conditions (Table S4). The most prevalent haplotypes in the panel were haplotype H2 (36%) and H1 (22%). Most European barleys carried haplotype H2 (65%) and haplotype H1 (24%) (Table S10). Barleys from WANA carried in total 11 haplotypes, with intermediate to low frequencies. Frequent haplotypes within American barleys were haplotype H1 (34%) and H8 (25%), while haplotype H6b was predominant in East Asia (23%). Six haplotypes were region-specific (Table S10).

Photoperiod non-responsive haplotypes should carry nucleotide T at both positions, at SNP22 and at SNP48, compared to the cultivar Igri as detected by Turner et al. (2005). Three

haplotypes (H1, H2, H56) carried nucleotide T at both positions. Interestingly, haplotype H10 carried T at SNP48 but G at SNP22 and was found in two cultivars (BCC533 and BCC759, from Nepal and India, respectively) (Fig. 3b; Fig. S2; Table S4).

To ascertain whether these two genotypes are photoperiod responsive or non-responsive, both were grown together with 39 further genotypes from the Diversity panel under long and short day conditions. For Hd under long and short-day conditions, repeatability was high with 0.98 and 0.91, respectively. While there was no difference among the haplotypes in short day condition, under long days a clear effect was observed. The seven genotypes carrying H10 were found to be photoperiod responsive (Fig. 4), as haplotype H10 reached heading date significantly earlier than its derived haplotype H1 under long day conditions ($p<0.001$, 22 days earlier), while under short days, the observed difference of 3.8 days was not significant. This demonstrates that haplotype H10 is photoperiod responsive and concludes SNP22 as the most likely causative SNP. Supporting our findings, both domesticated H10-genotypes of the GWAS panel were early under long day conditions from the tipping until the anther extrusion stage in the study of Alqudah et al. (2014) (Fig. S3). Accordingly, these two genotypes headed also earlier than the non-responsive genotypes of the GWAS panel in field trials in Germany (across four locations and two years), corroborating that haplotype H10 is indeed photoperiod responsive. Therefore, our data strongly suggest that SNP22 located in the CCT domain is the causal SNP.

Genetic diversity at PPD-H1 within the Diversity panel

We detected ninety haplotypes in 2057 re-sequenced and taxonomically confirmed SSD-derived genotypes (Fig. 5; Figs. S4, S5; Table S13). Higher diversity was found in wild than in domesticated barleys: 71 haplotypes in wild and 27 in domesticated barleys (23 haplotypes in landraces and 17 in cultivars). Eight haplotypes were shared between domesticated and wild barleys. 17 haplotypes were unique to domesticated barley including the non-photoperiod

309 responsive haplotypes H1 and H2. Most haplotypes detected in wild barley had low frequencies.
310 Only haplotypes H6 (40.44%), H4 (14.33%) and H7 (13.69%) were more frequent and were
311 found in 68.46% of the wild barleys (Fig. 5; Table S11). These major haplotypes were not
312 exclusive to wild barley indicating extensive post-domestication utilization.
313 Wild barleys from Israel possessed the highest genetic diversity (47 haplotypes), followed by
314 Turkey (19 haplotypes). Several haplotypes were region-specific. Sixty-three haplotypes were
315 unique to wild barley and not exploited in the domesticated barleys, which might harbor a key
316 to local environment adaptation. Interestingly, several of the non-exploited haplotypes clustered
317 together (Fig. 5; Table S11; Supplementary Information online).
318 Within domesticated barleys, the most frequent haplotypes were H1 (15%) and H2 (27%).
319 These two haplotypes were exclusively found in domesticated barley. Haplotype H1 and its
320 four derived haplotypes (H55, H56, H2, H47) all carried "T" at SNP22 and were considered as
321 photoperiod non-responsive (Fig. 5; Tables S4, 13).
322 Additionally, we detected 23 haplotypes in 128 genotypes that were classified previously based
323 on their passport data as wild barleys (Table S3; Jones et al. 2008). But careful inspection of
324 them by growing them in the field and studying plant morphology and gene sequence data, we
325 found that they either carrying introgressions from domesticated barley or are seed
326 contaminants from *ex situ* maintenance and seed sharing among genebanks as also reported in
327 Jakob et al. (2014). Thus, we classify them as admixtures and were excluded from the analysis
328 (Supplementary Information online).

329 Interestingly, in the Diversity panel, haplotypic diversity for wild (H=0.79) and domesticated
330 (H=0.86) barleys was comparable (Table S14). Nucleotide diversity (π) and Waterson's Theta
331 (θ) were lower for domesticated barley than for wild barley. Differences in nucleotide diversity
332 between wild and domesticated were larger at silent sites (π silent). Within the Diversity panel,
333 59 segregating sites were found for wild barley that were monomorphic in domesticated barley.

By contrast, domesticated barleys carried 13 segregating sites that were monomorphic in wild material. Tajima's *D* was negative (-1.55) ($P > 0.10$) for wild barleys but positive for domesticated barley (0.32, $P > 0.10$). Only, non-synonymous sites showed significant negative Tajima *D* values (-1.84, $P < 0.05$) in wild compared to domesticated barleys (0.93, $P > 0.10$) indicating the presence of rare alleles in wild compared to domesticated barleys. Further, higher Tajima's *D* values (non-Syn/Syn ratio) were observed for wild (3.55) compared to domesticated (3.01) barley. By comparing wild and domesticated barleys, a 32% loss of diversity ($L\pi_{silent}$) at the silent sites and 9% ($L\pi_{Total}$) at the total number of segregating sites was observed, indicating a relatively moderate diversity loss post-domestication.

Phylogenetic relationships and geographic distribution of photoperiod responsive and non-responsive haplotypes

Phylogenetic relationships between 90 haplotypes were visualized using a MJ-network (Fig. 5). The central position in the MJ network is occupied by haplotype H75 carried by two wild barley genotypes from Israel (Shilat) located on the western slopes of the Judea mountain ridge. This region was previously identified as a hybrid zone between the Desert and Coast wild barley ecotypes (Hübner et al. 2009, 2013). From here, major geographical groups of haplotypes are visible:

Geographical group 1*:* from H75 to H10 to all photoperiod non-responsive haplotypes: photoperiod non-responsive haplotypes (H1, H55, H56, H2, H47) were exclusively found in domesticated barley (and one *H. agriocrithon*), clustered together and originated from the photoperiod responsive haplotype H10.

Interestingly, haplotype H10 was found in 16 wild barleys (13x Israel, 3x Iran), two landraces (1x Turkey, 1x Nepal) and two cultivars (1x India, 1x Japan) (Table S4, Fig. S6). Among the wild barleys from Israel, 12 were collected in the south and east of Israel where the climate is dry and warm. They were characterized as Desert ecotype and showed early flowering (Hübner

et al. 2013). One genotype was collected in the hybrid zone between Desert and Northern ecotypes (FT138, Moledet). Most importantly, among the wild barleys from Israel were eight genotypes recently collected by Hübner et al. (2009) that provides convincing evidence that haplotype H10 exists in wild-stands in nature and that it is associated with the early flowering Desert ecotype. The remaining four genotypes from Israel were collected in the 1960s and 1970s. Also, these showed truly wild characteristics and were collected approximately from the same locations as the eight genotypes by Hübner et al. (2009) (Fig. 7; Table S4). In a recent study wild barleys collected apart 28 years from Isarel were shown earlier flowering phenotype another sign of natural selection iminged on wild stands of present day due to global warming (Qian et al. 2019).

Also among the wild genotypes harboring haplotype H10 were three genotypes from Iran, collected between 1952-1958 in the province of Khuzestan (Table S4).

Haplotype H10 was not found in Turkey despite extensively sampling of the entire distribution area of the species in the country (N=362). The disjunctive distribution of haplotype H10 suggests similar environmental conditions at the respective collection sites.

The non-responsive haplotype H1 was found in domesticated genotypes mostly from Central Europe (Table S4). Haplotype H2, was mainly found in genotypes originating from North-West Europe (Fig. S4). Thus, geographical distribution differences of non-responsive and responsive haplotypes were observed (Figs. S4, S7). Interestingly, the non-responsive haplotypes H1 and H2 were also found in domesticated barley from the Fertile Crescent including in Bedouin landraces, newly collected by Hübner et al. (2009) (Supplementary Information online).

Geographical group 2: H75 -> H7 -> H8 -> H92: Haplotype H7 was found in 256 genotypes of almost equal frequency in wild (13.7%) and domesticated (11.4%) barleys. Among the wilds, haplotype H7 was found in 33.3% of wild barley from Israel and 0.6% genotypes from Gaziantep/Turkey. Also, 88 landraces predominantly collected in North Africa and the Near East harbored this haplotype. Haplotype H8 was found in 0.5% of wild barley (4x ISR, 1x CYP)

385 but in 9.2 % of domesticated barley, representing mostly landraces collected in the eastern
386 Mediterranean. Haplotype H92 was mainly found in landraces from Algeria.
387 **Geographical group 3**: H75 -> H13 -> H4 -> H3 -> H95: Haplotype H13 was detected in 13
388 wild and one landrace barley from Chad. Haplotype H4 represents a major haplotype and was
389 detected in 230 genotypes (14.33% of wild and 8.55% of domesticated). Wild barleys were
390 collected in Greece (N=1), Israel (N=54, including the Desert type barley FT143), Jordan
391 (N=6), Lebanon (N=4), Syria (N=6) and Turkey (N=64). The remaining were 39 landraces
392 mainly from the Fertile Crescent but also from Libya (N=3), and 56 cultivars mainly from
393 Turkey. All wild barleys harboring haplotype H3 (N=18) were collected west of Gaziantep in
394 Turkey (Supplementary Information online).

395 ***The 6000 years old domesticated barley sample from the Yoram cave was photoperiod***
396 ***responsive***
397 We determined the allelic states of *PPD-H1* and *HvCEN* in the aDNA sample JK3014 extracted
398 from barley grains found in the Judean desert and dated to 6000 years BP (Mascher et al. 2016).
399 After aligning previously published ancient DNA sequences (sample JK3014) to the current
400 barley reference genome sequence (MorexV2, Monat et al. 2019), the genotypes of putative
401 causal variants in *PPD-H1* and *HvCEN* were determined (Table S15). The ancient barley
402 carried the ancestral alleles at *PPD-H1* (photoperiod responsive, G at SNP22) and *HvCEN*
403 ('early'-flowering, C [proline], at position 531 of Comadran et al. 2012). Multiple sequence
404 alignment analysis concluded that JK3014 carried the following haplotypes; H4 at *PPD-H1* and
405 IV at *HvCEN*. This haplotype combination has not been found in the comprehensive Diversity
406 panel consisting of 2057 genotypes collected in the last 150 years. However, H4 is one of the
407 major haplotypes of *PPD-H1* and shared by wild and domesticated barleys. Interestingly,
408 *HvCEN* haplotype IV was not found in wild barley, it is derived from haplotype II (Comadran

et al. 2012). Haplotype IV was found in 21 domesticated barleys mainly from Ethiopia (N=10) but also from Turkey (N=2), Syria (N=1) and Yemen (N=1) (Table S4).

Analysis of environmental data shed more light on the region of origin of photoperiod non-responsiveness

To characterize the collection sites, bioclimate variables were clustered by employing PCA. The precipitation-related variables i.e. Bio14, Bio17, Bio12 and Bio18, were separated from the temperature-related variables (Bio1, Bio5, Bio6, Bio9, Bio10, Bio11) with PC1. With PC2, Bio4, Bio7 were separated from Bio8, Bio13 and Bio16 (Fig. S9; Fig. 6). The PC1 separated mainly collection sites of non-responsive barley carrying the haplotypes H1 and H2 from responsive barleys (other haplotypes). Most collection sites of non-responsive genotypes had higher values of the precipitation-related variables and lower values of the temperature variables (Fig. S10). Interestingly, the PC2 separated non-responsive barleys containing haplotype 2, into two groups (Fig. 6a). One group consisted mainly of landraces from Ethiopia that showed strong correlations with precipitation-related variables Bio13 and Bio16. The second group comprised 40 landraces from 19 countries including Egypt and Israel (Table S16). Especially based on the temperature-related variables (Bio1, Bio2, Bio5, Bio9 and Bio10), the collection sites of H10 containing wild barley from Israel were more closely related to the collection sites of non-responsive H1 containing domesticated barley than the collection sites of H10 wild barley from Iran (Fig. 6a-c; Table S16).

In general, landraces containing H1 from the Near East were collected in regions with higher precipitation (Bio12, Bio14, Bio17) than the H10 wild barley (Fig. 6b). H10 containing landraces were collected from contrasting environments. Collection sites of Desert type wild barley from the Southern Levant were positively associated with Bio1, Bio6, Bio11, and Bio15 (Fig. 6c; Supplementary Information online).

Current environmental conditions at the collection sites of extant wild and landrace barley were compared with the collection site of the ancient Masada barley (JK3014) on the basis of pairwise Euclidean distances and bioclimatic variables. The following results were obtained (Fig. 7; Table S18): (i) the environmentally overall most similar collection site of all extant wild barley to Masada was the collection site of the wild barley population B1K-12 (Israel, Kidron stream), where all wild barley was described as Desert types (Hübner et al. 2009, 2013). The *PPD-H1* haplotypes found in this population were H6, H7, H28 and H66; and the haplotype combinations for *PPD-H1* and *HvCEN* were H66HII (FT045) and H7HII (FT047), which are photoperiod responsive and 'early' flowering. Most interestingly, also the closest wild barley to *btr1* was collected here (FT643, Pourkheirandish et al. (2015); and (ii) The environmentally second closest collection site, which is also the closest H10 collection site, to Masada is Almog, where FT639 (Desert type) was collected. From this we concluded that the *HvCEN* haplotype II (progenitor haplotype of IV), but also the *PPD-H1* haplotype 10 were found very close to the Masada cliff. The ecologically closest landrace collection site (with exact details in the passport data) to Masada was found in Egypt supporting Mascher et al. (2016) (Table S18; Supplementary Information online)

Vernalization requirement and phenotypic performance of genotypes containing haplotype H10 under long-day field conditions

Under vernalized and non-vernalized conditions, heading date of 843 genotypes of wild and domesticated barley was investigated to determine their vernalization requirement and to characterize key agronomic traits in the vernalized treatment. For all traits, the repeatability of data was high and ranged from 0.86 to 0.98.

In the non-vernalized treatment, 582 genotypes were heading, while the remaining 261 genotypes (30%) were not flowering and therefore considered winter types. On average, heading date was reached 15 days later under non-vernalized compared to vernalized conditions

among the set of 582 genotypes. Interestingly, a bimodal distribution of Hd differences was observed, indicating two phenotypic groups - spring (flowering without vernalization) and facultative growth habit (flowering without vernalization but earlier when vernalized) (Fig. S8a). Of all 843 barley genotypes, 97 wild accessions did not exhibit completely wild characteristics and were excluded from subsequent analysis, leaving 746 accessions for comparison (470 wild and 276 domesticated). A total of 204 accessions were classified as spring types (including 6 wild barleys from Israel) and 305 as facultative, while 237 accessions were classified as winter types.

The genotypes carrying the non-responsive haplotypes H1 and H2 at *PPD-H1* were mostly spring types. In contrast, the photoperiod responsive progenitor haplotype H10 was mainly found in facultative types (wild barley from Israel, N=9; and wild barley from Iran, N=3) but also in two spring types (wild barley FT147 from Israel, landrace FT537 from Turkey) and one winter type (wild barley FT002 from Israel) (Fig. S8b). The genotypes with haplotype H10 showed a short life cycle with the second earliest heading date (1. H66, 2. H10, 3. H26) and the earliest maturity date (1. H10, 2. H66, 3. H26) of all haplotypes under vernalized, long-day field conditions in Germany (Fig. S8c; Table S19), even if only wild barley was considered (Fig. S8d). In addition, the H10-containing genotypes were among the three genotypes with the shortest plant height, narrowest flag leaves, shortest main ear and narrowest main ear width (Table S19). From these data we conclude that plants containing haplotype 10 are well adapted to their local environmental conditions in the Southern Levant or in Khuzestan, and that they are characterized by facultative or even spring growth habit (Supplementary Information online).

Discussion

We present a GWAS of Hd across multi-location field trials that cover diverse latitudes and longitudes and detected significant associations at the *PPD-H1* genomic region. Using re-

483 sequencing covering causative SNPs of the *PPD-H1* gene, we found that SNP48 (Jones et al.
484 2008) is not causal. Further, we re-sequenced the largest SSD and geo-referenced Diversity
485 panel of barley to study *PPD-H1* diversity, phylogeny and domestication history. We arrived
486 at the following five key findings:

487 **1. Adaptation of barley depends upon the environment**

488 The onset of Hd, a surrogate measurement of flowering time in crops, takes place in favorable
489 environments. The photoperiod responsive *PPD-H1* gene triggers flowering when the day
490 length increases. In winter barley, the vernalization responsive *Vrn-H2* gene acts as repressor
491 of flowering, although many other gene interactions are also involved. With exposure to cold
492 temperature, the repression gradually reduces and thus *PPD-H1* promotes flowering during the
493 spring period. However, exposure to long photoperiod promotes early flowering in barley which
494 carries the dominant *Ppd-H1* allele, and so reduces the vegetative growth. In higher latitudes
495 (like in Central or Northern Europe), the photoperiod non-responsive *ppd-H1* allele enables
496 vegetative growth during long days (spring season) and thus ensures higher yield. The GWAS
497 analysis of multi-environment trials detected associations from genomic regions corresponding
498 to the major flowering-time genes *Ppd-H2/HvFT3*, *PPD-H1*, *HvCEN*, *HvLUX1*, *Sdw1*/*denso*,
499 *Vrn-H2*, *Vrn-H1*, *Vrn-H3/HvFT1* and *HvCO1* (Fig. 2). However, significant effects vary across
500 sites suggesting that for adaptation plants utilize common but variable effects on flowering
501 time, which was also indicated by significant SNP x site interactions. In a recent finding, *PpdH1*
502 and *HvCEN* genes elucidate the Genotype x Environment pattern when grown under spring and
503 winter-sown trials of barley (Bustos-Korts et al. 2019), as we also found in trials across the
504 diverse latitudes (Table S8-S9). Crop plants grown under variable environments experience
505 different photoperiod and therefore the influence of the regulator of Hd varies with the
506 environment (Göransson et al. 2019; Afsharyan et al. 2020). For instance, our lower latitude
507 trials (Turkey, Syria) have less annual variation in photoperiod compared to higher latitude

Germany where average day length increases by up to 4 hours during the spring growing season. Furthermore, temperature, humidity and soil type can also influence Hd. The environment specific associations observed in our study indicate that adjustment of Hd depends upon gene expression in each environment. Absence of major significant association at *PPD-H1* in Montana/USA could be due to the adaptation of the genotypes. In Montana, plants must withstand low temperatures and freezing during nighttime, and chilling during daytime until mid-June, followed by a terminal drought. Earlier flowering lines withstand terminal drought environments better. However, in our GWAS panel most domesticated genotypes originated from Europe and were adapted to European conditions and are photoperiod non-responsive and thereby developing more biomass. In the short Montana season, they experienced forced flowering, and this might be the reason that no major associations were found. Notwithstanding this, the small peaks that were observed in the *HvCEN*, *VrnH2* and *Vrn-H3*/ *HvFT1* regions and might be important for adaptation under drought conditions such as in Montana.

The region near the *Vrn-H3*/ *HvFT1* gene showed strong peaks at three sites (Germany, Turkey, USA). This is interesting as it was reported that *HvFT1* is the central regulator where signals are perceived that promote flowering (Nitcher et al. 2013). It is not known why no association was detected in this genomic region in Syria. Either it suggests that different loci in high linkage disequilibrium (LD) but in dispersion contribute to the effects observed in our study or that contrasting alleles within the gene cause these effects which could be supported by observing late heading effects in Syria. Further study is needed to ascertain the effects from the *Vrn-H3*/ *HvFT1* locus by re-sequencing the region. It has been shown recently that at least four identical copies of the *Vrn-H3*/ *HvFT1* gene are present in barley varieties that have spring growth habit whereas a single copy is present in most other barley varieties. Therefore, copy number variation causes huge effects on the expression of *Vrn-H3*/*HvFT1* (Nitcher et al. 2013) and may explain the variation seen in this region. Re-sequencing could shed more light on the observed

differences. As shown by Casas et al. (2011), four *HvFT1* haplotypes contributed to differences in flowering time but at the SNP level we could not ascertain the differences precisely.

Our findings confirm Maurer et al. (2015) and Herzig et al. (2018). Under the conditions of the field trials of Halle (Germany) and Dundee (Scotland), the photoperiod responsive wild barley alleles of *PPD-H1* studied in the barley Nested Association Mapping (NAM) population HEB-25 flowered substantially earlier than the non-responsive cultivated barley allele, respectively. In a later study with selected HEB-25 lines segregating for responsive and non-responsive alleles, the early flowering effect of wild barley *Ppd-H1* alleles was verified in Dundee (Scotland), Halle (Germany) and Al-Karak (Jordan), but disappeared under field conditions in Dubai (United Arab Emirates) and Adelaide (Australia) with Hd at day lengths below 13 hours (Wiegmann et al. 2019). The drought conditions in Al-Karak lead to a positive effect of the *Ppd-H1* responsive allele on grain yield.

Further study is needed, but a similar effect as in Al-Karak can be expected in Israel, where the Desert type of wild barley with haplotype 10 was found. In this region, where barleys flower early (before Easter; Passover = Easter is the holyday of the barley harvest), the difference in heading date between haplotypes that respond to the photoperiod, but also between responsive and non-responsive haplotypes could be much less pronounced than in higher latitudes.

Interestingly, the progenitor haplotype H10 of the non-responsive haplotype H1 is associated with the Desert ecotype, which is characterized by early flowering (very early in Almog, around January, when daylength is about 10.5h) to avoid terminal drought. Adaptation to water availability is likely to have led to an orchestra of early alleles at many flowering related loci.

As expected, further early haplotypes were found in the Judean Desert. These haplotypes occur together in the Desert type wild barley populations and with different frequencies; e.g. (i) Yeruham: 3x H10, 2x H28; (ii) Neomi: 5x H66; (iii) Shivta: 1x H10, 2x H7, 1x H26, 1x H69; or (iv) Kidron stream: 1x H6, 1x H7, 1x H28, 1x H66 (Table S4). All these haplotypes should have similar effects on Hd in the Southern Levant. Haplotype H10 may not have any advantage

559 over other *PPD-H1* haplotypes in the region. In fact, under field conditions in Germany, all
560 these haplotypes were the earliest for time to maturity, and with the exception of H6, also the
561 earliest in Hd (Table S19). Most genotypes carrying these haplotypes were classified as
562 facultative types (Fig. S8b). The climate in the Judean desert does not favor a strong
563 vernalization requirement, and we conclude that most Desert-type wild barley has a facultative
564 growth habit from which a spring growth habit has evolved. The only six wild barleys with
565 spring growth habit are from Israel: FT050 (H6), FT147 (H10), FT288 (H26), FT301 (H7) and
566 FT019 (H7), which suggests that the spring growth habit originated in the Southern Levant.
567 Support for this hypothesis comes from Saisho et al (2011), who classified seven of 161 wild
568 barley types as spring types, with the majority being facultative. However, no wild barley from
569 Israel was included in their study, and the taxonomic status remains unclear, as certain taxa of
570 hybrid origin, were included.

571 **2. SNP22 is the causal basis of the *ppd-H1* mutation**

572 Our GWAS results confirmed that *PPD-H1* is one of the most important genes in regulating
573 flowering and variation as this gene causes natural diversity of barley flowering. In our study
574 the two functional SNPs reported by Turner et al. (2005) and Jones et al. (2008) did not show
575 differences in terms of the significance level. This is in accordance with the observed near
576 perfect LD between these SNPs reported by Turner et al. (2005) and Jones et al. (2008).
577 However, within just 87 genebank accessions, Jones et al. (2008) concluded that the functional
578 SNP22 reported by Turner et al. (2005) is not the causal SNP.
579 Re-sequencing of the gene space spanning putative causative SNP22 and SNP48, we detected
580 recombination between the two, previously reported to be in complete LD (Fig. 3). In the
581 Diversity panel of 2057 barley accessions we found 20 (0,97%) accessions (16 wild, 4
582 domesticated) with such recombination (haplotype H10), compared to Jones et al. (2008) who
583 used a landrace panel of mostly European origin and found no recombination between these

584 SNPs. We show that haplotype H10 containing genotypes respond to photoperiod, which leads
585 to early heading under long-day conditions (vernalized) compared to genotypes with non-
586 responsive haplotypes (Fig. 4; Figs. S3, S8). Therefore, SNP22 in the CCT domain of Turner
587 et al. (2005) should be considered as the causal basis of the *ppd-H1* mutation, which is
588 supported by phylogenetic analysis. All photoperiod non-responsive haplotypes (H47, H2, H56,
589 H55 and H1) clustered together in the MJ network, thus suggesting a monophyletic origin of
590 photoperiod non-responsiveness in barley due to the G to T non-synonymous substitution at
591 *SNP22* (Fig. 5). *It is important* to note that in our study no phenotypically wild barley was found
592 that carried a photoperiod non-responsive haplotype (*ppd-H1*). Thus, our data indicate that all
593 extant wild barley is photoperiod responsive (*Ppd-H1*). This is supported by Baloch et al. (2013)
594 studying wild barley from Jordan and Iran.

595 **3. Photoperiod non-responsiveness originated from Desert type wild barley in the**
596 **Southern Levant**

597 In total, sixteen wild barleys harboring H10 were collected from Israel and Iran (and not from
598 the central part of the Fertile Crescent). Based on our analysis of environmental data, we show
599 that wild barley from Israel containing H10 (N=13) grow in more similar habitats to non-
600 responsive barley than H10 wild barley from Iran (N=3). For some of them, the haplotype of a
601 second important flowering time gene, *HvCEN* is known form the study of Comadran et al.
602 (2012). The following *PPD-H1* - *HvCEN* haplotype combinations were found in Israel (3x
603 H10III, 1x H10IX) and Iran (3x H10I). This result supports the origin of European non-
604 responsive barley from Desert type wild barley from the Southern Levant, likely carrying
605 H10III, as haplotype III at *HvCEN* was by far the most frequent haplotype in European non-
606 responsive spring and winter barley (Comadran et al. 2012).
607 But how can the occurrence of wild barley from Iran with haplotype 10 be explained? We
608 speculate that the province of Khuzestan in southwestern Iran was part of the ancient natural

609 distribution range of the species and that haplotype 10 survived the Last Glacial Maximum
610 (LGM) about 21k years ago in the region (Jakob et al. 2014).

611 Among the four domesticated barleys harboring the haplotype H10, we found a 2-rowed, naked
612 landrace from Turkey (FT537) that was collected by Jack Harlan in 1948. The allelic status at
613 *HvCEN* is not known. In contrast, the haplotype combination H10I was found in (i) one landrace
614 from Nepal (6-rowed, hulled); (ii) one cultivar from India (6-rowed, hulled); and (iii) one 2-
615 rowed, naked cultivar from Japan (Fig. S6). These four domesticated barleys probably originate
616 from H10I containing wild barleys from Iran. Further study is needed to investigate the
617 contribution of wild barley from Iran to the *btr2* genepool (Pourkheirandish et al. 2015).
618 Remains of non-brittle two-rowed barley dated to the Middle PPNB (10th century BP) were
619 found in sites across the Fertile Crescent: Jericho (Israel, only about 10 km from the haplotype
620 H10 collection site Almog), Tell Aswad (Syria), Jarmo (Iraq) but also Ali Kosh (Iran,
621 Khuzestan province, where H10 wild barley was collected) (Alizadeh (2003); Zohary et al.
622 (2013).

623 **4. No severe genetic bottleneck at the *PPD-H1* gene**

624 Often a severe genetic diversity change is observed when comparing wild and domesticated
625 barley populations. Signatures of domestication include reduced genetic variation compared to
626 the wild. We observed 14 haplotypes at *PPD-H1* (1376 bp fragment) within the diverse GWAS
627 panel of world-wide origin (Table S10). Further, in the Diversity panel, we found 90 haplotypes
628 within an 898 bp fragment (Table S11). This number of haplotypes is impressive compared to
629 most re-sequencing studies in barley, but it is strikingly low compared to the 121 haplotypes
630 observed in only 266 accessions (that consisted of only 72 wild barleys) by Jones et al. (2008).
631 Theoretically, this difference could be due to the smaller fragment re-sequenced in our study
632 (898 bp considered) compared to 3508 bp re-sequenced by Jones et al. (2008). Other reasons
633 could be issues with SNP calling and haplotype assignment by Jones et al. (2008).

Comparing nucleotide diversity between wild (n = 942) and domesticated barley (n = 1110), we observed only a 9% loss of diversity at *PPD-H1*. This is also contrary to the findings of Jones et al. (2008) that reported a severe bottleneck and 22.5% loss of diversity within their 72 wild and 194 domesticated barleys. Due to the small sample size, diversity values by Jones et al. (2008) should be noted with caution. We have 61 wild barley accessions in common with Jones et al. (2008) (Table S3; Table S4). Thus, we are convinced that the lower number of 90 haplotypes in our study is robust. More haplotypes described by Jones et al. (2008) probably result from issues with SNP calling and haplotype assignment. Nevertheless, recent studies (Cuesta-Marcos et al. 2010; Russell et al. 2016) reported also lower haplotype numbers at *PPD-H1* even from the larger-sequence lengths and diverse samples including wild barleys than Jones et al. (2008).

In our study we used comparative numbers of wild and domesticated barleys and comparatively smaller loss in genetic diversity was observed (Jakob et al. 2014). One reason could be that several major photoperiod responsive haplotypes are shared among wild and domesticated genotypes and that probably led to the observation of low nucleotide diversity change. One other reason could be that the domesticated group is a mix of spring, facultative and winter types, different row and caryopsis types and was collected from a wide range of environments. Thus, diversity (= expected heterozygosity on random mating) will be pushed higher in the domesticated group as a result of the winter – spring, 2-rowed – 6-rowed and other polymorphisms. There may be a general point that selection that maintains polymorphism within the domesticated group, as here, will also maintain diversity or at least reduce the loss. So, we get a reverse of the usual *pre- post-domestication* pattern with less loss than under neutrality. Nevertheless, overall, we observed fewer haplotypes (N=27) in domesticated barley (landraces N=23; cultivars: N=17) compared to truly wild barley (N=71).

Interestingly, recent investigation in sorghum using sequencing of the archaeological samples of wild and domesticated sorghum of different historical periods revealed that the surge in

diversity occurred over time and the formation of a domestication bottleneck is probably a myth (Smith et al. 2019; Brown 2019).We observed within wild and domesticated barleys segregating sites exclusive to either group. Segregating sites exclusive to the wild barley indicates that there are many mutations in wild barleys that were possibly not selected in domesticates. However, mutations exclusive to the domesticated barleys suggest that these mutations probably occurred after the initial domestication and/or outside the natural distribution range. Since wild and domesticated barleys can coexist together in the farmer's fields, natural gene flow may alter the values of genetic diversity. Such cases are reported to be relatively rare and are unlikely to be important in nature (Abdel-Ghani et al. 2004; Russell et al. 2011; Hübner et al. 2012). However, we believe that the rate of re-introgression of wild barley alleles probably occurred more frequently than previously thought.

To broaden the genetic basis for barley improvement at *PPD-H1*, haplotype information from this study could be considered. Potentially beneficial haplotypes could be introgressed from wild barley into the elite background (Dempewolf et al. 2017: https://www.cwrdiversity.org/project/pre-breeding/). Gene editing will provide another opportunity.

5. Photoperiod non-responsiveness most likely originated post domestication and in the Fertile Crescent

Barley domestication history is complex. In recent years publications suggest multiple domestication events that led to the present domesticates (Kilian et al. 2006; Morrell and Clegg 2007; Dai et al. 2012; Zeng et al. 2018; Pourkheirandish et al. 2015). One of the major events that led to the adaptation of barley to wider areas is the evolution of non-responsive barley (*ppd-H1*). Two contrasting hypotheses about the origin and spread of non-responsive barleys were published: (I) Turner et al. (2005) and Cockram et al. (2007) suggested that photoperiod non-responsiveness originated in domesticated barley (*'post' domestication, outside of the Fertile Crescent, during the spread of barley cultivation towards Northern Europe*); and (II) Jones et

686 al. (2008) concluded that the non-responsive phenotype originated in wild barley from Iran
687 (*'pre' domestication*).
688 In our study we showed that the origin of photoperiod non-responsive haplotypes was derived
689 from photoperiod responsive haplotype H10, and that the mutation leading to non-responsive
690 types was only found in domesticated barley. Our combined data indicate a monophyletic
691 natural mutation that most likely occurred in a domesticated *btr1Btr2*-type, 2-rowed, facultative
692 barley in the Southern Levant ('*post domestication in the Fertile Crescent'*).
693 A likely scenario would be that a domesticated barley with a photoperiod responsive allele at
694 *PPD-H1* e.g. H4, such as the Masada barley, hybridized with a Desert-type, facultative, wild
695 barley harboring haplotype 10, in the Southern Levant, probably where H10 barley still grows
696 today (Fig. 7). The fully fertile F1 hybrid would be a 'domesticated' barley harboring, for
697 example, the *PPD-H1 – HvCEN* haplotype combination H10II or H10III. Their offspring would
698 later receive the natural mutation at SNP22, also in the Southern Levant, probably under
699 irrigated cultivation (if we assume that the non-responsive haplotype H1 would have negative
700 effects on barley in this hot and dry region). In a second scenario, a domesticated *btr1Btr2*-type
701 barley, which already contains H10 would directly receive the mutation at SNP22.

702 The ancient Masada sample is important in this context. Mascher et al. (2016) concluded that
703 this 6000-year-old sample was a domesticated (*btr1Btr2*), 2-rowed and hulled barley. We found
704 that it carried the *PPD-H1 – HvCEN* haplotype combination H4IV and was therefore
705 photoperiod responsive (H4) and *'early* flowering' (IV). Interestingly, this haplotype
706 combination was not found in the comprehensive collection of 2057 wild and domesticated
707 barleys. The phylogenetically and phenotypically closest *PPD-H1 – HvCEN* haplotype
708 combination found in the Diversity panel was H4II, and found in one wild barley from Turkey
709 (Northern Levant, near Gaziantep); in four landraces (3x Libya, 1x Georgia) and in nine
710 cultivars from six countries outside the Fertile Crescent (Table S4). At *HvCEN*, the progenitor

haplotype of IV is II, which was found in Israel in 8 wild barleys (also in the environmentally closest collection sites to Masada, Fig. 7); and in one landrace obtained from a market in Jerusalem in 1964. Haplotype 4 was found in Israel in 54 wild barley. Our data suggest that the Masada barley evolved from local populations in the Southern Levant and was not introduced from elsewhere.

The probable source of the Masada barley found in the Yoram Cave is located in the Ein Gedi oasis, about 17 km north of the cave (David 2015; Fig. 7). This is the largest and most important oasis in the Judean Desert with an annual water flow of about 3.5 million cubic meters. There, abundant archaeological remains attest for agricultural irrigation systems of agricultural terraces fed with spring water (Hadas 2012). Although the irrigation systems date to the Roman-Byzantine period (1st-6th century CE), the coexistence of nearby archaeological sites from the Chalcolithic period indicates that the spring was flowing at the same time. In addition, a review of historical sources attests to barley cultivation in this irrigation system during the Roman period (Hadas 2012). Considering this, the barley of the Yoram Cave was probably grown under irrigated conditions in the Ein Gedi oasis. Apparently, this is the only example of irrigated cereal fields in ancient Israel.

Similarly, the origin of non-responsive barley could have occurred under irrigated conditions in the Southern Levant. The mutation probably took place in domesticated barley but also a wild barley with H10 could have been mutated and then hybridized with a domesticated barley, under irrigated conditions in mixed stands.

Few photoperiod non-responsive haplotypes, but with a high frequency and found only in domesticated barley, indicate that they are probably of relatively recent origin. Although the 6000-year-old Masada sample was photoperiod responsive, this does not exclude the possibility that the photoperiod non-responsiveness originated earlier than 6000 years ago or that photoperiod responsive and non-responsive domesticated barley co-existed in the Southern

Levant in the past as today (Table S4). Based on von Bothmer at al. (2003), *"Barley cultivation reached Spain ca. 7,000 years BP (Before Present), N Africa and Ethiopia ca. 8,000 years BP and northern Europe ca. 6,000 years BP."* Further study is needed to shed more light on the age of non-responsive haplotypes.

Conclusion

We showed that the photoperiod non-responsive adaptation to long day spring conditions in Europe originated from Desert type wild barley (H10) in the Judean desert and involved the selection of one *de novo* mutation (SNP22). Haplotypes H1 and H2 increased in frequency during the spread of civilization out of the Fertile Crescent towards Northern Europe under higher selection pressure. Haplotype H2 probably originated also *de novo* (synonymous substitution) during this range extension. Our data suggest that the spring growth habit evolved from the facultative habit in the Southern Levant. Finally, we conclude that all *btr1*-type barley evolved from Desert type wild barley near the Dead Sea.

Supplementary Material

Supplementary material mentioned in the text, comprising 19 supplementary tables and 13 supplementary figures are available online.

Sequence availability

New sequence data from this article are deposited in GenBank Data library under accession numbers provided in Tables S12 and S13 (Supplementary Material online): KF309068-309171.

Acknowledgements

We thank the following for providing seeds and/ or DNA material: Federal *ex situ* Genbank of Germany in Gatersleben (IPK), Germany; U.S. Department of Agriculture, Agricultural Service (USDA-ARS); International Center for Agricultural Research in Dry Areas (ICARDA); Max

759 Planck Institute for Plant Breeding Research (MPIPZ), Cologne, Germany; Australian Winter
760 Cereals Collection, Tamworth, Australia; Institut für Pflanzenbau und Pflanzenzüchtung,
761 Braunschweig, Germany; Aegean Agricultural Research Institute, Izmir, Turkey; Swedish
762 University of Agricultural Sciences (SLU), Sweden; NPGBI-Karaj, Iran; Seed and Plant
763 Improvement Institute (SPII), Karaj, Iran; the Exbardiv consortium; M. Abbasi, S. Michel, M.
764 Ristow, M. von Korff, R. Fritsch, S. Jakob and F. Salamini. We thank Sigi Effgen, Jutta Schütze
765 (both MPIPZ), Kerstin Wolf, Heike Harms, Birgit Dubsky, Christiane Kehler, Heloise Giraud,
766 Peter Schreiber, Jürgen Marlow, Michael Grau (all IPK) and all gardeners and students involved
767 for excellent technical support. The authors thank B. Steffenson, A. Tondelli, M. von Korff, M.
768 Koorneef and F. Salamini for discussions. This work was supported by the German Research
769 Foundation (DFG) Priority Program SPP1530 (FKZ KI 1465/6-1); the ERA-PG-funded project
770 Exbardiv (http://www.erapg.org), GABI Future GENOBAR
771 (https://www.pflanzenforschung.de), the Leibniz Institute of Plant Genetics and Crop Research
772 (IPK), Gatersleben and the Max Planck Institute for Plant Breeding Research, Cologne.

773 **References**

774 Abdel-Ghani AH, Parzies HK, Omary A, Geiger HH. 2004. Estimating the outcrossing rate of
775 barley landraces and wild barley populations collected from ecologically different
776 regions of Jordan. *Theor Appl Genet.* 109:588–595.
777 Abdel-Ghani AH, Sharma R, Wabila C, Dhanagond S, Owais SJ, Duwayri MA, Al-Dalain
778 SA, Klukas C, Chen D, Lübberstedt T, et al. 2019. Genome-wide association mapping in
779 a diverse spring barley collection reveals the presence of QTL hotspots and candidate
780 genes for root and shoot architecture traits at seedling stage. *BMC Plant Biol.* 19:216.
781 Afsharyan NP, Sannemann W, Léon J, Ballvora A. 2020. Effect of epistasis and environment
782 on flowering time in barley reveals a novel flowering-delaying QTL allele. *J Exp Bot.*
783 71:893–906.
784 Alizadeh A. 2003. Excavations at the prehistoric mound of Chogha Bonut, Khuzestan, Iran: S

785 Alqudah AM, Sharma R, Pasam RK, Graner A, Kilian B, Schnurbusch T. 2014. Genetic
786 dissection of Photoperiod response based on GWAS of pre-anthesis phase duration in
787 spring barley. *PloS One* 9 (11)
788 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search
789 tool. *J Mol Biol*. 215:403–410.
790 Badr A, Muller K, Schafer-Pregl R, Rabey H El, Effgen S, Ibrahim HH, Pozzi C, Rohde W,
791 Salamini F. 2000. On the origin and domestication history of Barley (*Hordeum vulgare*).
792 *Mol Biol Evol.* 17:499-510.
793 Baloch AW, Baloch W, Baloch GA, Mangan B, et al. 2013. The effect of Ppd-H1 gene's SNP
794 48 marker on heading dates in natural populations of barley from Middle East. *Global*
795 *Journal of Biodiversity Science and Management* 3(2): 233-236.
796 Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific
797 phylogenies. *Mol Biol Evol* 16:37-48.
798 Bayer M, Rapazote-Flores P, Ganal M, Hedley PE, Macaulay M, Plieske J, Ramsay L, Russell
799 J, Shaw PD, Thomas W, Waugh R. 2017. Development and evaluation of a barley 50k
800 iSelect SNP array. *Front Plant Sci.* 8:1792
801 Brown TA. 2019. Is the domestication bottleneck a myth? *Nat Plants* 5:337–338.
802 Bustos-Korts D, Dawson IK, Russell J, Tondelli A, Guerra D, Ferrandi C, Strozzi F, Nicolazzi
803 EL, Molnar-Lang M, Ozkan H, et al. 2019. Exome sequences and multi-environment
804 field trials elucidate the genetic basis of adaptation in barley. *Plant J*. 99:1172–1191.
805 Campoli C, Drosse B, Searle I, Coupland G, Von Korff M. 2012. Functional characterisation
806 of HvCO1, the barley (Hordeum vulgare) flowering time ortholog of CONSTANS. Plant
807 J. 69:868–880.
808 Casas A, Djemel A, Ciudad F, Yahiaoui S, Ponce L, Contreras-Moreira B, Gracia MP, Lasa J,
809 Igartua E. 2011. *HvFT1* (*VrnH3*) drives latitudinal adaptation in Spanish barleys. *Theor*
810 *Appl Genet.* 122:1293-1304.
811 Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ. 2007.
812 Control of flowering time in temperate cereals: genes, domestication, and sustainable
813 productivity. *J Exp Bot.* 58:1231-1244.
814 Comadran J, Kilian B, Russell J, et al. 2012. Natural variation in a homolog of Antirrhinum
815 *CENTRORADIALIS* contributed to spring growth habit and environmental adaptation
816 in cultivated barley. *Nat Genet.* 44:1388-1392.
817 Cuesta-Marcos A, Szucs P, Close TJ, Filichkin T, Muehlbauer GJ, Smith KP, Hayes PM. 2010.
818 Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in

barley: Implications for association mapping in germplasm arrays varying in size and structure. *BMC Genomics* 11:707.

Dai F, Nevo E, Wu DZ, Comadran J, Zhou MX, Qiu L, Chen ZH, Beiles A, Chen GX, Zhang GP. 2012. Tibet is one of the centers of domestication of cultivated barley. *Proc Natl Acad Sci U S* A. 109:16969-16973.

Dawson IK, Russell J, Powell W, Steffenson B, Thomas WTB, Waugh R. 2015. Barley: A translational model for adaptation to climate change. *New Phytol.* 206:913–931.

Dempewolf H, Baute G, Anderson J, Kilian B, Smith C, Guarino L. 2017. Past and Future Use of Wild Relatives in Crop Breeding. *Crop Sci.* 57:1070-1082

Göransson M, Hallsson JH, Lillemo M, Orabi J, Backes G, Jahoor A, Hermannsson J, Christerson T, Tuvesson S, Gertsson B, et al. 2019. Identification of ideal allele combinations for the adaptation of spring barley to northern latitudes. *Front Plant Sci.* 10:542.

Hadas G. 2012. Ancient agricultural irrigation systems in the oasis of Ein Gedi, Dead Sea, *Israel Journal of Arid Environments* 86 (2012) doi:10.1016/j.jaridenv.2011.08.015

Hall TA. 1999. BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT. Nucleic Acids Symp. Ser.

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

Haseneyer G, Stracke S, Paul C, Einfeldt C, Broda A, Piepho HP, Graner A, Geiger HH. 2010. Population structure and phenotypic variation of a spring barley world collection set up for association studies. *Plant Breeding.* 129:271-279.

Herzig P, Maurer A, Draba V, Sharma R, Draicchio F, Bull H, Milne L, Thomas WTB, Flavell AJ, Pillen K. 2018. Contrasting genetic regulation of plant development in wild barley grown in two european environments revealed by nested association mapping. *J Exp Bot.* 69:1517-1531.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 25:1965–1978

Hübner S, Bdolach E, Ein-Gedy S, Schmid KJ, Korol A, Fridman E. 2013. Phenotypic landscapes: Phenological patterns in wild and cultivated barley. *J Evol Biol.* 26:163–174.

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.

852 Hübner S. et al. 2012. Islands and streams: clusters and gene flow in wild barley populations
853 from the Levant. *Mol Ecol.* 21:1115–1129
854 Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol*
855 *Biol Evol*. 23:254-267.
856 Jakob SS, Rodder D, Engler JO, Shaaf S, Ozkan H, Blattner FR, Kilian B. 2014. Evolutionary
857 history of wild barley (*Hordeum vulgare* subsp. *spontaneum*) analyzed using multilocus
858 sequence data and paleodistribution modeling. *Genome Biol Evol.* 6:685–702.
859 Jones H, Leigh FJ, Mackay I, Bower MA, Smith LM, Charles MP, Jones G, Jones MK, Brown
860 TA, Powell W. 2008. Population-based resequencing reveals that the flowering time
861 adaptation of cultivated barley originated east of the Fertile Crescent. *Mol Biol Evol*.
862 25:2211-2219.
863 Kikuchi R, Kawahigashi H, Ando T, Tonooka T, Handa H. 2009. Molecular and Functional
864 Characterization of PEBP genes in barley reveal the diversification of their roles in
865 flowering. *Plant Physiol.* 149:1341–1353.
866 Kilian B, Özkan H, Kohl J, Haseler AV, Barale F, Deusch O, Brandolini A, Yucel C, Martin
867 W, Salamini F. 2006. Haplotype structure at seven barley genes: relevance to gene pool
868 bottlenecks, phylogeny of ear type and site of barley domestication. *Mol Genet Genom.*
869 276:230-241.
870 Lancashire PD, Bleiholder H, Boom T Van Den, Langeluddecke P, Stauss R, Weber E,
871 Witzenberger A. 1991. A uniform decimal code for growth stages of crops and weeds.
872 *Ann Appl Biol*. 119:561–601.
873 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping
874 and population genetical parameter estimation from sequencing data. Bioinformatics
875 27:2987–2993.
876 Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA
877 polymorphism data. *Bioinformatics* 25:1451-1452.
878 Long NV, Dolstra O, Malosetti M, Kilian B, Graner A, Visser RGF, van der Linden CG.
879 2013. Association mapping of salt tolerance in barley (*Hordeum vulgare* L.). *Theor Appl*
880 *Genet.* 126:2335–2351.
881 Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V,
882 Dockter C, Hedley PE, Russell J, et al. 2017. A chromosome conformation capture
883 ordered sequence of the barley genome. *Nature* 544:427–433.

884 Mascher M, Schuenemann VJ, Davidovich U, Marom N, Himmelbach A, Hübner S, Korol A,
885 David M, Reiter E, Riehl S, et al. 2016. Genomic analysis of 6,000-year-old cultivated
886 grain illuminates the domestication history of barley. *Nat Genet.* 48:1089-1093.
887 Maurer A, Draba V, Jiang Y, Schnaithmann F, Sharma R, Schumann E, Kilian B, Reif JC,
888 Pillen K (2015) Modelling the genetic architecture of flowering time control in barley
889 through nested association mapping. *BMC Genomics* 16 10.1186/s12864-015-1459-7.
890 McMaster GS, Moragues M. 2019. Crop development related to temperature and photoperiod.
891 In: Savin R., Slafer G. (eds) Crop Science. Encyclopedia of Sustainability Science and
892 Technology Series. Springer, New York, NY.
893 Monat C, Padmarasu S, Lux T, Wicker T, Gundlach H, Himmelbach A, Ens J, Li C,
894 Muehlbauer GJ, Schulman AH, et al. 2019. TRITEX: Chromosome-scale sequence
895 assembly of Triticeae genomes with open-source tools. *Genome Biol*. 20:284.
896 Monteagudo A, Igartua E, Contreras-Moreira B, Gracia MP, Ramos J, Karsai I, Casas AM.
897 2019. Fine-tuning of the flowering time control in winter barley: The importance of
898 HvOS2 and HvVRN2 in non-inductive conditions. *BMC Plant Biol*. 19:113.
899 Morrell PL, Clegg MT. 2007. Genetic evidence for a second domestication of barley (*Hordeum*
900 *vulgare*) east of the Fertile Crescent. *Proc Natl Acad Sci U S A*. 104:3289-3294.
901 Nitcher R, Distelfeld A, Tan C, Yan L, Dubcovsky J. 2013. Increased copy number at the
902 *HvFT1* locus is associated with accelerated flowering time in barley. *Molecular*
903 *Genetics Genomics* 288:261-275.
904 Pasam RK, Sharma R, Walther A, Özkan H, Graner A, Kilian B. 2014. Genetic diversity and
905 population structure in a legacy collection of spring barley landraces adapted to a wide
906 range of climates. *PLoS One* 9:1–29.
907 Pasam, RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A. 2012.
908 Genome-wide association studies for agronomical traits in a world wide spring barley
909 collection. *BMC Plant Biol* 12:16.
910 Piperno DR, Weiss E, Holst I, Nadel D. 2004. Processing of wild cereal grains in the Upper
911 Palaeolithic revealed by starch grain analysis. *Nature* 430:670–673.
912 Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, Sameri M, Azhaguvel P, Sakuma
913 S, Dhanagond S, Sharma R, et al. 2015. Evolution of the Grain Dispersal System in
914 Barley. *Cell* 162:527–539.
915 Qian C, Yan X, Shi Y, Yin H, Chang Y, Chen J, Ingvarsson PK, Nevo E, Ma XF. 2020.
916 Adaptive signals of flowering time pathways in wild barley from Israel over 28
917 generations. *Heredity* 124:62–76.

R Core Team. 2018. A Language and Environment for Statistical Computing. R Found. Stat. Comput.

Renaud G, Stenzel U, Kelso J. 2014. LeeHom: Adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Re*s. 42:e141.

Russell J, Mascher M, Dawson IK, Kyriakidis S, Calixto C, Freund F, Bayer M, Milne I, Marshall-Griffiths T, Heinen S, et al. 2016. Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat Genet*. 48:1024–1030.

Saisho D, Ishii M, Hori K, Sato K. 2011. Natural variation of barley vernalization requirements: implication of quantitative variation of winter growth habit as an adaptive trait in East Asia. *Plant Cell Physiol*. 52:775-784.

Sharma R, Draicchio F, Bull H, Herzig P, Maurer A, Pillen K, Thomas W, Flavell AJ. 2018. Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding. *J Exp Bot*. 69:3811–3822.

Smith O, Nicholson W V., Kistler L, Mace E, Clapham A, Rose P, Stevens C, Ware R, Samavedam S, Barker G, et al. 2019. A domestication history of dynamic adaptation and genomic deterioration in Sorghum. *Nat Plants* 5:369–379.

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* 10:1–12.

Steffen A, Fischer A, Staiger D. 2014. Determination of photoperiodic flowering time control in Arabidopsis and barley. In: Methods in Molecular Biology. p. 285–295.

Stracke S, Haseneyer G, Veyrieras JB, Geiger H, Sauer S, Graner A, Piepho HP. 2009. Association mapping reveals gene action and interactions in the determination of flowering time in barley. *Theor Appl Genet.* 118:259-273.

Takahashi R, Hayashi J, Yasuda S, Hiura U. 1963. Characteristics of the wild and cultivated barleys from Afghanistan and its neighboring regions. Ber. Ohara Inst. Landw Biol., Okayama Univ. 12:1-23.

Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS. 2004. Selection versus demography: A multilocus investigation of the domestication process in Maize. *Mol Biol Evol*. 21:1214–1225.

Tondelli A, Xu X, Moragues M, Sharma R, Schnaithmann F, Ingvardsen C, Manninen O, Comadran J, Russell J, Waugh R, et al. 2013. Structural and temporal variation in genetic diversity of European spring two-row barley cultivars and association mapping of

quantitative traits. *Plant Genome* 6:1–14.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310:1031-1034.

von Bothmer, R, Sato K, Komatsuda T, Yasuda SFG. 2003. The domestication of cultivated barley. In: R. von Bothmer, Th. van Hintum, H. Knüpffer and K. Sato (eds), Diversity in Barley (*Hordeum vulgare*). In: TvH Roland von Bothmer, Helmut Knüpffer and Kazuhiro Sato, editor. Diversity in Barley (Hordeum vulgare): Elsevier B.V. p. 9-27.

Weiss E, Wetterstrom W, Nadel D, Bar-Yosef O. 2004. The broad spectrum revisited: Evidence from plant remains. *Proc Natl Acad Sci U S A.* 101:9551–9555.

Wiegmann M, Maurer A, Pham A, March TJ, Al-Abdallat A, Thomas WTB, Bull HJ, Shahid M, Eglinton J, Baum M, Flavell AJ, Tester M, Pillen K. 2019. Barley yield formation under abiotic stress depends on the interplay between flowering time genes and environmental cues. Sci Rep 9:6397 https://doi.org/10.1038/s41598-019-42673-1

Xu X, Sharma R, Tondelli A, Russell J, Comadran J, Schnaithmann F, Pillen K, Kilian B, Cattivelli L, Thomas WTB, et al. 2018. Genome-Wide Association Analysis of Grain Yield-Associated Traits in a Pan-European Barley Cultivar Collection. *Plant Genome* 11:170073

Zeng X, Guo Y, Xu Q, Mascher M, Guo G, Li Shuaicheng, Mao L, Liu Q, Xia Z, Zhou J, et al. 2018. Origin and evolution of qingke barley in Tibet. *Nat Commun*. 9:5433.

Zohary et al. 2013. Domestication of Plants in the Old World; Fourth Edition: Oxford Univ. Press, Oxford.

974 **Figure legends**

975 **Fig. 1:** Phenotypic distribution of days to heading (Hd) over four multi-location field trials.
976 Violin plots of Hd (in days) across locations.

977 **Fig. 2:** Manhattan plots of days to heading (Hd) from four trial sites are displayed. The
978 horizontal red line shows the significance threshold based on $-\log_{10}P=4.0$. (a) Germany; (b)
979 USA; (c) Turkey; (d) Syria. Important co-localized candidate genes are indicated.

980 **Fig. 3:** Schematic overview of *PPD-H1* gene structure and SNPs. **(a)** The *PPD-H1* gene
981 consists of 8 exons. Conserved domains are indicated (Pseudo-receiver and CCT). SNP
982 positions are shown. The 50K SNP chip physical positions are given as in Bayer et al. (2017).
983 Details are provided in Table S7; bold=significant SNPs). The locations of SNP22 and SNP48
984 are indicated. Primer-binding sites are shown by arrows (P5F + P5R; PP05 + PP04). **(b)** SNPs
985 based on re-sequencing at *PPD-H1* in the GWAS panel and corresponding haplotypes (and
986 their frequencies in brackets) are provided.

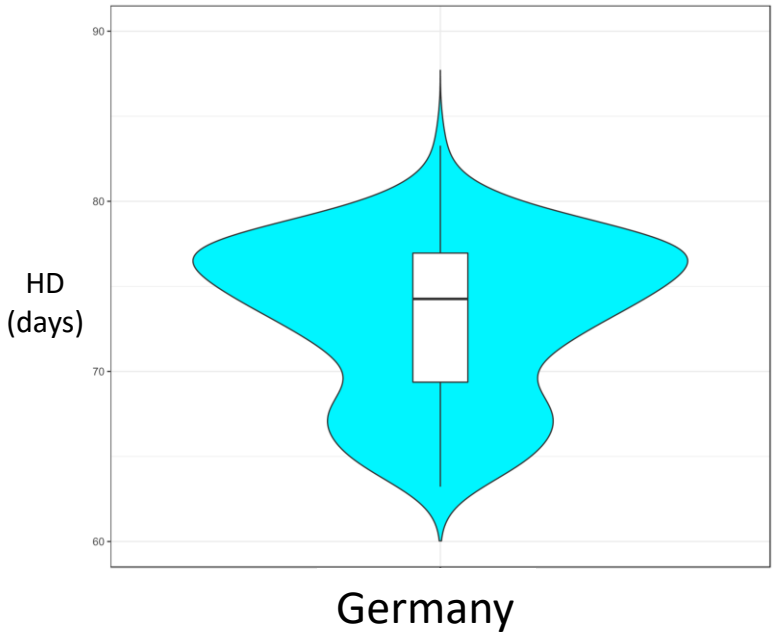
987 **Fig. 4:** Comparison of phenotypic response (in days to heading at BBCH55) to photoperiod for
988 41 genotypes grown under long and short-day conditions. Differences in HD under long and
989 short conditions are smaller for non-responsive (*ppd-H1*) genotypes compared to photoperiod
990 responsive (*Ppd-H1*) genotypes. This provides evidence that haplotype H10 containing
991 genotypes are photoperiod responsive. Boxplots are based on phenotypic BLUEs of genotypes.

992 **Fig. 5:** Allele mining at *PPD-H1*. Median-joining network derived from re-sequenced DNA
993 haplotypes of 2057 geo-referenced barley genotypes. 90 haplotypes were found and are
994 represented by arbitrarily given roman numerals. Circle sizes correspond to the frequency of
995 that particular haplotype. Red, haplotype found in wild barley; dark blue, domesticated barley;
996 yellow, *H. vulgare agriocrithon*. Distance in bp between haplotypes is indicated by Arabic
997 numerals and visible at higher magnification. Eight haplotypes were shared among wild and

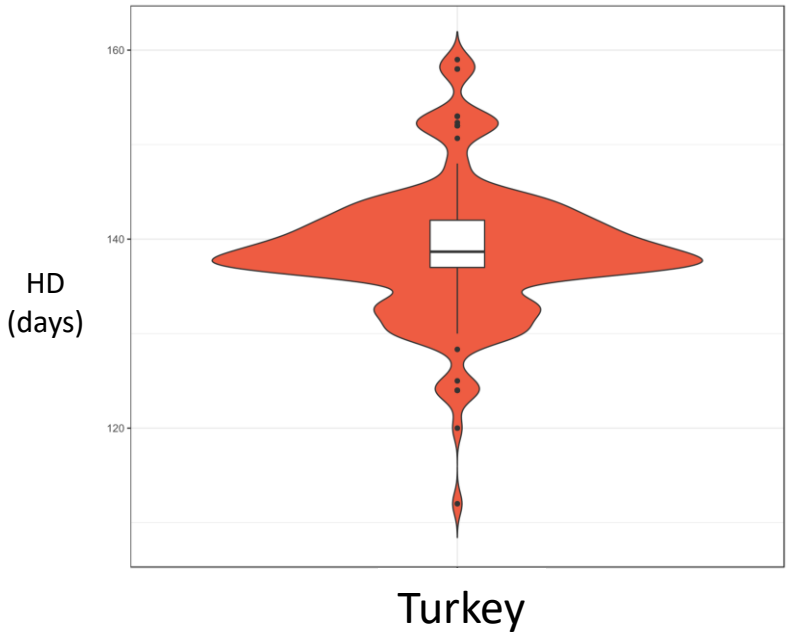
domesticated barleys. Photoperiod non-responsive (*ppd-H1*) haplotypes (H1, H2, H47, H55 and H56) are indicated. Black dots indicate median vectors. Numbers within circles correspond to the number of individuals carrying that haplotype.

Fig. 6: Principal component analysis (PCA) biplot of bioclimatic variables for collection sites of 1375 genotypes (942 wild and 432 landrace barleys, and the Masada sample). **a)** non-responsive (H1, H2) and responsive groups are separated. Two groups of H2 collection sites are visible. H10.Wild.Israel: wild barley from Israel containing haplotype; H10.Wild.Iran: wild barley from Iran containing haplotype 10. **b)** collection site characteristics of non-responsive barley (.NE – landraces from the Near East) and H4 containing landrace barley are highlighted. **c)** Zooming in on the characteristics of the collection sites of extant wild barley from Israel compared to the collection sites of the Masada sample and landraces containing H4. Wild barley yieldcontaining six haplotypes are indicated. Large circles indicate the median PC projection for each haplotype. See Tables S16, S17 and Figure S9 for more information.

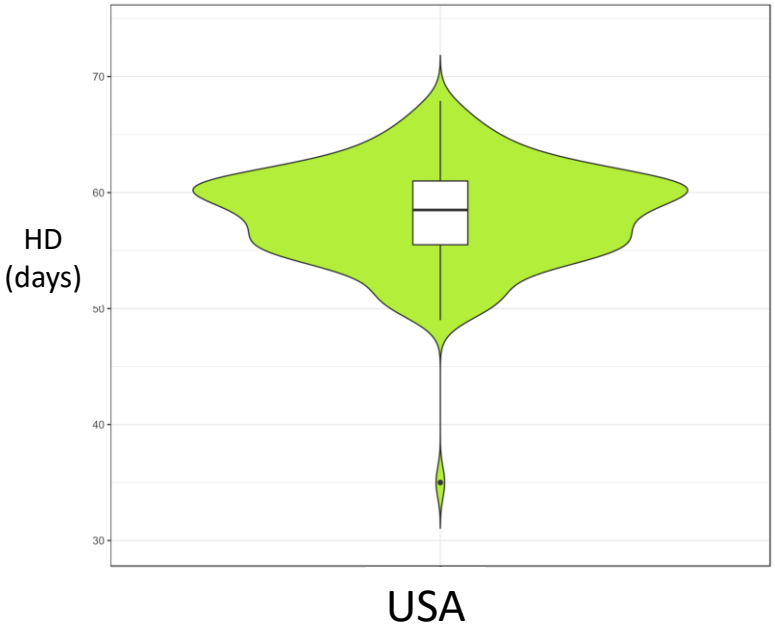
Fig. 7: Distribution of *PPD-H1* haplotype 10 in wild barley from Israel, relevant excavation sites, and the environmentally closest collections sites to Masada. a) haplotype 10 collection sites are indicated by black dots. b) **Red dot** – Kidron stream, the environmentally overall most similar collection site of extant wild barley to Masada. FT643 the closest wild barley to *btr1* based on Pourkheirandish et al. (2015) was collected here; **Green dot** – Almog, the environmentally second closest and also the closest *PPD-H1* haplotype 10 collection site to Masada; **Blue dot** – Neomi, the environmentally third closest collection site to Masada. FT013, FT015 and FT016 from this collection site were also among the closest wild barleys to *btr1*. Star – location of Jerusalem; yellow triangles – excavation sites (Ohalo II, Masada, Ein Gedi). See Tables S4, S18).



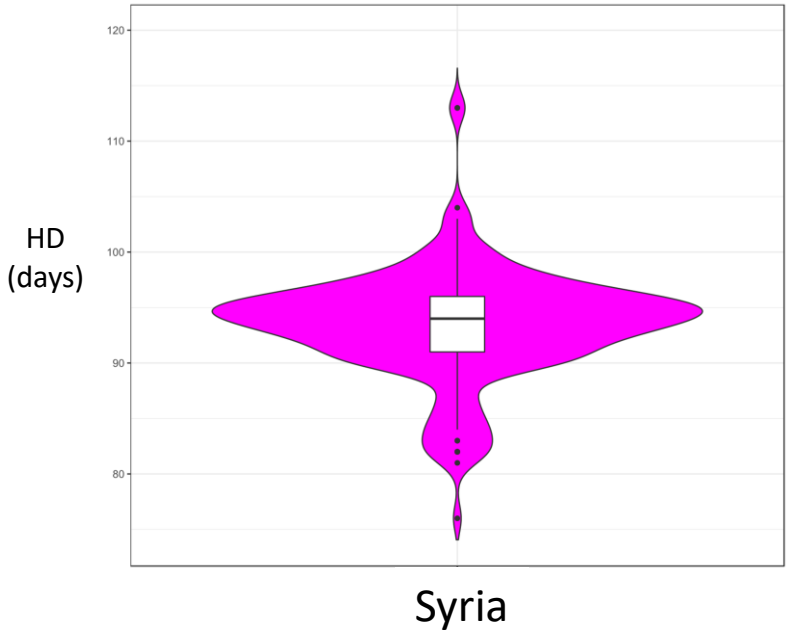
Germany



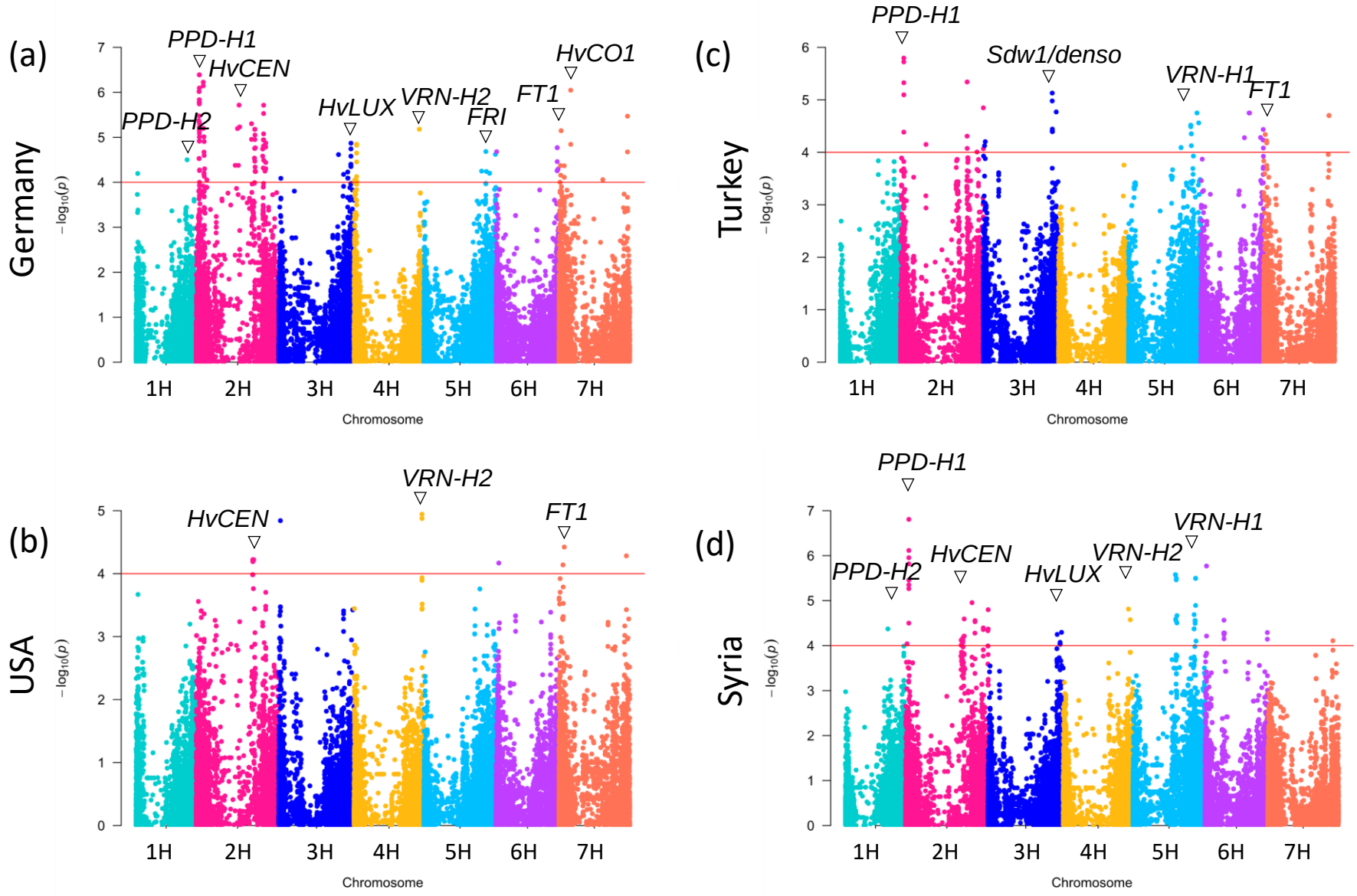
Turkey

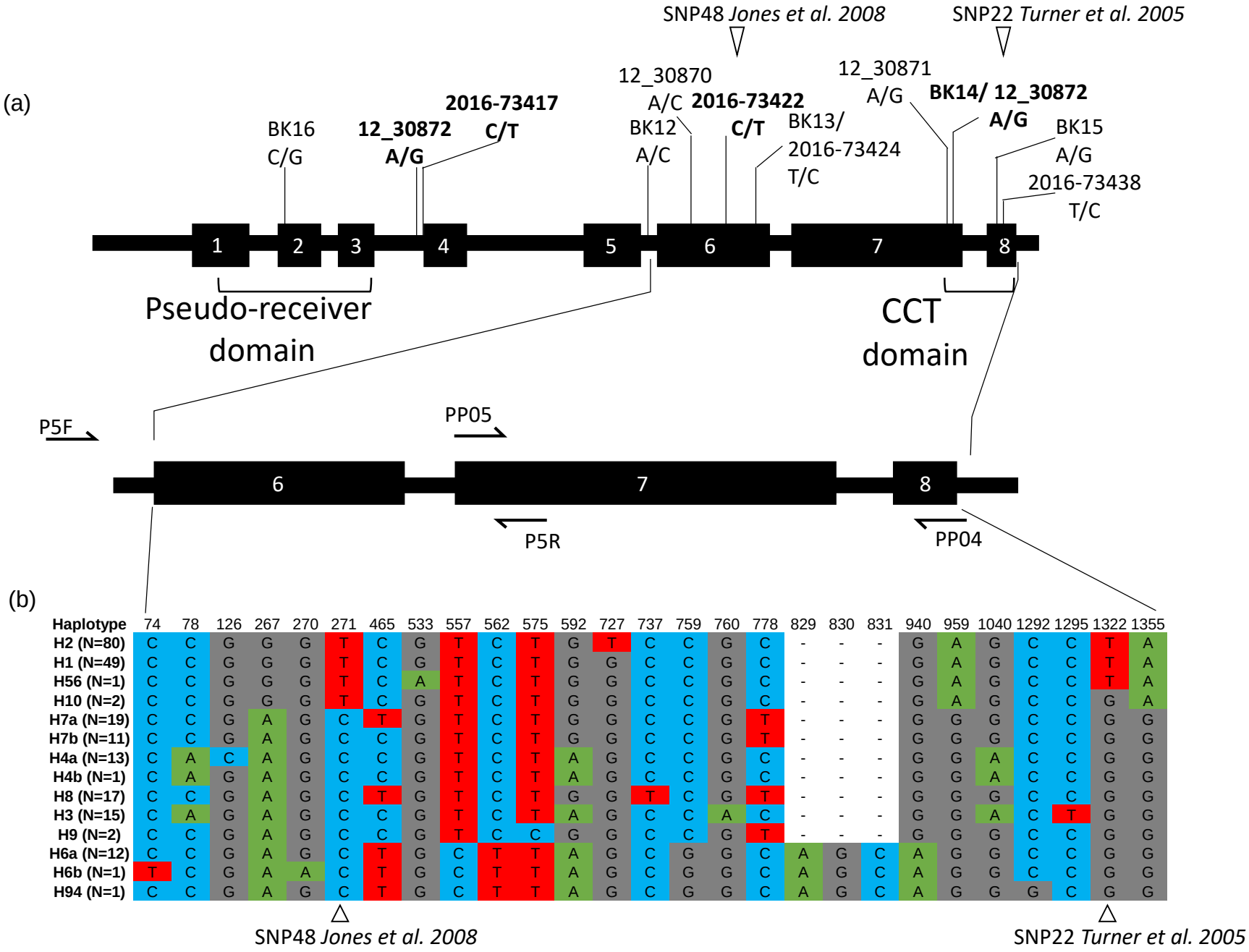


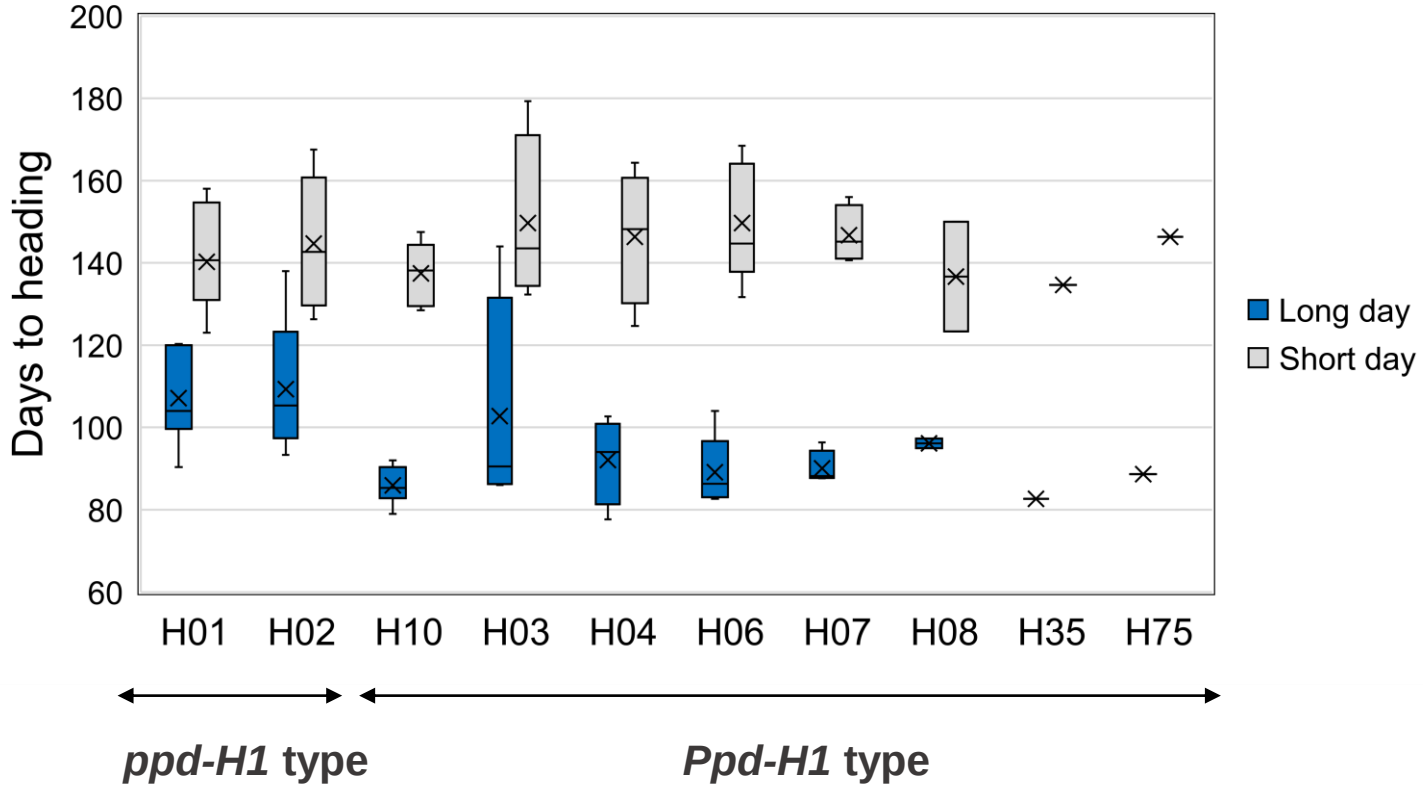
USA

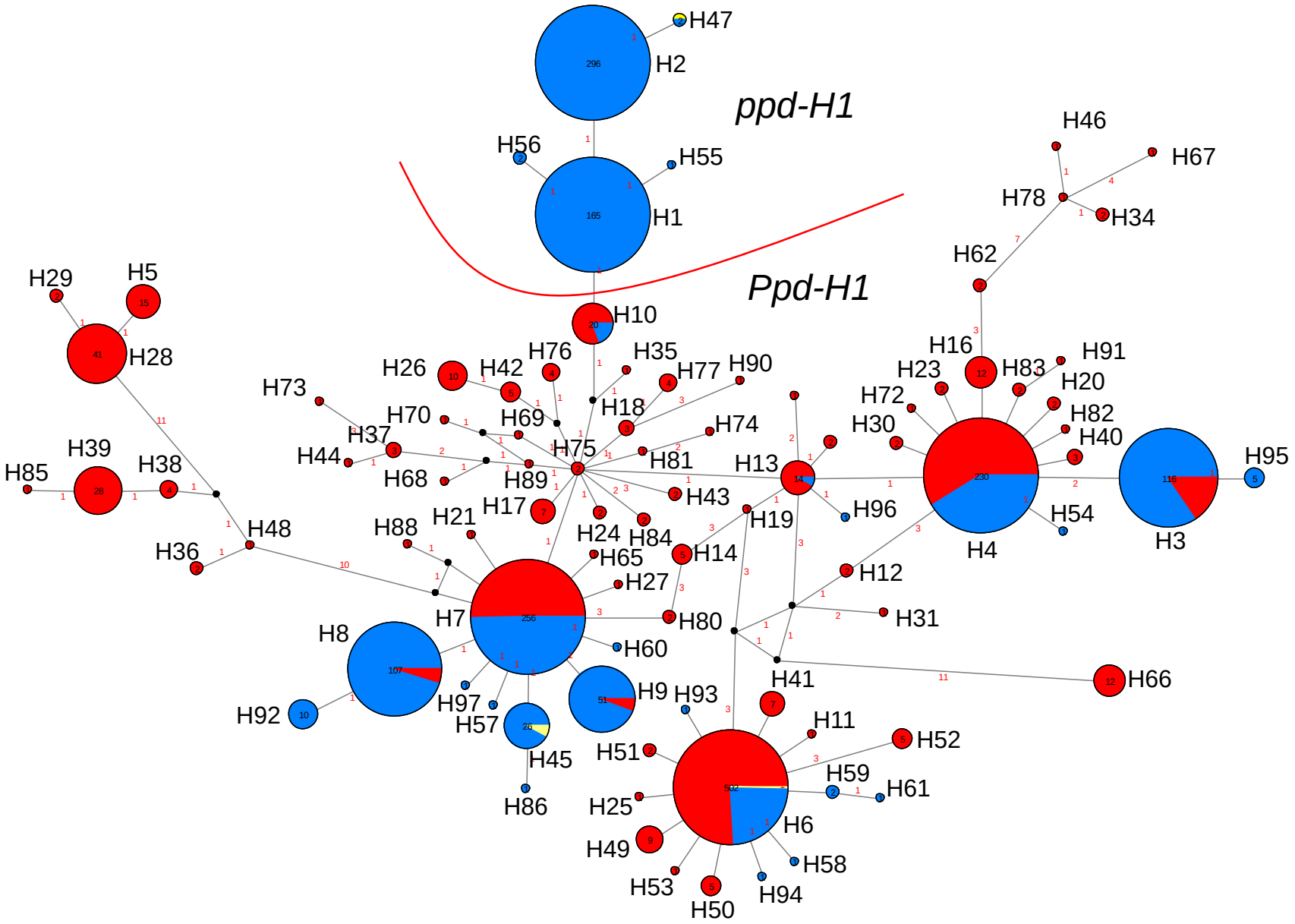


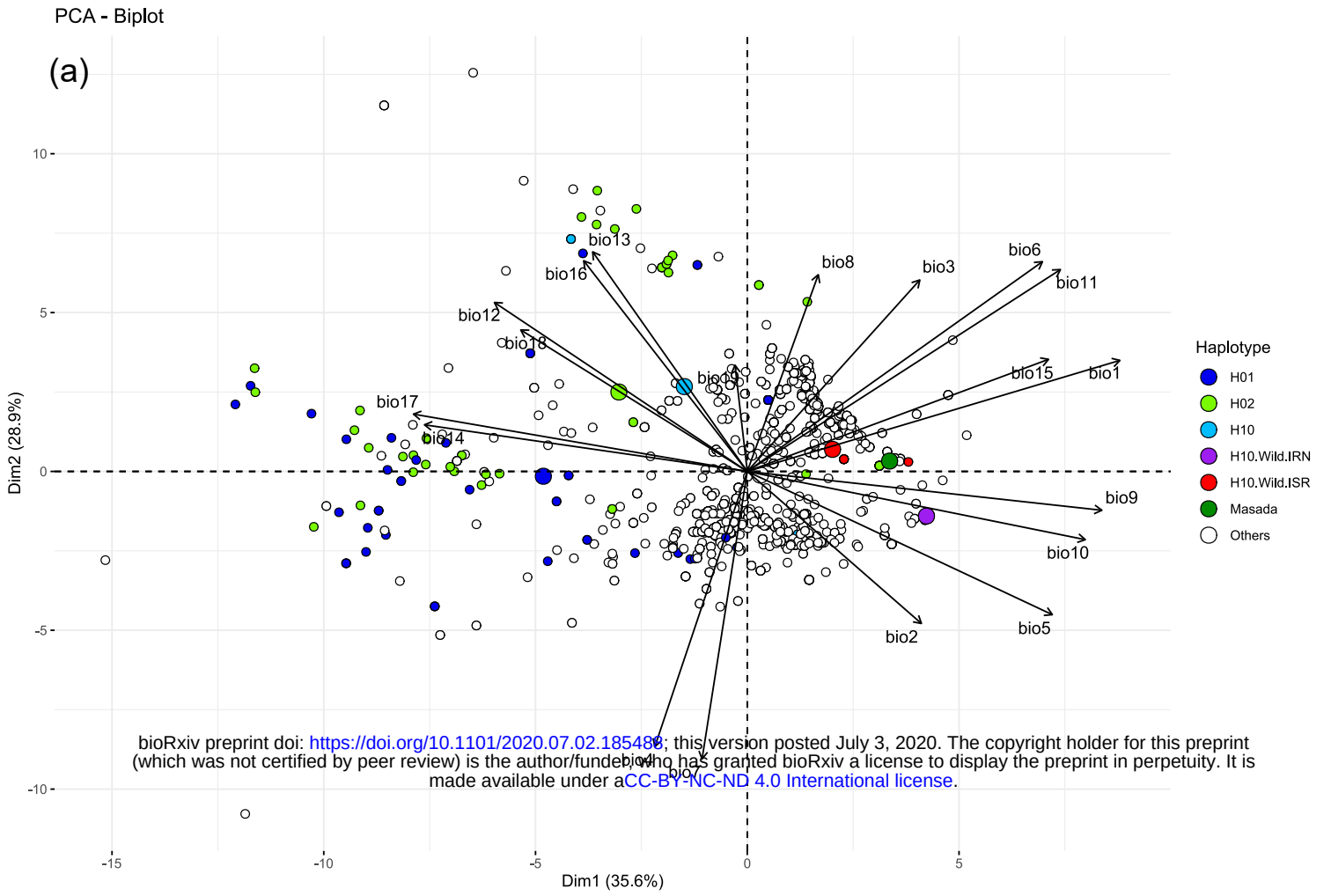
Syria

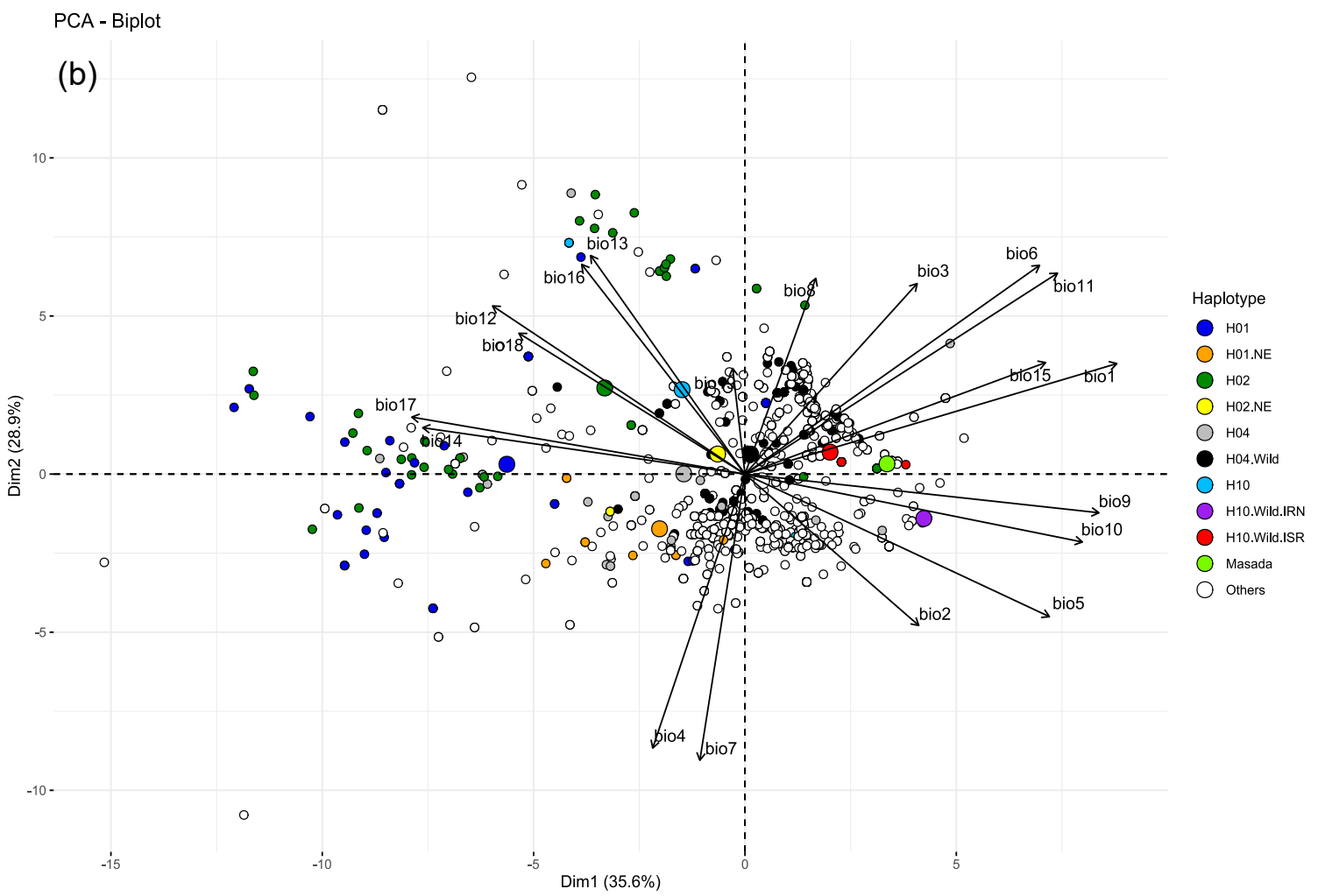

(a)
Germany
PPD-H1
HvCEN
PPD-H2
HvLUX
VRN-H2
FRI
FT1
HvCO1
(b)
USA
VRN-H2
HvCEN
FT1
(c)
Turkey
PPD-H1
Sdw1/denso
VRN-H1
FT1
(d)
Syria
PPD-H1
PPD-H2
HvCEN
HvLUX
VRN-H2
VRN-H1
−log10(p)
1H
2H
3H
4H
5H
6H
7H
Chromosome

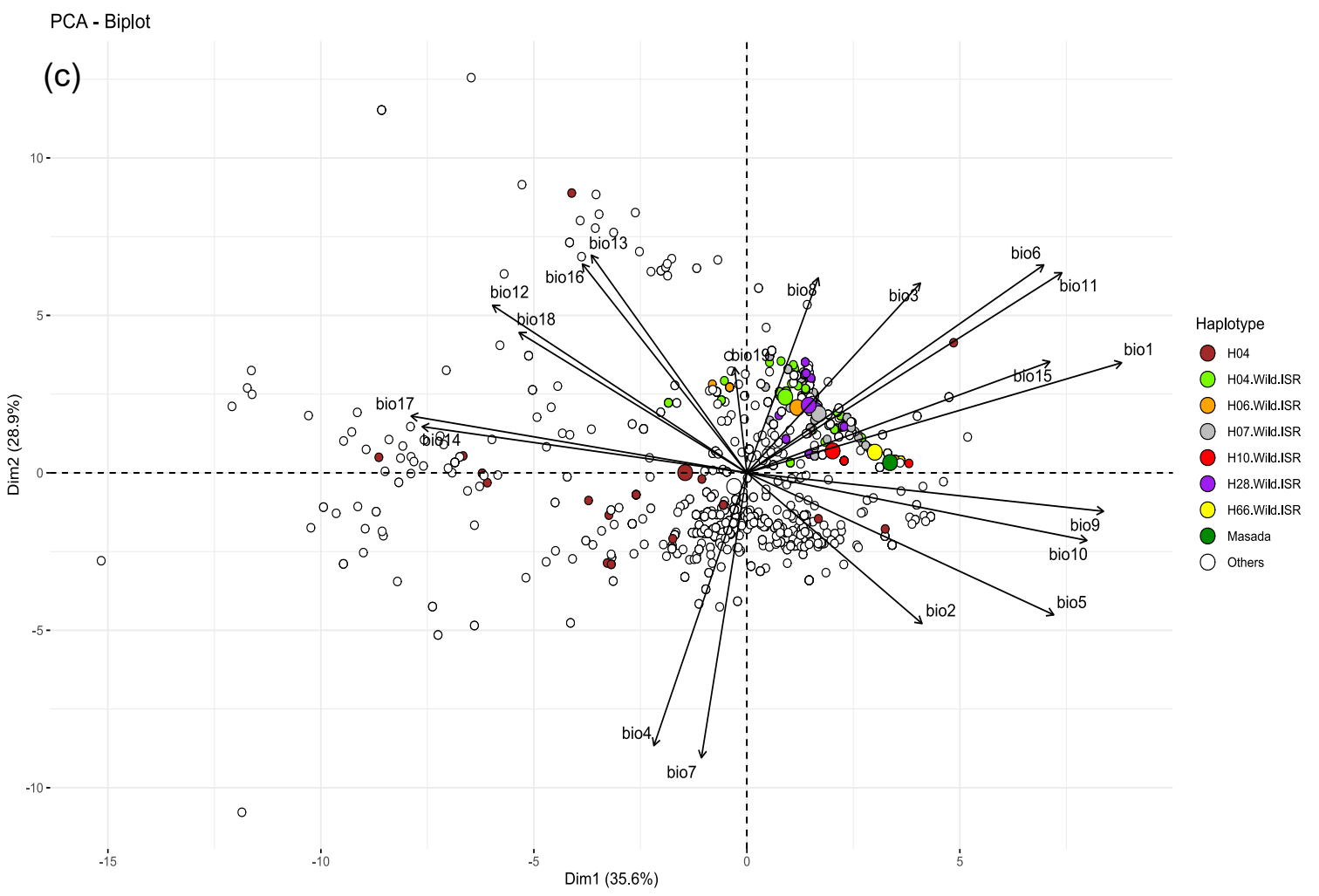

(a)
SNP48 Jones et al. 2008
SNP22 Turner et al. 2005
12_30870
A/C
2016-73422
C/T
12_30871
A/G
BK14/ 12_30872
A/G
BK16
C/G
12_30872
A/G
2016-73417
C/T
BK12
A/C
BK13/
2016-73424
T/C
BK15
A/G
2016-73438
T/C
1
2
3
4
5
6
7
8
Pseudo-receiver
domain
CCT
domain
P5F
PP05
6
7
8
P5R
PP04
(b)
Haplotype 74 78 126 267 270 271 465 533 557 562 575 592 727 737 759 760 778 829 830 831 940 959 1040 1292 1295 1322 1355
H2 (N=80) C C G G G T C G T C T G T C C G C - - - G A G C C T A
H1 (N=49) C C G G G T C G T C T G G C C G C - - - G A G C C T A
H56 (N=1) C C G G G T C A T C T G G C C G C - - - G A G C C T A
H10 (N=2) C C G G G T C G T C T G G C C G C - - - G A G C C G A
H7a (N=19) C C G A G C T G T C T G G C C G T - - - G G G C C G G
H7b (N=11) C C G A G C C G T C T G G C C G T - - - G G G C C G G
H4a (N=13) C A C A G C C G T C T A G C C G C - - - G G A C C G G
H4b (N=1) C A G A G C C G T C T A G C C G C - - - G G A C C G G
H8 (N=17) C C G A G C T G T C T G G T C G T - - - G G G C C G G
H3 (N=15) C A G A G C C G T C T A G C C A C - - - G G A C T G G
H9 (N=2) C C G A G C C G T C C G G C C G T - - - G G G C C G G
H6a (N=12) C C G A G C T G C T T A G C G G C A G C A G G C C G G
H6b (N=1) T C G A A C T G C T T A G C G G C A G C A G G C C G G
H94 (N=1) C C G A G C T G C T T A G C G G C A G C A G G G C G G
SNP48 Jones et al. 2008
SNP22 Turner et al. 2005

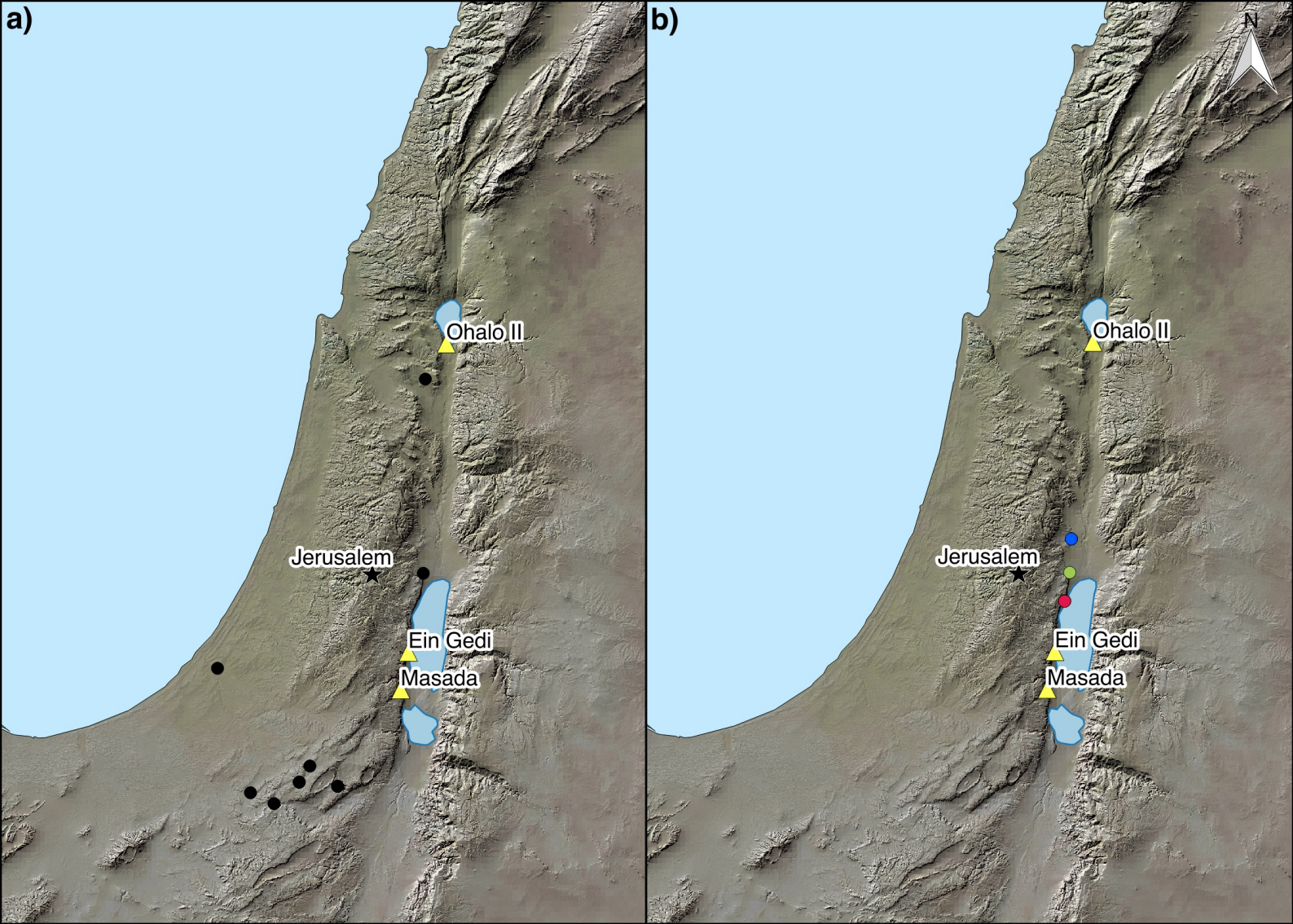

200
180
160
140
120
100
80
60
Days to heading
H01
H02
H10
H03
H04
H06
H07
H08
H35
H75
Long day
Short day
ppd-H1 type
Ppd-H1 type


ppd-H1
Ppd-H1
H47
H2
296
H56
H55
H1
165
H46
H67
H78
H34
H62
H29
H5
H28
41
H10
20
H16
H91
H23
H83
H20
H72
H82
H30
H40
H95
H26
H42
H76
H35
H77
H90
H73
H18
H74
H70
H69
H37
H75
H81
H44
H68
H89
H13
H43
H39
H38
H85
28
H17
H96
H54
H4
H3
H21
H24
H84
H88
H65
H14
H19
H48
H36
H12
H27
H7
256
H80
H31
H8
107
H60
H66
H41
H92
H97
H9
H93
H57
H45
H11
H52
H51
H59
H61
H25
H6
H49
H58
H53
H50
H94
H86








a)
Ohalo II
Jerusalem
Ein Gedi
Masada
b)
N
Ohalo II
Jerusalem
Ein Gedi
Masada