Chapter 6
Hordeum

Peter L. Morrell and Michael T. Clegg

6.1 Basic Botany of the Species

Wild barley (*Hordeum vulgare* ssp. *spontaneum*), the progenitor of cultivated barley (*H. vulgare* ssp. *vulgare*), is an annual, diploid grass species with a chromosome number of $n = 7$ and an estimated rate of self-fertilization of 98%. *Hordeum* belongs to the *Triticeae* tribe of the grass family that also includes wheat (*Triticum*), and together, these two grasses account for roughly 34% of the world's cereal production, although barley accounts for only about 6% compared to 28% for wheat. Both species are believed to be among the world's first cereal species to be domesticated by Neolithic humans.

The principle uses of barley are for malting in beer and whiskey production and for animal feed. According to a 2007 FAO report, Russia is the largest producer of barley followed by Canada: *http://faostat.fao.org/site/339/default.aspx*.

6.2 Taxonomy and Distribution

6.2.1 Taxonomy

The genus *Hordeum* includes roughly 30 species that could be considered as a potential resource for genetic improvement of barley. Among these species, only *H. bulbosum* is generally capable of producing fertile progeny when crossed to the domesticated species, *H. vulgare* ssp. *vulgare* (Bothmer et al. 1983). Thus, *H. bulbosum* is regarded as the only species in the secondary gene pool of cultivated barley (Pickering and Johnston 2005). *Hordeum bulbosum* is an obligately outcrossing (self-incompatible) (Lundqvist 1962) species with a wide geographic distribution.

Wild barley, *H. vulgare* ssp. *spontaneum*, is the progenitor of cultivated barley and is fully interfertile with *H. vulgare* ssp. *vulgare* (hereafter referred to as cultivated barley or simply as "barley"). As indicated by the nomenclature, they are members of the same biological species.

6.2.2 Distribution

The natural distribution of wild barley ranges from the Mediterranean portion of the Middle East, across the Zagros Mountains, and into adjacent Southwest Asia, a distance east to west of ~3,500 km. The eastern and western portions of the species range are relatively low-elevation regions; wild barley has limited cold tolerance and is rare above 1,500-m elevation (Zohary and Hopf 2000). Wild barley populations are abundant in the western portion of the range, but the species is rare at higher elevations (e.g., parts of the Zagros and the continental plateau in Turkey and Iran) and sporadic in the eastern portion of the range (Zohary and Hopf 2000). The Zagros Mountains, together with a series of smaller mountain ranges, trending northwest to southeast, roughly bisect the range; many mountain peaks rise above 3,000 m and the tallest single peak is 4,500 m. Thus, the Zagros Mountains represent a significant disruption in the natural range of wild

P.L. Morrell (✉)
Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA
e-mail: pmorrell@umn.edu

C. Kole (ed.), *Wild Crop Relatives: Genomic and Breeding Resources, Cereals*,
DOI 10.1007/978-3-642-14228-4_6,

barley, and the mountains present a substantial barrier to the movement of animals that serve as potential seed dispersers. Seed dispersion is facilitated by the spikelets of wild barley that have long, barbed lemma awns, well suited for attachment to animal fur (Bothmer et al. 1995). Moreover, a number of large animal species occur natively within the range of wild barley, including goats, boars, deer, and gazelles, providing a means for the dispersion of wild barley propagules.

Hordeum bulbosum is a perennial grass that occurs in a variety of habitats in the Mediterranean region (Bothmer et al. 1995). The geographic range of *H. bulbosum* is larger than that of *H. vulgare* ssp. *spontaneum*, extending around the Mediterranean basin, including North Africa and southern Europe, and east into western Central Asia, including present day Afghanistan and Tajikistan (Bothmer et al. 1995; Pickering and Johnston 2005). *Hordeum bulbosum* occurs as both an $n = 7$ diploid and $n = 14$ autotetraploid. Diploid forms of *H. bulbosum* predominate west of Greece and Egypt, while tetraploid forms are more common to the east (Bothmer et al. 1995).

6.3 Utilization of Transpecific Genetic Resources

Hordeum bulbosum has many desirable traits that could be used for barley improvement (Johnston et al. 2009). These include disease and pest resistance and tolerance for abiotic stress: traits that have immediate value in current barley breeding programs. Other traits from *H. bulbosum* such as perenniality and self-incompatibility have the potential to alter the crop for more sustainable production and maintenance of genetic diversity (Johnston et al. 2009).

Although direct crosses between *H. vulgare* and *H. bulbosum* are possible, most successful crossing schemes involving *H. bulbosum* include complex combinations of diploid, triploid, and tetraploid progeny (reviewed in Pickering and Johnston 2005, also see Johnston et al. 2009). The difficultly of hybridization between these species appears to be reflected in limited evidence of natural hybridization between *H. bulbosum* and *H. vulgare* ssp. *spontaneum*. In unpublished results, using a *H. bulbosum*-specific molecular marker, Johnston and Pickering found no evidence of introgression from *H. bulbosum* into *H. vulgare* ssp. *spontaneum* in almost 500 accessions from the Near East and Far East where the two species grow sympatrically (Pickering and Johnston 2005).

The largest impact of *H. bulbosum* on barley improvement occurred subsequent to the discovery that crosses of diploid *H. bulbosum* with *H. vulgare* could result in progeny containing a single haploid *H. vulgare* genome (Kasha and Kao 1970). This resulted in the widespread use of *H. bulbosum* for the production of doubled haploid barley lines. The "*Hordeum bulbosum* method" has been widely deployed to produce genetically identical progeny from barley crosses, which permits the testing of individual accessions across a large number of environments (Tinker et al. 1996). Highly replicated, doubled haploid barley populations have been used by many investigators to test new analytical methods for quantitative trait loci (QTL) mapping (e.g., Xu 2003; Zhang et al. 2005; Xu and Jia 2007).

Gametophytic self-incompatibility (GSI) occurs in diverse grass species, including *H. bulbosum* (Baumann et al. 2000). GSI in the grasses has been characterized as a two-locus system with two independent, multiallelic loci identified as the *S* and *Z* loci. The loci and genetic mechanism responsible for GSI in the grasses have yet to be isolated (Yang et al. 2008). Genetic mapping and gene expression-based approaches have been used to attempt to isolate causative loci in *H. bulbosum* (Gudu et al. 2002; Kakeda et al. 2008). Genetic maps and annotated expressed sequence tag (EST) sequence libraries for barley permit comparisons of genetic map positions for the *H. bulbosum S* locus to the barley genome, with higher marker density, physical map positions (Kunzel et al. 2000), and higher resolution genetic maps than would be available for many other self-incompatible grasses (Kakeda et al. 2008). However, the lack of a barley genome sequence means that actual gene-level identification of putatively causative loci must be based on comparison to the rice genome sequence (Kakeda et al. 2008), where gene order may not be conserved.

6.4 Genome Characteristics and Resources

The size of the barley genome has been estimated as 5.5 Gb, or roughly twice the size of the human genome. Despite the large genome size, the diploid

barley genome is more accessible and less complex than the genomes of most other cultivated species in the Triticeae tribe of the grass family, which also includes wheat and rye. Comparative sequence data from bacterial artificial chromosomes (BACs) from rice and barley suggest that colinear portions of the rice and barley genomes differ by large numbers of transposable elements (Dubcovsky et al. 2001). Much of the genome size of barley (and wild barley) can be attributed to the large number of transposable elements, including the copia-like retroelement BARE-1 (barley retroelement 1) (Manninen and Schulman 1993; Suoniemi et al. 1996), which is extremely abundant in the barley genome. Copy number of BARE-1 varies by >2-fold within *H. vulgare* ssp. *spontaneum* and ssp. *vulgare* and is a major contributor to variation in genome size (Vicient et al. 1999; Kalendar et al. 2000) with copy number being associated with eco-geographic variation within the habitat occupied by populations of *H. vulgare* ssp. *spontaneum* (Kalendar et al. 2000).

Currently, approximately ~24,000 EST sequences from *H. vulgare* ssp. *spontaneum* have been deposited in GenBank. An additional ~400,000 EST sequences are reported for *H. vulgare* ssp. *vulgare*. Much of the EST data for barley was retained as original sequence trace files, which were organized into database-driven software known as HarvEST that permits EST clustering *http://harvest.ucr.edu/*. Thus, barley EST data, including EST sequences from wild barley, can readily be browsed, visually inspected, and used for BLAST searches and single nucleotide polymorphism (SNP) identification.

Planning for barley genome sequencing is underway, but currently, neither a genome sequence nor a complete physical map of the genome is available (Schulte et al. 2009).

6.5 Conservation Initiatives

A number of germplasm repositories maintain collections of wild barley and related species for ex situ preservation. The *Gramene* website provides access to these centers at *http://www.gramene.org/species/hordeum/barley_germplasm.html*. Among the larger collections are those maintained at the United States Department of Agriculture (USDA) at Aberdeen, Washington, USA and by the International Center for Agricultural Research in the Dry Areas (ICARDA) in Aleppo, Syria. The *Hordeum vulgare* ssp. *spontaneum* collection at the USDA lists 1,152 accessions from 16 countries, and 192 *H. bulbosum* accession from 20 countries as currently available. The ICARDA collection currently reports 1,799 accessions of *Hordeum vulgare* ssp. *spontaneum* from 23 countries and 60 *H. bulbosum* accessions from 12 countries. Other large collections such as Plant Gene Resources of Canada include large number of accessions (e.g., 533 *Hordeum bulbosum* accessions); however, repositories often include duplicates of accessions in other repositories. In heavily collected regions, it is possible that multiple accessions derive from either the same or adjacent local plant populations.

All of the major repositories maintain online databases where "passport" data for each accession is available. This information includes geographic region of origin, generally including latitude and longitude and inferred elevation. This information is searchable online through the Germplasm Resources Information Network at *http://www.ars-grin.gov/npgs/index.html* and similar resources (see the *Gramene* link above for other repository websites). For many accessions, phenotypic traits, including disease and insect resistance or susceptibility, are recorded for individual accessions. For *Hordeum bulbosum*, the passport data includes ploidy, which is especially important for utilization of individual accessions of this species. However, as with other passport data, ploidy has not been assessed for all accessions. Many repositories have moved toward a system based on core collections (cf. Brown 1989), which have relatively complete phenotypic and collection data.

6.6 Role in Classical and Molecular Genetic Studies

Because wild barley is predominantly inbreeding, individual plants grown from seed provided by genetic stock centers have very low heterozygosity and thus constitute sets of naturally, partially inbred lines. This accelerates the creation of inbred lines that can be genotyped once and then used repeatedly for phenotypic analysis of traits of interest. Large projects such

as the *Arabidopsis thaliana* 1,001 genomes resequencing effort are applying this approach on a massive scale; resequencing whole genomes from many lines so that the resequencing data can then be used in a variety of studies designed to more directly link genotypes and phenotypes (Weigel and Mott 2009).

Similar efforts have been initiated in wild barley, though initially on a smaller scale. One example is the Wild Barley Diversity Collection (WBDC), developed at the University of Minnesota (Steffenson et al. 2007). The WBDC includes 318 samples representing the entire geographic range of wild barley. The WBDC has been genotyped using 558 DaRT markers and with >3,000 SNPs on the Illumina Golden Gate Genotyping platform (Close et al. 2009).

Another research group has reported establishment of a collection of wild barley lines from Israel. The Barley1K collection, as described by Hübner et al. (2009) is a hierarchical sample of 1,020 wild barley accessions from 51 populations from diverse ecogeographic locations within Israel. For each collection locality, climatic data, soil conditions, etc. are reported. Accessions were also genotyped at 42 microsatellite loci. The high-density of sampling in the Barley1K collection will be used to examine allelic variation relevant to the traits important to agriculture.

Wild barley introgression lines, which contain small portions of the wild barley genome in the genetic background of a cultivated line, provide an opportunity to introgress favorable alleles from wild barley lines into breeding populations. A number of papers have reported successful production of wild barley introgression lines. The design of these populations varies, but all have the goal of introducing small chromosomal segments from the wild barley genome in an effort to introgress favorable alleles into domesticated barley that may improve biotic or abiotic stress response, yield, or contribute to other favorable agronomic traits.

Many of the studies take the form of advance backcross QTL (AB-QTL) experiments (Tanksley and Nelson 1996) in early generations (e.g., Pillen et al. 2003; von Korff et al. 2008) and transition toward the creation of recombinant inbred lines (RIL) or similar populations in later generations (e.g., Baum et al. 2003; Grando et al. 2005; Yun et al. 2005, 2006; Inostroza et al. 2007).

Recently, a set of 110 diploid *Hordeum bulbosum* introgression lines has also been reported (Johnston et al. 2009). The introgression lines cover all but one chromosome arm in the *Hordeum* genome. A set of 46 sequence-based markers were used to identify introgression. The lines were derived from three different barley cultivars (Emir, Golden Promise, and Morex) and four *H. bulbosum* accessions. The use of multiple *H. bulbosum* lines is interesting; perhaps because of the effort involved in creating AB-QTLs and RIL populations, most wild barley introgression projects have sampled a single wild barley accession. Much additional diversity within the *H. vulgare* ssp. *spontaneum* and *H. bulbosum* remains unexplored.

While wild barley has frequently been used as a parent in QTL mapping experiments, intercompatibility with domesticated barley obviates the need for independent genetic maps specific to wild barley. A restriction fragment length polymorphism (RFLP)-based genetic map for *H. bulbosum* has been reported (Jaffe et al. 2000; Salvo-Garrido et al. 2001). Comparison with genetic maps from *H. vulgare* suggests that the genomes of the two species are largely colinear. The recombination rate within barley centromeric regions is higher than in *H. bulbosum* while the opposite is true for telomeric regions. The creation of a *H. bulbosum* genetic map should facilitate the isolation of loci of interest in *H. bulbosum*, including loci that contribute to self-incompatibility and genes that regulate haploid formation and intercompatibility in crosses with barley (Salvo-Garrido et al. 2001).

6.7 Role in Crop Improvement Through Traditional and Advanced Tools

Efforts to introgress favorable alleles from wild barley into barley breeding populations have been reported by many different research groups (Baum et al. 2003; Matus et al. 2003; Pillen et al. 2003, 2004; Grando et al. 2005; von Korff et al. 2005, 2006; Gyenis et al. 2007; Inostroza et al. 2007; Schmalenbach et al. 2008, 2009; von Korff et al. 2008). The majority of these studies report AB-QTL experiments, where wild barley is used as a donor parent. A minimum of two to three generations of backcrossing to the cultivated barley recurrent parent is generally required to recover agronomic phenotypes (cf. Pillen et al. 2004). Barley variety development typically involves crossing of

parents with favorable alleles, followed by multiple generations of selection for favorable agronomic traits prior to the development of a new inbred line. A number of independent studies have reported that wild barley lines were found to contribute favorable alleles in QTL mapping experiments, with the proportion of favorable alleles contributed by wild barley ranging from ~25 to 40.9% (von Korff et al. 2006; Inostroza et al. 2009; Schmalenbach et al. 2009). In a number of cases, the favorable alleles contributed by wild barley are for yield and yield component traits, such as grain number per spike and ear length (Schmalenbach et al. 2009). Another study reported that 37.5% of favorable alleles for the very complex phenotype, malting quality, were contributed by the wild barley parent in an AB-QTL population (von Korff et al. 2008). These studies suggest that while wild barley may be particularly valuable as a source of disease resistance alleles that can be readily introgressed into barley breeding programs (cf. Yun et al. 2006), there are a number of other phenotypes where wild barley can contribute favorable alleles. However, in general, it has been difficult to develop lines resulting from introgression that exceed the productive capability of existing cultivars (Inostroza et al. 2009).

QTL mapping based on large numbers of RILs derived from crosses with wild barley also permits the identification of genomic regions where alleles contributed by the cultivated parent are required in order to recover favorable agronomic phenotypes, an approach Inostroza et al. (2009) describe as a form of "genetic knockdown" experiment.

6.8 Distribution of Genetic Diversity in Wild Barley

A number of studies have examined the nature of adaptations that contribute to ecotypic variation in wild barley (e.g., Volis et al. 2002, 2004; Verhoeven et al. 2004, 2008). These studies have included comparison of potentially adaptive traits in wild barley populations from the extremes of the range, including populations from Israel and Turkmenistan (Volis et al. 2000) and comparison of ecotypic variation within single regions (Volis et al. 2002). Wild barley ecotypes from desert, semi-steppe batha (shrubland-like plant community), Mediterranean grasslands, and mountains have been compared under common garden conditions to identify the extent of differential response to water stress and other environmental factors (Volis et al. 2000, 2002). A density experiment using the various barley ecotypes found that plants from the Mediterranean grassland ecotype was the superior competitor, with the greatest yield and reproductive output (Volis et al. 2004). There is a complex set of traits associated with local adaptation in wild barley, involving subtle differences in factors such as drought and frost tolerance, plant resource allocation, and timing of reproduction (Volis et al. 2004). Examination of the genetic basis of differentiation between wild barley populations in QTL populations (Verhoeven et al. 2004, 2008; Poorter et al. 2005) showed that QTLs putatively associated with adaptation, and increased fitness in one ecological environment does not necessarily diminish fitness in another. This suggests a complex basis for local adaptation dependent on variation at a large number of underlying traits (Verhoeven et al. 2004, 2008; Poorter et al. 2005).

There is a long history of work on biochemical and molecular diversity in barley that is well reviewed in Bothmer et al. (2003). A particularly comprehensive early study examined more than 1,500 accessions of wild and cultivated barley on a worldwide scale, but based on just four esterase encoding loci. This study found substantial levels of genetic diversity and also that most wild alleles had been incorporated into cultivated barley (Kahler and Allard 1981), indicating little loss of genetic diversity following domestication. Interestingly, Kahler and Allard (1981) noted some genetic differentiation among collections between East Asian and Middle South Asian accessions and European and North American collections. Later, isozyme surveys tended to assay many more loci, but often with a more restricted geographic focus. An exception can be found in Nevo et al. (1986) where 27 loci were analyzed in 2,125 individuals from portions of Israel, Turkey, and Iran. While this work uncovered hints of geographic differentiation, the main focus was on environmental genetic correlations. Clear evidence for geographic patterns of differentiation did not arise until the beginning of DNA-based resequencing studies. Resequencing studies are especially informative because they reveal all mutations in a gene sample and the mutations can easily be classified into amino acid substitutions and

synonymous changes. In addition, indel mutations are easily resolved. Finally, complete haplotype data can be resolved, thereby yielding the linkage phase of mutations within a gene. Haplotype data contain valuable information on linkage disequilibrium (Watterson 1975) and they allow the investigator to reconstruct the temporal sequence of mutational differences between haplotypes.

Various statistics can be calculated from these data that relate to effective population size and that can potentially be used to test for selection at a locus. One useful statistic is θ, a function of the number of polymorphic sites in a gene; for neutral genes, θ is approximately equal to $4N_e\mu$, where N_e is the effective population size and μ is the mutation rate per nucleotide site (Watterson 1975). Another common statistic is π, a function of the pairwise frequency of site variants (Tajima 1983). A statistic known as Tajima's D is a function of the difference between θ and π. In populations at equilibrium between genetic drift and mutation, the difference is expected to be zero for neutral loci (Tajima 1989). Nonzero values of Tajima's D may arise owing to demographic changes or owing to natural selection. Some of these statistics will be referred to below.

Wild barley, along with *Arabidopsis thaliana*, was among the first plant species in which extensive resequencing data were collected (Wright and Gaut 2005). The most comprehensive survey to date was based on a sample of 45 accessions carefully selected to span the entire geographic range of wild barley. The first locus examined in wild barley, alcohol dehydrogenase 1 (*Adh1*), had relatively modest levels of diversity with θ for synonymous sites of 0.005. An excess of low frequency nonsynonymous substitutions was observed in the sample, but there was no evidence of recombination among sampled haplotypes, and no evidence of population structure (Cummings and Clegg 1998). These results were largely consistent with expectations for a self-fertilizing species. Resequencing data based on a subset of 25 accessions from two additional *Adh* loci identified higher levels of diversity and evidence of recombination at *Adh2* (Lin et al. 2002). In sharp contrast, a resequencing study of *Adh3* (Lin et al. 2001) revealed extensive divergence between haplotypes, marked by population structure between accessions collected east and west of the Zagros Mountains (Lin et al. 2001) and evidence of recombination between haplotypes (Lin et al. 2002). Estimates of diversity for synonymous sites for *Adh2* and *Adh3* were 0.0159 and 0.0325, respectively; dramatically higher than the level of diversity observed at *Adh1*. Thus, both the geographic distribution of haplotypes and level of diversity for the other two independently segregating *Adh* loci bore little resemblance to that identified at *Adh3*, even though the same 25 accessions comprised the sample. Indeed, sampled haplotypes at *Adh1* and *Adh2* seemed to be distributed almost at random across the range of the species, suggesting that the rate of migration was at least of the order of the temporal history of the coalescent process. The distribution of haplotype diversity at *Adh3* implied a barrier to gene flow across the range of wild barley, while the broad geographic distribution of some haplotypes at *Adh1* and *Adh2* implies migration sufficient to distribute all haplotypes across the range. Because all genes in a genome move together as diploid individuals or, in the case of pollen flow, as haploid gametes, the results for the *Adh* loci presented a paradox. To ask whether the unusual pattern was unique to *Adh*3, an expanded survey of additional loci was undertaken.

The subsequent study was designed to investigate the strength of migration across the geographic range and to characterize heterogeneities in spatial patterns of genetic diversity. The number of loci resequenced was expanded by six additional loci based on the same sample of 25 accessions used for the *Adh* genes (Morrell et al. 2003). Owing to small sample size, the study focused on broad-scale geographic patterns and examined levels of diversity and rate of migration between the eastern, western, and intermediate "Zagros" portion of the range of wild barley. Using a maximum-likelihood, coalescence-based approach, average migration rate per locus among regions was estimated as slightly greater than one migrant per generation. This rate of migration should be sufficient to result in a homogeneous geographic distribution of haplotype variation. However, geographic structure was evident at roughly half of the sampled loci, with the *G3pdh* locus showing evidence of extensive haplotype divergence, similar to that reported for *Adh3*. As noted above, wild barley has long, barbed awns that promote dispersal of disarticulated spikelets and thus is well adapted for dispersal by both small and large mammals. Thus, a relatively high rate of migration among wild barley populations is plausible for the species. Why do haplotypes at some loci in wild barley

disperse across the entire range on a time scale of roughly N_e while other loci exhibit strong patterns of genetic differentiation? The answer to this question is still unresolved, but it seems likely that some kind of geographic selection is affecting patterns of diversity at a substantial subset of loci.

At first sight, it appears surprising that large isozyme surveys, such as those cited above, failed to detect clear genetic differentiation east and west of the Zagros Mountains. Put differently, why should resequencing, based on a relatively small sample of accessions, reveal sharp patterns of geographic differentiation not apparent at the isozyme level? The answer can probably be found in the power of haplotype data. The mutational distance among haplotypes is readily apparent in haplotype data and this allowed the detection of divergent haplotype classes by geographic region. There is no comparable measure of mutational distance between isozyme alleles; so haplotype data are much more powerful for the detection of geographic structure.

6.9 Linkage Disequilibrium

Linkage disequilibrium (LD) in wild barley was estimated from the same 25 accessions discussed above based on resequencing data from 18 loci (Morrell et al. 2005). Three loci with very marked population structure contributed to strong admixture LD, but when the preponderance of the data is considered without these loci, the decay of LD in wild barley is relatively rapid, falling to half the initial value in roughly the first 1,000 bp. Relatively rapid decay of LD was also observed in an independent sample of four loci in 34 accessions from across the range of wild barley (Caldwell et al. 2006). Resequencing of individual loci, such as the *Ppd-H1* related to flowering time adaptation, has identified large numbers of recombination events, again suggestive of rapid decay of LD (Jones et al. 2008). Surprisingly, the estimated decay of LD is comparable to that found in the outcrossing crop species maize (Morrell et al. 2005). This result presents a paradox because predominantly self-fertilizing species are expected to be characterized by extensive LD. The reason for this expectation is that inbreeding leads to high levels of homozygosity, thus suppressing effective recombination (Nordborg 2000). It is not clear why LD is so limited in wild barley, although one plausible hypothesis is that wild barley evolved self-fertilization relatively recently in its evolutionary history (perhaps within the past 20,000–40,000 years). Some support for this hypothesis can be found in the fact that *H. bulbosum*, the sister species of wild barley, has a self-incompatible breeding system, and molecular clock calculations indicate that the two species diverged from a common ancestor roughly seven million years ago (Blattner 2004).

Admixture LD presents both a problem and an opportunity for those who wish to use association mapping to locate major genes of interest. The problem arises because the pooling of haplotypes across heterogeneous subpopulations induces LD in the pooled sample independent of map distance. Thus, LD is not a reliable indicator of map distance, and one must first determine the extent of population substructure before making inferences based on LD. The opportunity arises because when properly characterized, admixture LD can be employed to look for gene-trait associations.

6.10 Recombination

Recombination is composed of two processes: meiotic crossing-over involving a symmetrical exchange between sister chromatids and gene conversion where short tracks of a few hundred base pairs are asymmetrically exchanged between synapsing chromatids. Resequencing data permit detailed investigation of recombinational processes because the coalescent history represented in a sample includes many thousands of meioses and therefore provides much greater statistical power than even large genetic crosses. This feature of resequencing data was exploited by Morrell et al. (2006) to estimate the frequency of gene conversion events relative to crossover events in wild barley and several other species. The result showed that gene conversion events are at least as frequent as crossover events in barley recombination. This finding adds an additional complication for association mapping, because the effect of gene conversion is to recombine adjacent sequences while not affecting more distantly linked genes. LD may not be a simple linear function of map distance because of this effect, so LD is likely

to be a biased predictor of gene-trait distances because LD may be lower at short distances than at intermediate distances. This will clearly confound association mapping and lead to some errors.

6.11 Genetic Evidence and the Origins of Domesticated Barley

The question of how and why Neolithic humans began to domesticate plants and animals is certainly one of the great historical mysteries of all time. There is clear archeological evidence of domestication of barley by 10,500 BP in the Fertile Crescent region of the Middle East (Bothmer et al. 1995; Willcox 2005). Evidence of human utilization of barley may extend to 19,000 years BP (Zohary and Hopf 2000). As one of the first plant species to be domesticated in the Fertile Crescent region, barley sits at the nexus of this historical conundrum. There has been a long debate about whether barley was domesticated once in the Fertile Crescent and spread east into Asia and north into Europe or whether barley might have been domesticated multiple times (Zohary 1999; Zohary and Hopf 2000). For many years, the idea of a unique domestication was appealing because the transition to agriculture must have been associated with many changes in human culture that, among other things, required abandoning a nomadic existence for an existence tied to particular plots of land. Moreover, it seems parsimonious to argue that the genetic changes induced by such a transition occurred just once. Archeological evidence clearly supports an origin of barley domestication in the Fertile Crescent (Willcox 2005), but there have been intriguing hints of a second origin largely based on genetic data.

One of the strongest and most intriguing pieces of evidence for multiple origin of a crop comes from "domestication traits" (Sang 2009; Zohary 1999). Wild barley is characterized by a brittle rachis that disarticulates at seed maturity. This facilitates the dispersion of barley seed and is therefore thought to be adaptive for the wild plant. Domesticated barley has a mutation that produces a nonbrittle rachis, so that the ears of grain are retained on the stalk at maturity, thus facilitating seed collection by humans (hence the term "domestication trait"). It is intriguing that the nonbrittle phenotype is actually controlled by two distinct genetic loci, either of which can produce the nonbrittle phenotype (Takahashi and Yamamoto 1949; Takahashi 1964). Why would human domesticators have selected for a second mutation when one serves the purpose? The best explanation for this puzzle is that the two mutations were selected independently in two geographically separate populations, thus implying a second independent domestication. There is strong evidence for multiple origins of another key domestication trait, the two-rowed versus six-rowed phenotype (Komatsuda et al. 2007). However, another domestication trait, "naked" or hulless barley grains, appears to have a single origin (Taketa et al. 2008).

The extensive genetic differentiation east and west of the Zagros Mountains afforded an opportunity to apply a different test for multiple domestications of barley. The reasoning is that if landraces of barley were domesticated more than once, they should be genetically similar to their geographic area of domestication and subsequent cultivation. Morrell and Clegg (2007) applied this reasoning by resequencing seven loci that exhibited geographic differentiation in wild barley in a set of 32 cultivated barleys. A genetic assignment approach, implemented in the program Structure (Pritchard et al. 2000; Falush et al. 2003), was applied to the data to ask whether there was concordance between wild and landrace barley from particular geographic regions. The data clearly showed strong concordance, supporting a multiple domestication hypothesis.

The cultivated sample also included a few modern cultivars of barley. Cultivars of European or North American origin tended to show a Fertile Crescent origin while cultivars from Asia indicated an origin east of the Zagros Mountains. The sample mesh was insufficient to ask whether population substructure exists on a finer scale. It also did not pinpoint the geographic location(s) of domestication events, although the data hint that the origin of eastern landraces was in the western foothills of the Zagros or points farther east. Locations of early Neolithic agropastoral settlements suggest three general regions in which the secondary domestication could have taken place; in the foothills of the Zagros, at such sites as Ali Kosh and Jarmo at Mehrgarh (in present day Pakistan) or in the piedmont zone between the Kopet Dag mountain range and Kara Kum Desert (east of the Caspian Sea, in present day Turkmenistan).

The discovery of multiple domestications is of more than academic interest. Multiple origins imply that unique genes of agricultural value may be uncovered in different ecogeographical regions. This places a premium on maintaining genetic resource collections from different ecogeographical regions and on conducting more detailed surveys with a much finer sample mesh to uncover the finer details of both genetic substructure and domestication history. Ultimately, it will be possible to trace the geographic origins of many major adaptive haplotypes and to ask how their introduction into otherwise adapted cultivars may improve barley quality and yield.

References

Baum M, Grando S, Backes G, Jahoor A, Sabbagh A, Ceccarelli S (2003) QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' x *H. spontaneum* 41-1. Theor Appl Genet 107: 1215–1225

Baumann U, Juttner J, Bian XY, Langridge P (2000) Self-incompatibility in the grasses. Ann Bot 85:203–209

Blattner FR (2004) Phylogenetic analysis of *Hordeum* (Poaceae) as inferred by nuclear rDNA ITS sequences. Mol Phylogenet Evol 33:289–299

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Caldwell KS, Russell J, Langridge P, Powell W (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. Genetics 172:557–567

Close TJ, Bhat PR, Lonardi S, Wu YF, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S, Bozdag S, Roose MA, Moscou MJ, Chao S, Varshney R, Szucs P, Sato K, Hayes PM, Matthews DE, Kleinhofs A, Muehlbauer GJ, DeYoung J, Marshall DF, Madishetty K, Fenton RD, Condamine P, Graner A, Waugh R (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10:582

Cummings MP, Clegg MT (1998) Nucleotide sequence diversity at the alcohol dehydrogenase 1 locus in wild barley (*Hordeum vulgare* ssp. *spontaneum*): an evaluation of the background selection hypothesis. Proc Natl Acad Sci USA 95:5637–5642

Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiol 125:1342–1353

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587

Grando S, Baum M, Ceccarelli S, Goodchild A, El-Haramein FJ, Jahoor A, Backes G (2005) QTLs for straw quality characteristics identified in recombinant inbred lines of a *Hordeum vulgare* x *H. spontaneum* cross in a Mediterranean environment. Theor Appl Genet 110:688–695

Gudu S, Laurie A, Kasha J, Xia J, Snape W (2002) RFLP mapping of a *Hordeum bulbosum* gene highly expressed in pistils and its relationship to homoeologous loci in other Gramineae species. Theor Appl Genet 105:271–276

Gyenis L, Yun SJ, Smith KP, Steffenson BJ, Bossolini E, Sanguineti MC, Muehlbauer GJ (2007) Genetic architecture of quantitative trait loci associated with morphological and agronomic trait differences in a wild by cultivated barley cross. Genome 50:714–723

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

Inostroza L, del Pozo A, Matus I, Hayes P (2007) Drought tolerance in recombinant chromosome substitution lines (RCSLs) derived from the cross *Hordeum vulgare* subsp *spontaneum* (Caesarea 26-24) x *Hordeum vulgare* subsp *vulgare* cv Harrington. Agri Técnica 67:253–261

Inostroza L, Pozo A, Matus I, Castillo D, Hayes P, Machado S, Corey A (2009) Association mapping of plant height, yield, and yield stability in recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp *vulgare* background. Mol Breed 23:365–376

Jaffe B, Caligari PDS, Snape JW (2000) A skeletal linkage map of *Hordeum bulbosum* L. and comparative mapping with barley (*H. vulgare* L.). Euphytica 115:115–120

Johnston PA, Timmerman-Vaughan GM, Farnden KJ, Pickering R (2009) Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. Theor Appl Genet 118:1429–1437

Jones H, Leigh FJ, Mackay I, Bower MA, Smith LMJ, Charles MP, Jones G, Jones MK, Brown TA, Powell W (2008) Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. Mol Biol Evol 25:2211–2219

Kahler AL, Allard RW (1981) Worldwide patterns of genetic variation among 4 esterase loci in barley (*Hordeum vulgare* L). Theor Appl Genet 59:101–111

Kakeda K, Ibuki T, Suzuki J, Tadano H, Kurita Y, Hanai Y, Kowyama Y (2008) Molecular and genetic characterization of the S locus in *Hordeum bulbosum* L., a wild self-incompatible species related to cultivated barley. Mol Genet Genom 280:509–519

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci USA 97:6603–6607

Kasha KJ, Kao KN (1970) High frequency haploid production in barley (*Hordeum vulgare* L.). Nature 225:874–876

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M (2007) Six-rowed barley originated from a mutation

in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci USA 104:1424–1429

Kunzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412

Lin JZ, Brown AH, Clegg MT (2001) Heterogeneous geographic patterns of nucleotide sequence diversity between two alcohol dehydrogenase genes in wild barley (*Hordeum vulgare* subspecies *spontaneum*). Proc Natl Acad Sci USA 98:531–536

Lin JZ, Morrell PL, Clegg MT (2002) The influence of linkage and inbreeding on patterns of nucleotide sequence diversity at duplicate alcohol dehydrogenase loci in wild barley (*Hordeum vulgare* ssp. *spontaneum*). Genetics 162:2007–2015

Lundqvist A (1962) Self-incompatibility in diploid *Hordeum bulbosum* L. Hereditas 48:138–152

Manninen I, Schulman AH (1993) BARE-1, a copia-like retroelement in barley (*Hordeum vulgare* L.). Plant Mol Biol 22:829–846

Matus I, Corey A, Filichkin T, Hayes PM, Vales MI, Kling J, Riera-Lizarazu O, Sato K, Powell W, Waugh R (2003) Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp *vulgare* background. Genome 46: 1010–1023

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

Morrell PL, Lundy KE, Clegg MT (2003) Distinct geographic patterns of genetic diversity are maintained in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite migration. Proc Natl Acad Sci USA 100:10812–10817

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

Morrell PL, Toleno DM, Lundy KE, Clegg MT (2006) Estimating the contribution of mutation, recombination and gene conversion in the generation of haplotypic diversity. Genetics 173:1705–1723

Nevo E, Beiles A, Zohary D (1986) Genetic resources of wild barley in the Near East: structure, evolution and application in breeding. Biol J Linn Soc 27:355–380

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154:923–929

Pickering R, Johnston PA (2005) Recent progress in barley improvement using wild species of *Hordeum*. Cytogenet Genome Res 109:344–349

Pillen K, Zacharias A, Leon J (2003) Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). Theor Appl Genet 107:340–352

Pillen K, Zacharias A, Leon J (2004) Comparative AB-QTL analysis in barley using a single exotic donor of *Hordeum vulgare* ssp *spontaneum*. Theor Appl Genet 108:1591–1601

Poorter H, van Rijn CP, Vanhala TK, Verhoeven KJ, de Jong YE, Stam P, Lambers H (2005) A genetic analysis of relative growth rate and underlying components in *Hordeum spontaneum*. Oecologia 142:360–377

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Salvo-Garrido H, Laurie DA, Jaffe B, Snape JW (2001) An RFLP map of diploid *Hordeum bulbosum* L. and comparison with maps of barley (*H. vulgare* L.) and wheat (*Triticum aestivum* L.). Theor Appl Genet 103:869–880

Sang T (2009) Genes and mutations underlying domestication transitions in grasses. Plant Physiol 149:63–70

Schmalenbach I, Korber N, Pillen K (2008) Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. Theor Appl Genet 117:1093–1106

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

Schulte D, Close TJ, Graner A, Langridge P, Matsumoto T, Muehlbauer G, Sato K, Schulman AH, Waugh R, Wise RP, Stein N (2009) The International Barley Sequencing Consortium-At the threshold of efficient access to the barley genome. Plant Physiol 149:142–147

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

Suoniemi A, Anamthawat-Jonsson K, Arna T, Schulman AH (1996) Retrotransposon BARE-1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. Plant Mol Biol 30:1321–1329

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 3:585–595

Takahashi R (1964) Linkage study of two complementary genes for brittle rachis in barley. Ber Ohara Inst landw Biol Okayama Univ 38:81–90

Takahashi R, Yamamoto J (1949) Studies on the classification and the geographic distribution of barley varieties, VIII. Nogaku Kenkyi 38:81–90

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

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Tinker NA, Mather DE, Rossnagel BG, Kasha KJ, Kleinhofs A, Hayes PM, Falk DE, Ferguson T, Shugar LP, Legge WG, Irvine RB, Choo TM, Briggs KG, Ullrich SE, Franckowiak JD, Blake TK, Graf RJ, Dofing SM, Saghai-Maroof MA, Scoles GJ, Hoffman D, Dahleen LS, Kilian A, Chen F, Biyashev RM, Kudrna DA, Steffenson BJ (1996) Regions of the genome that affect agronomic performance in two-row barley. Crop Sci 36:1053–1062

Verhoeven KJ, Vanhala TK, Biere A, Nevo E, van Damme JM (2004) The genetic basis of adaptive population differentiation: a quantitative trait locus analysis of fitness traits in two

wild barley populations from contrasting habitats. Evolution 58:270–283

Verhoeven KJ, Poorter H, Nevo E, Biere A (2008) Habitat-specific natural selection at a flowering-time QTL is a main driver of local adaptation in two wild barley populations. Mol Ecol 17:3416–3424

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. Plant Cell 11:1769–1784

Volis S, Mendlinger S, Orlovsky N (2000) Variability in phenotypic traits in core and peripheral populations of wild barley *Hordeum spontaneum* Koch. Hereditas 133:235–247

Volis S, Mendlinger S, Ward D (2002) Differentiation in populations of *Hordeum spontaneum* Koch along a gradient of environmental productivity and predictability: plasticity in response to water and nutrient stress. Biol J Linn Soc 75:301–312

Volis S, Verhoeven KJ, Mendlinger S, Ward D (2004) Phenotypic selection and regulation of reproduction in different environments in wild barley. J Evol Biol 17:1121–1131

von Bothmer R, Flink J, Jacobsen N, Kotimaki M, Landstrom T (1983) Interspecific hybridization with cultivated barley (*Hordeum vulgare* L.). Hereditas 99:219–244

von Bothmer, R, Jacobsen N, Baden C, Jørgensen RB, Linde-Laursen I (1995) An ecogeographical study of the genus *Hordeum*. In: Systematic and ecogeographic studies on crop genepools, vol 7. IBPGR, Rome, Italy, 127 p

von Bothmer R, Sato K, Komatsuda T, Yasuda S, Fischbeck G (2003) The domestication of cultivated barley. In: von Bothmer R, van Hintum T, Knüpffer H, Sato K (eds) Diversity in barley (*Hordeum vulgare*). Elsevier, San Diego, CA, USA, pp 9–27

von Korff M, Wang H, Leon J, Pillen K (2005) AB-QTL analysis in spring barley. I. Detection of resistance genes against powdery mildew, leaf rust and scald introgressed from wild barley. Theor Appl Genet 111:583–590

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

von Korff M, Wang HJ, Leon J, Pillen K (2008) AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp *spontaneum*). Mol Breed 21:81–93

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276

Weigel D, Mott R (2009) The 1001 Genomes roject for *Arabidopsis thaliana*. Genome Biol 10**:**ARTN 107

Willcox G (2005) The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres. Veg Hist Archaeobot 14:534–541

Wright SI, Gaut BS (2005) Molecular population genetics and the search for adaptive evolution in plants. Mol Biol Evol 22:506–519

Xu S (2003) Estimating polygenic effects using markers of the entire genome. Genetics 163:789–801

Xu S, Jia Z (2007) Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics 175:1955–1963

Yang B, Thorogood D, Armstead I, Barth S (2008) How far are we from unravelling self-incompatibility in grasses? New Phytol 178:740–753

Yun SJ, Gyenis L, Hayes PM, Matus I, Smith KP, Steffenson BJ, Muehlbauer GJ (2005) Quantitative trait loci for multiple disease resistance in wild barley. Crop Sci 45:2563–2572

Yun SJ, Gyenis L, Bossolini E, Hayes PM, Matus I, Smith KP, Steffenson BJ, Tuberosa R, Muehlbauer GJ (2006) Validation of quantitative trait loci for multiple disease resistance in barley using advanced backcross lines developed with a wild barley. Crop Sci 46:1179–1186

Zhang M, Montooth KL, Wells MT, Clark AG, Zhang D (2005) Mapping multiple quantitative trait loci by Bayesian classification. Genetics 169:2305–2318

Zohary D (1999) Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the Near East. Genet Resour Crop Evol 46:133–142

Zohary D, Hopf M (2000) Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley. Oxford University Press, Oxford, UK, 316 p