

Mapping and exploiting the barley genome: techniques for mapping genes and relating them to desirable traits

Hélène Pidon, Nils Stein

▶ To cite this version:

Hélène Pidon, Nils Stein. Mapping and exploiting the barley genome: techniques for mapping genes and relating them to desirable traits. Achieving sustainable cultivation of barley, Burleigh Dodds Science Publishing, pp.123-156, 2020, Burleigh Dodds Series in Agricultural Science, 10.19103/AS.2019.0060.05. hal-04081555

HAL Id: hal-04081555 https://hal.inrae.fr/hal-04081555

Submitted on 25 Apr 2023 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



CH05 - Mapping and exploiting the barley genome

Publisher:	Burleigh Dodds Science Publishing			
Manuscript ID	BDSP-BK-2018-0237.R1			
Manuscript Type:	Book			
Date Submitted by the Author:	12-Feb-2019			
Complete List of Authors:	Stein, Nils; Leibniz-Institut fur Pflanzengenetik und Kulturpflanzenforschung Gatersleben Pidon, Hélène; Leibniz-Institut fur Pflanzengenetik und Kulturpflanzenforschung Gatersleben,			
Keywords:	Gene mapping, markers, sequencing			



1 Mapping and exploiting the barley genome – techniques for mapping genes

2 and relating them to desirable traits

- 3 Hélène Pidon¹ and Nils Stein^{1,2}
- ⁴ ¹Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany
- ⁵ ²Georg-August-University, Center of integrated Breeding Research (CiBreed), Göttingen, Germany
- 6 7 stein@ipk-gatersleben.de
- , 8

stellie pr Saters

- 9 Abstract:
- 10 Barley is the fourth most cultivated cereal in the world and is thus an important crop in breeding. A
- 11 large variety of traits are needed for resistance to various stresses as well as to meet the
- 12 requirements of its several uses. To breed efficiently for those traits, identification of the causal
- 13 genes and availability of genetic markers for selection are important targets for research. Gene
- 14 mapping approaches have undergone radical changes in the past decades. In this chapter, we give an
- 15 overview of these changes emphasizing the importance of the availability of barley reference
- 16 genomes and sequencing technologies. An overview of the methods and resources for mapping
- 17 barley genes is given, covering developments over time, and concluding with future trends.
- 18 Key words:
- 19 Gene mapping, markers, sequencing, genome
- 20 Content:
- 21 1 Introduction
- 22 2 What are your markers? New possibilities in the genomics era
- 23 3 Classical mapping strategies still efficient but improving in the genomics era
- 24 4 The association mapping boom
- 25 5 Multiparental populations: The perfect balance?
- 26 6 From an interval to the causal gene: from high-resolution mapping to gene cloning
- 27 7 Emerging mapping strategies: fast NGS-enabled technologies
- 28 8 Conservation of barley germplasm
- 29 9 Genetic and genomic resources of barley
- 30 10 Case Study: From *rym4* to *rym11*, illustration of paradigm shift in disease resistance mapping and
- 31 cloning
- 32 11 Conclusion and future trends
- 33 12 Where to look for further information
- 34 13 References
- 35

36 1 Introduction

- 37 Barley has been proposed as a model species for Triticeae (Linde-Laursen et al., 1997) given its high
- economic importance, being the fourth most cultivated cereal (FAOSTAT, 2018) and at the same time
- having a less complex genome than the tetra- and hexaploid wheats. It is diploid and highly
- 40 inbreeding, and a large number of barley genetic stocks are available. Barley can adapt to various
- 41 biotic and abiotic stresses and has various uses, from malting to animal feeding, thus breeders target

42 a large number of traits. The molecular mapping of genes and traits started more than three decades ago and, since then, genetic and genomic resources for barley were systematically developed and 43 44 improved. Marker-assisted selection was implemented to accelerate breeding, however, low marker density and only loose linkage of markers to functional genes were serious limitations in early 45 46 attempts to isolate functional genes underlying traits of interest. The complexity of the 5.1 Gbp 47 barley genome, packed full with repetitive elements (International Barley Genome Sequencing 48 Consortium, 2012; Mascher et al., 2017) rendered this step slower than in smaller genome plant 49 species like Arabidopsis or rice. The first gene in barley isolated by positional cloning was Mlo, a gene 50 conferring resistance to powdery mildew (Büschges et al., 1997). As of today at least 65 cloned barley 51 genes have been documented, including 28 for architecture traits alone (Hansson et al., 2018). Forty-52 two of them were cloned after 2010, demonstrating the importance of improved genomic resources 53 in barley gene isolation. The introduction of next-generation sequencing (NGS) technologies 54 constituted a paradigm shift, bringing barley genetics into the genomic era where marker discovery 55 and genotyping is easy and fast. Recently, an annotated reference sequence was published (Mascher 56 et al., 2017), enabling a range of new methods to rapidly and precisely map genes at moderate cost, 57 and to provide breeders with tools for genomics-assisted breeding. Species diversity and germplasm 58 collections are key for finding new diversity for traits desirable in breeding but, despite the careful 59 survey of the species that started centuries ago, genebanks' descriptions remain limited. NGS also 60 opened the doors to their complete genetic characterization, providing better-informed access and 61 use of genetic resources in breeding programs.

62 In this chapter, we review the methods currently available for mapping genes and quantitative trait

63 loci (QTLs) in barley, from high-throughput genotyping to gene identification, with a focus on how

64 the application of NGS technologies in barley studies can accelerate gene mapping. We conclude by

65 considering prospects for the future.

66 2 What are your markers? New possibilities in the genomics era

When an interesting phenotype has been identified, the first steps are to determine whether it is mono- or polygenic, and its dominance. This is achieved by studying the segregation of the trait by phenotyping a segregating population, typically derived from crossing homozygous parental genotypes. Segregating progeny are then genotyped to map the gene. Genetic markers used in gene mapping have undergone constant progress for almost fifty years. From the use of metabolites, allozymes and isozymes, DNA markers emerged almost 40 years ago, allowing higher marker density and flexibility.

- 74 The first DNA markers were restriction fragment length polymorphisms (RFLPs) based on fragment
- 75 length differentiation of DNA digested by restriction endonuclease (Botstein et al., 1980; Graner et
- 76 al., 1991; Tanksley et al., 1989). For visualization, a Southern blot of the digested fragments was
- 77 hybridized to a radioactively labeled DNA probe and then exposed to x-ray films, resulting in
- 78 differential banding profiles of the autoradiogram. Widely used in the 1980s, hybridization-based
- 79 markers entailed laborious procedures and labeling of probes with radioisotopes. The development
- of PCR technology in 1985 (Saiki et al., 1985) prompted the replacement of those markers by PCR-
- 81 based markers (Figure 1). Initially, RFLP DNA fragments were sequenced (Michalek et al., 1999) and
- 82 converted into PCR-based markers like sequence tagged site markers (STS) (Olson et al., 1989). Since
- 83 this time, the development of alternative PCR-based assays like random amplified polymorphic DNA

84 (RAPD) (Kleinhofs et al., 1993; Williams et al., 1990) or amplified fragment length polymorphisms

- 85 (AFLP) (Becker et al., 1995; Vos et al., 1995), both of which are more frequent in the genome,
- 86 diminished the importance of RFLPs and their derived STS-markers for genetic analysis. However,
- those new markers were also labor-intensive and poorly reproducible in the case of RAPD (table 1).
- 88 The development of microsatellites (or simple sequence repeats, SSR) (Becker and Heun, 1995; Litt
- and Luty, 1989; Saghai Maroof et al., 1994) was another PCR-facilitated breakthrough for genotyping.
- 90 Based on the length of a tandemly repeated nucleotide motif of 1 to 6 bases, SSR markers have been
- 91 the markers of choice for trait mapping for years and are still used in numerous current projects
- 92 because of their advantages: abundant, codominant, polyallelic universal, robust, easy to automate
- and transferable between genotypes (Gupta and Varshney, 2000).

Figure 1: Cumulative number of papers per year per type of marker. Based on the number of papers
in PubMed using the word "barley" as well as the name of the different markers in their title or
abstract.

97

Genotyping entered the high-throughput era with diversity array technology (DArT) (Jaccoud et al., 98 99 2001), developed in 2004 for barley (Wenzl et al., 2004). This hybridization-based method generates 100 whole-genome fingerprints by scoring the presence/absence of DNA fragments and provided 101 genotype info at several thousand loci in parallel. DArT markers are reliable and were extensively 102 used in several species for establishing high-density consensus maps (Wenzl et al., 2006). But, with 103 the development of automated, multiplex sequencing technologies, and the inherent reduction of 104 costs, higher-throughput genotyping technologies emerged. Through Sanger sequencing technology, 105 expansive expressed sequence tag (EST) (Zhang et al., 2004) datasets were produced from different 106 genotypes. They were gathered in databases like HarvEST (http://www.harvest-web.org/) and 107 provided the first generic genome-wide datasets for single nucleotide polymorphism (SNP) mining 108 and marker development, eventually leading to the development of the first SNP arrays in barley. 109 SNPs represent the most abundant source of variation in the genome, making them very good 110 polymorphisms for gene mapping. SNP arrays allow the genotyping of a large number of 111 polymorphisms by array capture of total DNA, but require prior SNP discovery through DNA or RNA 112 sequencing. Successive EST assemblies provided such a basis for the development of Illumina 113 platforms BOPA1 and BOPA2, later merged into the Illumina GoldenGate assay containing almost 3,000 SNPs (Close et al., 2009; Muñoz-Amatriaín et al., 2011). The barley SNP array resources became 114 115 precursors to higher density formats like the Illumina 9K iSelect chip (Comadran et al., 2012) and, more recently, the 50k iSelect SNP array (Bayer et al., 2017) that allows for parallel genotyping of 116 117 almost 45,000 SNPs.

118 The SNP arrays allow for high-throughput genotyping, but their production requires a lot of resources 119 and includes a bias toward alleles that were present in the population used for polymorphism 120 detection. Further improvement of NGS technologies introduced an ongoing decline in costs. This 121 provided the basis for establishing the first barley genome sequences (International Barley Genome 122 Sequencing Consortium, 2012; Mascher et al., 2017), which enable sequence-based genotyping at 123 the genome scale. This has removed the bottleneck of the marker discovery process and allows for 124 direct SNP identification in the population of interest. Deep sequencing of the whole barley genome 125 entails a heavy sequencing load, and its resolution would be higher than that needed for the purpose 126 of SNP calling and genetic mapping. Here, several complexity reduction methods helped to introduce 127 more cost-efficient solutions. One is to sequence only the coding fraction of the genome, either

128 through RNAseq or through exome capture sequencing (Mascher et al., 2013). Exome capture is 129 based on the use of oligonucleotide baits specific to a set of exons predicted from the sequence of 130 cv. Morex (International Barley Genome Sequencing Consortium, 2012). Those baits are then 131 hybridized to genomic DNA to capture the respective gene space for sequencing. The region targeted 132 for sequencing is thus greatly reduced, and sequencing load commensurately so, allowing for higher 133 coverage sequencing and, ultimately, highly accurate genotype calling. However, even if this is an 134 efficient method for some applications (Mascher et al., 2013), it is labor-intensive. Another approach 135 involves randomly selecting regions for sequencing based on the presence of restriction sites, greatly 136 lowering the cost and labor intensity. Three such methods are used in barley: RAD-seq 137 (Chutimanitsakun et al., 2011), genotyping by sequencing (GBS) (Poland et al., 2012) and DArTseq 138 (DArT sequencing: www.diversityarrays.com). RAD-seq consists in cutting the DNA with restriction 139 enzymes like Sbfl or EcoRI (Baird et al., 2008). The fragments are then barcoded through ligation to 140 adapters, and the samples from several individuals are pooled. Subsequent steps are performed on 141 the pooled samples, decreasing the labor and cost. The pooled fragments are sheared and size-142 selected to achieve a length suitable to the sequencing platform, followed by a selective PCR used to 143 select only the fragment harboring the restriction site. GBS follows a similar protocol, but with fewer 144 steps. It utilizes two enzymes: a rare-cutter (Pstl) and a frequent-cutter (Mspl). Fragments are 145 barcoded as follows: forward adaptors are designed with the Pstl adaptor whereas reverse adaptors 146 are matching Mspl. Unlike RAD-seq, no shearing or size selection step is required after pooling the 147 fragments. The short Pstl-Mspl fragments are instead selected by a selective PCR amplification step 148 with a short extension time, and sequenced on an NGS platform. DArTseq has a similar protocol to 149 GBS, but uses different enzymes in pairs with PstI (Wenzl et al., 2004). Both GBS and DaRTseq 150 complexity reduction are slightly deviated to genic regions because of the use of enzymes more 151 efficient in under-methylated regions.

152 The prevalence of these different methods over time is illustrated on figure 1, and their distinctive 153 characteristics are listed in table 1. The polymorphisms detected by these sequencing-based 154 methods can be used directly for mapping, or to design markers to genotype a larger number of plants with a lower marker density. Those markers are classically used in the process of high-155 156 resolution mapping (see part 6). The most popular ones are *cleaved amplified polymorphic sequences* 157 (CAPS) (Konieczny and Ausubel, 1993), but newer methods like detection of SNP-specific 158 hybridization by fluorophores is growing. Of those systems, the two most frequent are Taqman 159 (Holland et al., 1991) and KASPar (LGC genomics, Semagn et al., 2014) markers. The genotype at a locus is determined in a PCR with pairs of fluorescently labeled primers or probes, with one of each 160 161 pair being specific to an allele. After a single PCR, the reading of the fluorescence is performed in a 162 qPCR thermal cycler or a fluorescent plate reader. The relative fluorescence levels from each 163 member of a marker pair allows clear distinction of heterozygotes and homozygotes. These systems 164 thus allow a rapid, reliable and efficient genotyping of a large number of samples at a limited number 165 of loci.

166 Table 1: Characteristics of the most used markers in barley, currently and in the past.

167 3 Classical mapping strategies - still efficient but improving in the genomics era

Access to high-throughput genotyping conferred considerable benefits for linkage mapping. It
 accelerated data generation for mapping by generating high numbers of markers in a single

170 experiment, and allowed an increase in population size owing to reduced labor intensity in 171 comparison to previous genotyping methods. Linkage mapping of a given trait is usually performed 172 with the help of designed bi-parental mapping populations. The first step of mapping is to assess 173 whether the trait is controlled primarily by a single major gene, or if it is a quantitative trait, likely 174 influenced by many genes. This is determined by studying the segregation and distribution of the 175 phenotype in developed populations. If the trait is controlled a single gene, the segregation pattern is 176 also the way to determine the status of dominance of the gene. A trait governed by a single gene can 177 be mapped in a simple population like F_2 or back-cross (BC). But in case of a quantitative trait, the 178 unmonitored segregation in the numerous heterozygous regions limits the replication of 179 phenotyping and can impact its reproducibility, ultimately reducing the power of QTL detections. 180 Recombinant inbred lines (RILs) (Burr et al., 1988), obtained by repeatedly selfing F2 lines, were 181 created to increase the total number of recombination events while obtaining mostly homozygous 182 lines. Due to this last characteristic, phenotyping can then be repeated several times with a high level 183 of reproducibility. However, the generation of this kind of population is tedious as it requires 6 to 8 184 generations and has encouraged the development of faster ways to obtain homozygous lines. One of 185 these is the doubled haploid (DH) line (Clapham et al., 1973; Kao et al., 1991), which are produced by 186 inducing haploid lines from F1 plants and converting them into diploids. DH lines are based on a 187 single gamete, therefore they only have one meiosis to accumulate recombinations. The total 188 number of recombination is therefore lower than in RILs, but it is the fastest way to obtain

189 homozygous lines and is now widely used (see chapter 7).

190 Linkage mapping is used for both single genes and QTLs and can be performed with different 191 algorithms. The single-marker analysis identifies the gene interval or QTLs though calculation of the 192 effect of each marker of the trait by t-test or ANOVA (Beckmann and Soller, 1988; Edwards et al., 1987; 193 Thoday, 1961), but it lacks precision. Interval mapping (Haley et al., 1992; Lander and Botstein, 1989) 194 is based on the simultaneous analysis of two markers at a time, where the two marker positions define 195 an interval within which the likelihood of a gene or QTL being present is estimated. The resulting 196 position and effect of each detected locus is thus more precise. The composite interval mapping (CIM) (Zeng, 1993) approach takes into account that a trait can be influenced by several QTLs, thus reducing 197 198 the error. To improve the power of linkage mapping, several statistical models, such as multiple 199 interval mapping (Kao et al., 1999), were derived from CIM, including more factors like additive, 200 dominance or epistasis effects.

The availability of high-throughput genotyping and access to a reference genome sequence of barley now allows better linkage mapping by increasing the number of markers included in such studies, while making reference-guided mapping possible, using the information of the reference sequence to define the order of markers. But, even if increasing the number of markers can improve the mapping resolution, this improvement is limited by the number of recombination events that occurred at the locus, and the genetic resolution depends mostly on the size of the population and the frequency of recombination.

208 4 The association mapping boom

The genetic resolution achieved by linkage mapping is limited. It is directly linked to the
recombination frequency in the population, which remains low because of the low number of
generations that can be achieved. To alleviate this problem a natural population can be used to

212 perform genome wide association studies (GWAS). First used in humans where experimental

- populations cannot be developed (Spielman et al., 1993), it also has advantages in plants traits
 mapping: The recombination events that can be identified in natural populations are the historical
- recombinations that occurred during the evolutionary history of the species, representing a much
- 216 higher number of generations than the few collected in experimental populations (Zhu et al.,
- 217 2008). High-throughput genotyping allows to easily retrieve a large number of polymorphisms in a
- 218 population and the genomic era became the one of GWAS. This mapping strategy is detailed in
- chapter 8 and will only be generally described here.
- 220 GWAS can be applied for major genes but the kind of population used make it better suited to
- 221 quantitative traits. It is taking advantage of the linkage disequilibrium (LD) that exists between the
- causative locus and genotyped loci nearby to identify small haplotype blocks associated with the
- 223 phenotype of interest in the population. The resolution depends on the extent of LD in the genomic
- region of interest, and on the density of the genotyping. The optimal marker density is then linked to
- LD in the population: the higher the LD extent, the lower the optimal number of markers. One big
- advantage of GWAS is the use of a population that does not need to be created by crossing, saving
- time. However, the use of natural populations can also be a drawback as LD can be affected by the
- 228 genetic structure of the population. In each subpopulation, allele frequencies evolve independently,
- creating spurious haplotype associations caused by the happenstance over-representation of the
- phenotype in a subpopulation that also bears similar genotypes at many loci---and not to any true
- association between the marker and the phenotype. This bias can be controlled by adding
- parameters like genetic structure or relatedness as cofactors to the model (mixed linear model,
 MLM) (Yu et al., 2006). These corrections may also increase the risk of false negative associations, as
- 255 Willing (Tu et al., 2000). These confections may also increase the risk of faise negative associations, as
- traits really linked to the population structure will not be detected. Other models like CMLM (Zhang
 et al., 2010), ECMLM (Li et al., 2014) or MLMM (Segura et al., 2012) were developed to try to further
- 236 compensate for population structure.
- 237 In linkage mapping, the size of the population is crucial, as it counterbalances population structure.
- 238 Compared with linkage mapping, GWAS usually allows a better resolution than linkage mapping but,
- has a reduced capacity to detect effects linked to rare variants---which for certain traits can
- 240 represent the largest source of phenotypic variation. If the allele frequency at the locus is too low in
- the population, either the quality filter applied before the analysis may delete the informative
- 242 markers (the minimum allele frequency is often set at 0.05), or the statistical power of the analysis
- 243 itself will not allow detection. GWAS studies are thus better suited for either large-effect-low-
- 244 frequency or low-effect-large-frequency loci.

245 5 Multiparental populations: the perfect balance?

- 246 To combine the advantages of biparental and natural diversity populations, multiparental
- 247 populations were developed. The allele frequency and the QTL effects are higher than in a natural
- 248 diversity panels while the effects of population structure are alleviated, so it is easier to detect rare
- variants or QTLs with smaller effect compared with classical GWAS. Compared to classical QTL
- 250 mapping, having more than two parents also adds some historical recombination events to the
- 251 population, thus increasing the resolution of the QTLs. It also widens the represented diversity and
- 252 can allow access to more causal loci than those segregating in a biparental population. However,
- 253 some drawbacks of biparental and natural populations still remain. As the number of population

founders is limited, they do not represent the same proportion of the species' diversity than aclassical GWAS panel. Moreover, as for linkage mapping population, the creation of a multiparental

256 population requires several generations of crosses.

257 The two major types of multiparental populations were developed in barley: Multi-parent advanced 258 generation inter-cross (MAGIC) (Sannemann et al., 2015) and two nested association mapping (NAM) 259 (Maurer et al., 2015; Nice et al., 2016; chapter 11) populations. NAM populations are made by 260 crossing multiple inbred lines with a single one and subsequently deriving RILs or DH lines from the progeny (Yu et al., 2008). The common parent genotype normalizes the genetic background, and 261 262 common-parent specific markers are used to genotype the population. So even as diversity is 263 introduced by the diverse parental lines, the use of a common parent reduces the creation of new 264 haplotypes, since no new haplotypes are recreated between the unique parents. MAGIC populations 265 (reviewed by Huang et al., 2015) imply a greater number of crosses than NAM populations and are 266 thus more labor-intensive to make, but include more diversity. These populations are made by 267 intercrossing several founders (typically eight) and selfing the lines, then using the progeny of those 268 crosses to obtain RILs. Allele frequency is more balanced than in NAM populations and haplotype 269 diversity is higher, but these populations are more time-consuming to develop and cannot be 270 extended to additional founders as NAM populations can.

6 From an interval to the causal gene: from high-resolution mapping to gene

272 cloning

273 Whether linkage mapping or GWAS is used, the resolution achieved is rarely sufficient to directly 274 identify the causal gene or polymorphism. High-resolution mapping of the locus of interest is most 275 often needed to reduce the region and obtain an interval containing a limited number of candidate 276 genes. However, in the case of a quantitative trait, not all QTLs are suitable for high-resolution 277 mapping. QTLs with small effects or low statistical support should be avoided, as the chance of 278 success in high-resolution mapping is low. High-resolution mapping relies on increasing 279 recombination in a gene interval saturated with markers. Single-gene traits can be mapped at a high 280 resolution in F2 populations. To obtain higher resolution, RILs can be derived specifically from plants 281 recombining in the gene interval (segmental recombinants to obtain so-called segmental RILs). However, for quantitative traits, the locus should first be "mendelized", meaning that a high-282 283 resolution mapping population must be developed where the inheritance of the phenotypic variation 284 follow Mendel's law: two distinct phenotypic classes, easily distinguishable, must be obtained. To 285 reach this goal, the number of segregating loci affecting the trait must be reduced by inserting the 286 two alleles at a locus in a common genetic background. Near isogenic lines (NILs) are thus produced. 287 NILs are preferably homozygous lines, often containing a single locus from a donor parent (RILs, DHs, 288 F_2/F_3 , or original accessions containing the trait of interest) in the genetic background of a 289 phenotypically distinct recurrent parent. NILs are produced by crossing those two parents, then 290 performing backcrossing and/or selfing. A population is then created by crossing the obtained NIL 291 with the recurrent parent, and high-resolution mapping is performed in this new population.

With the tools of today, saturating a locus interval with markers is usually unproblematic. Most of the time, high-resolution mapping will result in an interval containing several candidate genes. The next step is typically a candidate gene approach. With the availability of an annotated reference genome, retrieving lists of candidate genes is straightforward. If one (or a small number) of those 296 genes stands out for its predicted function, it can be prioritized for further analysis. However, this 297 simple case may not occur. It is possible that the causal gene has no obvious predicted function. 298 Another possible reason is that phenotypic traits can be due to presence/absence variations (PAV) 299 (Feuk et al., 2006) and the sequencing of the interval in parents may still be needed to find the right 300 candidate gene. With the decreasing cost of de novo assemblies, we will, in a very near future, see 301 even several high quality genome sequences published per species. Already initiated in maize (Lu et 302 al., 2015; Hirsch et al., 2016; Yang et al., 2017; Sun et al., 2018; Springer et al., 2018), this 303 development empowers us with possibilities to compare the structure and genes present in mapping 304 intervals in a greater proportion of the diversity, and thus decreasing the risk of missing the causal 305 gene that can be absent in one genome (Monat et al., 2018; Worley et al., 2017). To identify the 306 polymorphisms that can explain the phenotypic differences between parents, a preferred candidate 307 gene (or the complete set of genes in the interval) is usually resequenced. Since the trait can also be 308 due to differential gene expression, it is possible to investigate gene expression differences by using 309 microarray (Close et al., 2004), qPCR or RNA-seq (Costa-Silva et al., 2017).

310 Such approaches allow a researcher to identify a reduced number of candidate causal genes for

further validation. A common gene validation method in barley is mutant analysis to identify the

312 effects of mutations that impair or alter a gene's function. This approach typically relies on methods

- 313 like RNA interference, chemical or physical mutagenesis or T-DNA insertion. In the case of chemical
- and physical mutagenesis, the mutation is not targeted and screening of the mutants can be
- 315 performed by targeting induced local lesions in genomes (TILLING, later detailed). Another approach
- is the use of genetic transformation to either overexpress the candidate gene in a plant without the
- 317 trait of interest, or to supplement a mutant. Contemporary genome-editing technologies like TALENS
- 318 (Gurushidze et al., 2014; Wendt et al., 2013) and CRIPR-Cas9 (Lawrenson et al., 2015) are becoming a
- 319 method of choice to generate mutants in barley at specific loci (Lawrenson et al., 2015; Holme et al.,
- 320 2017; Kapusi et al., 2017; Kumar et al., 2018). However, these methods should allow us to achieve a
- 321 larger spectrum of genome editing in a near future, from complete gene insertion or knockout to
- 322 single base replacement (Bortesi and Fischer, 2015), which would allow to unravel the role of each
- 323 candidate polymorphism in the trait by single nucleotide edition.

324 7 Emerging mapping strategies: fast NGS-enabled technologies

325 While the standard methods described so far are very efficient, even faster methods were 326 introduced in the genomics era. The classic quantitative genetic method to identify rapidly a gene or 327 QTL interval is by bulk segregant analysis (BSA) (Giovannoni et al., 1991; Michelmore et al., 1991). It 328 is based on the genotyping of individuals selected from the two tails of the phenotypic distribution of a mapping population or of a diversity set (Figure 2). Plants from both ends of the distribution can be 329 330 genotyped individually, but further economic savings can be made by genotyping pooled DNA from 331 groups of individuals with similar phenotypes. The frequency or the strength of signal at markers 332 near the causal locus shows a significant deviation between the two pools, allowing identification of 333 the genes or QTLs loci, with a very limited genotyping effort. BSA has been used for the first time in 334 the early 90s in barley (Barua et al., 1993) but, while early applications were confounded by 335 insufficient marker density resulting in a high rate of false positives, state-of-the-art high-throughput 336 genotyping methods have alleviated this problem. The accurate phenotyping of the entire original 337 population and the real cost and labor are the only remaining limitations (Gallais et al., 2007; Sun et 338 al., 2010).

Page **9** of **31**

339 Figure 2: Schematic representation of Bulk Segregant Analysis (BSA). A population is generated by

crossing two parents with distinct phenotypes. The offspring are phenotyped and two bulks,

341 comprising the individuals presenting the most extreme phenotypes of each class, are created. The

342 two bulks are sequenced and their genotypes are graphed along the chromosomes. Candidate

343 intervals are defined as loci where the two bulks show a divergent genotype.

344 Another fruitful method is the use of mutants exhibiting a trait of interest, or the loss of a trait of 345 interest, to find the mutation responsible for the trait in a reverse genetics approach. TILLING 346 (McCallum et al., 2000; see chapter 10) combines chemical mutagenesis with high-throughput 347 genome-wide screening for point mutations in genes of interest. The original TILLING method used 348 denaturing HPLC for mutation discovery. For higher throughput, a protocol using Li-Cor DNA analyzer 349 was soon developed (Till et al., 2003). Until recently, barley TILLING populations were mostly 350 characterized either by dHPLC (Caldwell et al., 2004) or Li-Cor (Gottwald et al., 2009; Kurowska et al., 351 2012; Lababidi et al., 2009; Szarejko et al., 2017; Talamè et al., 2008). However, both still involve 352 inconvenient steps, such as acrylamide gel preparation and amplicon purification, and require 353 labeled primers that can impact mutation detection by reducing PCR efficiency. For these reasons, protocol using Fragment AnalyzerTM (Advanced analytical technologies) were developed (Kang et al., 354 355 2018; Szurman-Zubrzycka et al., 2018). A further future direction of TILLING in barley is NGS-356 sequencing-based approaches. Such a protocol was developed by Tsai et al. (2011), consisting in 357 sequencing different PCR target amplicons in DNA pools. This method is still expensive and so far 358 restricted to species where the mutation density is high, like wheat (Krasileva et al., 2017). With 359 decreasing costs of sequencing, it is probable that TILLING will move to detection of mutations by 360 sequencing in a near future.

361

362 Combining sequencing and mutant screening is the goal of mapping-by-sequencing, following 363 SHOREmap (Schneeberger et al., 2009) or similar analysis like MutMap (Abe et al., 2012). The idea is 364 to combine EMS mutagenesis with whole genome sequencing in an approach similar to BSA. A 365 mutant is generated, most often by EMS. As EMS mutation causes several mutations in the genome, 366 the simple sequencing of one mutant is not sufficient to identify the causal mutation; the mutant 367 must be crossed with the non-mutated parent and a BSA analysis is performed. Complete sequencing 368 of the genome of the mutants allows to retrieve all polymorphisms (Abe et al., 2012; Schneeberger et al., 2009). Since the barley genome is large and complex, less expensive methods using genome 369 370 complexity reduction were developed.

371 Mapping-by-sequencing was developed and performed successfully in barley (Jost et al., 2016; Mascher et al., 2014; Pankin et al., 2014). After phenotyping, plants with the mutant phenotypes and 372 373 those with wild type phenotypes were separated into two pools. DNAs from both pools were 374 subjected to exome capture (Mascher et al., 2013) as a complexity reduction method and 375 subsequently sequenced on a Illumina platform. The obtained reads were mapped against the 376 reference genome and SNPs are detected. Allele frequency in the two pools was then mapped along 377 the genome and a clear imbalance of allele frequency between the two pools could be observed. This 378 method allows precise mapping of a gene with one sixth of the sequencing load needed for similar 379 resultion using WGS sequencing. However, the barley exome capture assay published in Mascher et 380 al. (2013) was estimated to capture around 86 % of the high-confidence exons and no exome capture 381 assay can capture 100 % of the genes, thus there is always an increased risk of missing the target 382 gene by this approach.

383 For the specific case of pathogen resistance, Steuernagel et al. (2016) described a method combining 384 sequencing of captured targets in the genome , and mutagenesis to clone genes. Successfully applied 385 in bread wheat, this method, called MutRenSeq, aimed at cloning dominant pathogen resistance 386 genes. Those genes are most frequently represented by NLR genes, whose structure is highly 387 conserved in the genome and can be enriched in sequencing libraries and specifically sequenced in a 388 capture assay called RenSeq (Jupe et al., 2013). MutRenSeq consists in producing loss of resistance 389 mutants by EMS mutagenesis. Independent susceptible mutant lines are then sequenced using the 390 RenSeq protocol (Figure 3). The NLR genes are assembled and aligned to those of the resistant 391 parent. If a gene harbors polymorphisms in all susceptible mutant lines, it is identified as the candidate gene for resistance. The application of RenSeq reduces the amount of data to be 392 393 sequenced to less than 0.1% of the genome. However, even in the case of dominant resistance, the 394 causal gene may belong to a different gene family. Thus, there is also in this case an elevated risk of 395 missing the target gene.

396 To alleviate the risk of missing the target gene, either because it is missing from the capture but also 397 in case it is due to PAV or large structural variation that cannot be detected by conventional mapping 398 to a reference genome, methods using local *de novo* assembly were designed. One of these methods 399 is MutMap-Gap that combines the sequencing reads mapped in the gene interval with the unmapped 400 reads to perform a small-scale de novo assembly (Takagi et al., 2013). But while this method is 401 efficient in a small genome like rice, the sequencing load to achieve a sufficient coverage in barley 402 would be very high and a complexity reduction method is required. Taking advantage of 403 chromosome flow-sorting technology and efficient de novo assembly methods (ChromSeq) (The 404 International Wheat Genome Sequencing Consortium, 2014), Sánchez-Martín et al. (2016) proposed 405 a method called MutChromSeq. It consists in generating EMS mutants and screening those for loss of 406 the desirable trait, then flow-sorting the chromosomes of the interesting mutants to select the 407 chromosome that contains the locus of interest (Figure 3). Then, this single chromosome is 408 sequenced and assembled de novo. The mutation overlap in independent mutants allows rapid and 409 precise identification of the gene underlying the trait. While chromosome flow-sorting is generally 410 complex, it is well optimized for barley (Doležel et al., 2012), reduces the cost of sequencing, and 411 achieves a greater depth of sequencing to perform a better assembly. These methods are very 412 efficient for genes producing strong phenotypes, but more quantitative traits still elude such 413 approaches. Moreover, they depend on the identification of loss-of-function mutants, which can be 414 difficult to obtain. Targeted chromosome-based cloning via long range assembly (TACCA) (Thind et 415 al., 2017) is designed to avoid these flaws. Like MutChromSeq, it is based on sequencing of the flow-416 sorted candidate chromosome and its assembly, which needs to reach high contiguity. To achieve it, 417 Dovetail Chicago(TM) sequencing libraries are prepared and sequenced in addition to the classic 418 Illumina libraries to perform long range scaffolding with state-of-the-art assembly pipelines (Putnam 419 et al., 2016). The resulting chromosome sequence is used to resolve structural variation in the gene 420 interval and to design additional markers. While slower than MutChromSeq, this method is faster 421 and more precise than a classical high-resolution mapping and can be applied to all kinds of traits.

Figure 3: Schematic representation of two fast-enabled gene mapping methods. Seeds from the
accession presenting the trait of interest are mutagenized, usually by EMS, and the mutants are
phenotyped. A few mutants presenting a loss of function for this trait are selected. MutRenSeq
consists in capturing the NLR genes in these mutants and sequencing them in a method called

426 RenSeq. NLR sequences are then *de novo* assembled and aligned to the susceptible reference. The

427 candidate genes are the NLRs presenting a mutation in all the loss-of-function mutants sequenced.

- 428 MutChromSeq and TACCA consists in performing chromosome sorting on the loss-of-function
- 429 mutants, as well as the wild type parent, to select the chromosome harboring the locus coding for
- 430 the trait of interest. These sorted chromosomes are sequenced, *de novo* assembled and aligned. The
- 431 candidate genes are the ones showing a mutation in all loss-of-function mutants.

432 8 Conservation of barley germplasm

433 The genus Hordeum comprises cultivated barley H. vulgare spp. vulgare, its wild subspecies H. vulgare spp. spontaneum, and over 30 wild relatives constituting the second and the tertiary gene 434 435 pools that can be a useful source of diversity (Blattner, 2009). Because of its economic importance, 436 barley germplasm has been collected and stored for about a century and now ranks third in terms of 437 the number of accessions kept in ex situ genebanks after wheat and rice (Commission on genetic 438 ressources for food and agriculture; FAO, 2010). Knüpffer (2009) estimated that over 450,000 439 Hordeum accessions are kept in ex situ collections distributed over 204 genebanks, including almost 440 300,000 accessions from *H. vulgare* spp. *vulgare* and around 32,000 of *H. vulgare* spp. *spontaneum*. 441 Among these genebanks, seven currently hold more than 20,000 accessions: the PGRC in Canada 442 (http://www.agr.gc.ca/pgrc-rpc), the NSGC in the US (http://www.ars-grin.gov/npgs), the EMBRAPA 443 CENARGEN in Brazil (www.embrapa.br/en/recursos-geneticos-e-biotecnologia), the ICARDA in Syria 444 (www.icarda.org/research-sub/biodiversity-and-its-utilization), the John Innes Center in the UK 445 (www.jic.ac.uk/research/germplasm-resources-unit/), the IPK Gatersleben in Germany (www.ipk-446 gatersleben.de/en/gbisipk-gaterslebendegbis-i/), and the VIR in Russia (www.vir.nw.ru/). These

resources are available to any researcher, and allow the mining of new traits, genes, and alleles.

448 From those genebanks, panels of germplasm and core collections are identified to represent the 449 largest diversity possible from a size-limited subsample of accessions. Core collections allow the 450 screening of a species' diversity with minimal effort. The International Barley Core Collection (BCC) is 451 constituted of about 1500 accessions, including some *H. vulgare* spp. spontaneum and some wild 452 relatives (Knüpffer and van Hintum, 2003). It was characterized genetically and phenotypically for 453 some traits. Other core collections include the NSGC Barley Core (Muñoz-Amatriaín et al., 2014) with 454 1860 accessions or the Landrace Collection LRC1485 that consists of 1485 spring barley landraces 455 selected specifically for the study of adaptation to climate (Pasam et al., 2014). Some of these panels 456 are also constituted to represent the diversity in the wild relatives like the Wild Barley Diversity 457 Collection that consist in 318 H. vulgare spp. spontaneum accessions (Steffenson et al., 2007) or the 458 Barley1K collection that comprises 1020 wild barleys sampled in 75 sites in Israel (Hübner et al., 459 2009). All these collections are genetically characterized, for example by the 9k iSelect array in the 460 case of the NSGC Barley Core (Comadran et al., 2012), as well as phenotypically for some traits, and 461 have been used for GWAS or for germplasm mining.

But, as carefully as a core collection is constructed, it cannot contain all the diversity contained in a species, therefore limiting their use for discovery of very rare alleles. IPK Gatersleben recently completed GBS sequencing of its entire barley collection (Milner et al., Unpublished data). The related webtool (http://bridge.ipk-gatersleben.de) permits to mine the germplasm of the barley collection and design subsets of it based on criteria like genetic or geographic diversity, but also to deliver direct access to the SNPs discovered via an intuitive visual interface. This initiative aims at unlocking the collection for allele mining and diversity studies and is already applied to other genebanks, expanding the genomic information of genetic resources and allowing association mapping to be performed on asubset of a genebank without additional genotyping.

471 9 Genetic and genomic resources of barley

Characterization of barley germplasm has been ongoing for a long time and resources are stored in
online databases. GrainGenes (https://wheat.pw.usda.gov/GG3/barley_blvd) (Carollo et al., 2005) is
a database harboring molecular and phenotypic information of Triticeae and Avena, including genetic
maps, markers and germplasm information. Some databases like PLEXdb (http://www.plexdb.org/)
(Wise et al., 2007) or GENEVESTIGATOR (https://genevestigator.com/gv/doc/content.jsp) provide
access to the pattern of expression of the genes represented on the Affymetrix microarray. RNA-Seq
data from Morex cultivar is also available at morexGenes (https://ics.hutton.ac.uk/morexGenes) as

- 479 well as a barley epigenome browser (https://ics.hutton.ac.uk/barley-epigenome).
- 480 The systematic sequencing of the barley genome (International Barley Genome Sequencing
- 481 Consortium, 2012; Mascher et al., 2017; Mayer et al., 2011) ran in parallel with the development of
- 482 new tools to access the data. BARLEX (http://barlex.barleysequence.org) (Colmsee et al., 2015)
- 483 includes the different types of data produced for the barley sequence assembly (Mascher et al.,
- 484 2017). Centered on the minimum tilling path of the sequenced BACs, it gives access to information
- such as the number of contigs and sequences of each individual BAC assembly, tables of published
- annotated genes, mapping of the 9k iSelect markers on the genomes, and a link to a BLAST server
- 487 (http://webblast.ipk-gatersleben.de/barley_ibsc/) allowing for homology searches against the
- 488 different published assemblies as well as other databases like full-length cDNA, ESTs, or regions
- 489 sequenced by exome capture.

490 A second way to access the genome assembly data is through the genome browser displayed by the 491 common tool of Gramene (http://www.gramene.org) (Tello-Ruiz et al., 2016) and Ensembl Plants 492 (http://plants.ensembl.org) (Kersey et al., 2016) where annotation of genes, as well as full cDNA and 493 transcriptomic data, can be visualized along the genome. These web tools also allow comparative 494 analysis between genomes to be performed. A third useful database for exploring barley genomic data 495 is PGSB PlantsDB (http://pgsb.helmholtz-muenchen.de/plant/barley/index.jsp) (Spannagl et al., 2016). 496 It contains visualization of the different versions of sequence assembly, information on annotated 497 genes as well as synteny comparison with the genomes of Brachypodium distachyon and Oryza sativa.

498 10 Case Study: From *rym4* to *rym11*, illustration of paradigm shift in disease 499 resistance mapping and cloning

500 To illustrate recent advances in positional cloning, we compared the mapping of two resistance 501 genes, performed in the same labs. The first is the *rym4/rym5* locus that was cloned in 2005 (Stein et 502 al., 2005). The second is *rym11*, cloned almost ten years later, in 2014 (Yang et al., 2014b). Both 503 genes confer complete recessive resistance to soil-borne Bymoviruses BaMMV and BaYMV and 504 encode for susceptibility factors.

The *rym4* gene was introgressed from landrace 'Ragusa' into commercial German barley varieties as
early as the 1980s but its genetic basis was unknown. This recessive resistance gene was proven to
be linked or allelic to *Rym1* by allelic test in a cross with *Rym1* variety Mokusekko 3 (Friedt et al.,
1987). Crosses between the *rym4* variety Franka and barley accessions exhibiting a strong dominant

509 phenotype mapped to a chromosome (genetic marker stocks) proved that rym4 was neither on 510 chromosomes 4H, 1H nor 5H. (Kaiser and Friedt, 1992, 1989), but placed rym4 distal on the long arm 511 of chromosome 3H by segregation analysis in crosses between the rym4 'Ogra' and 'Sonate' varieties 512 and trisomic lines (lines trisomic for specific chromosomes) and telotrisomic lines. From these early 513 attempts, successive innovations in molecular markers allowed the precise characterization of the 514 genetic interval. In the 1990s, several studies used RFLP, isozymes and RAPD markers and mapped 515 rym4 and rym5 in the same interval at the distal end of chromosome 3HL (Graner et al., 1999, 1995; 516 Graner and Bauer, 1993; Konishi et al., 1997; Ordon et al., 1995). Despite the increase in numbers of 517 marker types and of population sizes, the initial interval of 2.4 cM (Graner and Bauer, 1993) was not 518 significantly reduced during this decade. The first high-resolution mapping of the locus rym4/rym5 519 (Pellio et al., 2005) took advantage of advances in molecular markers to screen large populations: 520 3,884 F₂-plants for rym5, and 1,040 F₂ individuals for rym4 mapping. DNA bulks from individuals of these DH populations were used for AFLP, RAPD and RFLP marker saturation of the regions and RILs 521 522 were derived from recombinants between the common flanking markers of rym4 and rym5. The new 523 markers were converted into STSs and used to genotype the RILs, reducing rym4 and rym5 intervals 524 to 0.05 cM and 0.013 cM, eventually leading to the cloning of the gene (Stein et al., 2005). The Morex 525 BAC library (Yu et al., 2000) was screened with the markers of the rym4/rym5 region, and a 650 kb 526 physical map was constructed. The six BAC clones cosegregating with resistance were sequenced to 527 full length and annotated (Wicker et al., 2005), and two open reading frames (ORFs) were found. One 528 of them, Hv-eIF4E, was homologous to genes involved in recessive Potyvirus resistance in 529 dicotyledonous species and was sequenced in 56 barley accessions known to carry either rym4, rym5, 530 an unknown source of resistance, or no resistance. This sequencing revealed several non-531 synonymous polymorphisms in rym4 and rym5 accessions and this gene was confirmed by 532 complementation via Agrobacterium-mediated transformation of a rym4 NIL with either the full 533 length cDNA or the genomic DNA of a susceptible accession, thus inducing susceptibility. Kanyuka et 534 al. (2005), performed sequencing of RT-PCR products of different variants and evidenced that rym6 is 535 another natural allele of Hv-eIF4E. They also carried out TILLING on the gene and retrieved one 536 mutant allele, with which resistance segregates. Additional allele mining was later performed, 537 identifying several other alleles of resistance at this gene (Yang et al., 2017, 2014a).

538 In contrast, rym11 mapping was comparatively fast. The preliminary mapping of rym11 was 539 performed on two populations of 48 F₂ and 101 BC₁F₂, respectively (Bauer et al., 1997). The authors 540 performed BSA with RFLPs, RAPDs and one SSR marker, which localized the gene to within a large 541 interval of 16.4 cM, close to the centromeric region of chromosome 4HL. The first fine-mapping study (Nissan-Azzouz et al., 2005) analyzed three DH populations of 57, 191 (IPK1) and 161 (IPK2) DH lines, 542 543 respectively. The authors successively performed an initial mapping with RFLP markers, two BSA 544 analyses with RAPD and AFLP markers respectively, and a more complete mapping with SSR, RAPD 545 and AFLP markers. Based on the results, rym11 was mapped in a 3.7 cM interval in IPK1 and a 10.7 546 cM interval in IPK2 population. Lüpken et al. (2013) used a genomics-informed approach to resolve 547 the interval in a population comprising 5,102 F2 plants. To carry out the mapping, they used 548 publically available SSR and SNP markers and designed additional SNP and STS markers with the 549 information available in EST databases, in the barley genome zipper (Mayer et al., 2011), as well as in 550 the WGS sequence assembly of Morex (International Barley Genome Sequencing Consortium, 2012). 551 Polymorphisms were revealed either on gels or by pyrosequencing, leading to a final interval of 0.074 552 cM. The gene rym11 was eventually isolated by chromosome walking (Yang et al., 2014b). Barley BAC

553 clones (Ariyadasa et al., 2014; Schulte et al., 2011), sequenced in conjunction with the barley genome 554 sequencing project (Mascher et al., 2017) were identified by sequence comparison to the flanking 555 markers. Two overlapping BAC contigs were identified, resulting in a physical contig covering the 556 complete genetic interval. Based on the available BAC sequence data, new markers were designed to 557 reduce the interval to 0.0196 cM, representing 1.25 Mbp, using the same $5,102 F_2$ population 558 previously explored by Lüpken et al. (2013). Annotation of this interval revealed four open reading 559 frames (ORFs). These ORFs were resequenced, and only one gene exhibited polymorphism between 560 susceptible and resistant genotypes. This gene encoded a putative protein disulfide isomerase like 5-561 1 (HvPDIL5-1), and the resistant plants' allele contained a 1.3 kb deletion in its promoter and first exons. It was validated as the resistance gene by TILLING in an EMS-induced mutant population and 562 563 by complementation via Agobacterium-mediated transformation of the resistant genotype with the 564 cDNA of the gene from a susceptible accession. Moreover, the gene HvPDIL5-1 was sequenced in 565 1,732 diverse barley accessions, both domesticated (*H. vulgare* ssp. vulgare) and wild (*H. vulgare* ssp. 566 spontaneum), and identified three more natural haplotypes conferring resistance, all harboring a 567 premature stop codon in HvPDI5-1. This allele mining was later completed in a larger collection (Yang 568 et al., 2017).

569 These studies illustrate how developments in molecular marker technology have revolutionized gene 570 mapping studies, allowing the genotyping of large populations more easily and, by increasing the 571 number of useful markers available, allowing the discovery of intervals smaller than 1 cM. Increased 572 availability of sequence data represented a second game-changer. It enabled for instance the design 573 of numerous markers in the interval of rym11, sufficient to achieve a mapping resolution below 0.1 574 cM. It has also accelerated the acquisition of physical maps. The tedious chromosome walking carried 575 out for rym4 included the screening of a BAC library and the sequencing of the identified BACs, 576 whereas for rym11 the sequences were already available and BAC clones screening was performed in 577 silico. Moreover, TILLING and allele mining for both loci provided a fast way of validation as well as alternative alleles that could be of use for breeding. The rym11 mapping is now already five years old 578 579 and new methods of sequencing have rendered gene mapping studies even faster and more precise 580 (Jost et al., 2016; Mascher et al., 2014).

581 11 Conclusion and future trends

- 582 Advances in sequencing technology and the availability of the barley reference genome have
- 583 underpinned genomics-informed barley gene mapping and cloning. Genotyping is progressively
- 584 faster, easier and relatively more cost-efficient. New NGS-based gene mapping methods are in place
- 585 and more are appearing. A paradigm-changer for gene mapping in the coming years will be the
- 586 spread of *de novo* assemblies of complete barley genomes or flow-sorted chromosomes. More
- 587 reference-quality *de novo* assemblies of different barley haplotypes are likely to come as a result of
- 588 pan-genome analysis (Hurgobin and Edwards, 2017). Already initiated
- 589 (https://www.pflanzenforschung.de/de/plant-
- 590 2030/fachinformationen/projektdatenbank/strukturelle-genomvariation-haplotypendiversitaumlt-
- und-419), which will continue to unlock the diversity of barley, including structural changes, so far
- 592 unreachable with a single reference, and increasing the efficiency of gene cloning.
- In addition to classical gene and QTL mapping, the future will probably see an increase in expression
 QTL (eQTL) (Damerval et al., 1994) analysis. Barley gene expression is affected by complex regulation.

595 More than half of the genes have been found to be differently regulated between samples, and more than two thirds have been shown to exhibit alternative splicing (International Barley Genome 596 597 Sequencing Consortium, 2012). Moreover, quantitative traits are often regulated by the differential 598 expression of genes. In eQTL analysis, gene expression is quantified, and related to the phenotype. 599 Whole genome eQTL analysis has been performed on barley (Potokina et al., 2008) using RNA 600 binding on an Affimetrix chip. However, it is now possible to perform it by RNA-seq, achieving a 601 better throughput. Moreover, the methylome can also be sequenced with next generation 602 sequencing. Called whole genome bisulfite-sequencing (BS-seq), this method is based on sodium-603 bisulfite treatment, that converts unmethylated cytosine of the genomic DNA to uracil, followed by 604 whole genome sequencing. With this method, a single base methylome map of Arabidopsis thaliana 605 was generated (Lister et al., 2008). It was also performed recently on barley Morex accession (Wicker 606 et al., 2017), and could thus be applied to trait mapping in a near future.

However, despite the huge progress made in 30 years, efficient gene mapping is still facing several
bottlenecks. Time of generation during population construction, or to obtain progenies for

609 phenotyping, is slowing down a lot of projects. Watson et al. (2018) developed 'speed breeding' to

- 610 shorten the breeding cycle. This is achieved by increasing daily light exposure and harvesting the
- 611 plants at early growth stages. By this process, the number of barley generations in a year was
- 612 doubled, from three to six. A low-cost version of 'speed-breeding' makes this protocol available to
- the whole community and will probably be intensively used in the upcoming years (Ghosh et al.,
- 614 2018). Another challenge faced by trait mapping is computing power. The amount of sequencing
- data available grows as the price of sequencing decreases. Previous data analysis systems are no
- 616 longer sufficient to process the data generated (for review see Yin et al., 2017), and storage quickly
- becomes a limiting factor in a lot of labs. To keep up with modern analysis and data requirements,
- 618 labs must invest in appropriate bioinformatics platforms and trained staff.

619 As genotyping methods develop and the cost of sequencing decreases, genotyping ceases to be a 620 primary bottleneck in trait genetics. However, whether it is for GWAS or for mutant screening, many 621 of plants need to be phenotyped and this task can be arduous. Therefore, phenotyping has become a 622 more primary concern, particularly in the case of quantitative yield and stress response traits that 623 necessitate field phenotyping, and are strongly influenced by GxE effects. Progress in imaging using 624 sensors and drones (for review Tardieu et al., 2017) are beginning to alleviate this problem, but 625 development of so-called phenomics is still in its infancy and remains very costly. However, even if 626 limited in the traits to which it can be applied, phenomics will become a major part of trait mapping in 627 the future, allowing mapping of some complex traits currently far out of reach.

12 Where to look for further information

- 629 Further reading
- Bazakos et al. (2017): Annual review of plant biology: a good review on quantitative genetics
 methods. From phenotyping, to populations and QTL analysis.
- Grover and Sharma (2016): a good review on the different type of markers and their
 compared qualities.
- Jiao and Schneeberger (2017): a review summarizing the third generation sequencing that
 render assembly of barley possible.

636 637	 Rifkin (2012): a very complete book on QTL mapping, from population construction to marker selection and statistical analysis of the data 					
638	• Stein and Muehlbauer (2018): summarizes the state-of-the-art in barley genome analysis					
639						
640	Key journals/conferences					
641	• PAG (Plant and Animal Genome): the biggest conference on genetics and genomics. Held in					
642	January every year in San Diego, USA.					
643	International Barley Genetics symposium (IBSG): symposium on barley genetics held every					
644	four years					
645						
646	Theoretical and applied genetics (TAG)					
647	Molecular breeding					
648	Plant breeding					
649	Molecular Plants					
650	Plant Journal					
651	Plant Physiology					
652	• PNAS					
653	Plant Cell					
654	Nature Genetics					
655	Major international research projects					
656 657	International Barley Sequencing Consortium (IBSC)					
658	13 Acknowledgments					
659	We warmly thank Mark Timothy Rabanus-Wallace for language editing.					
660	14 References					
661						
662	Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K.,					
663	Mitsuoka, C., Tamiru, M., Innan, H., Cano, L., Kamoun, S., Terauchi, R., 2012. Genome					
665	sequencing reveals agronomically important loci in rice using Mutiviap. Nat. Biotechnol. 30, 174–8. https://doi.org/10.1038/pht.2095					
666	Arivadasa, R., Mascher, M., Nussbaumer, T., Schulte, D., Frenkel, Z., Poursarebani, N., Zhou, R.,					
667	Steuernagel, B., Gundlach, H., Taudien, S., Felder, M., Platzer, M., Himmelbach, A., Schmutzer,					
668	T., Hedley, P.E., Muehlbauer, G.J., Scholz, U., Korol, A., Mayer, K.F.X., Waugh, R., Langridge, P.,					
669	Graner, A., Stein, N., 2014. A Sequence-Ready Physical Map of Barley Anchored Genetically by					
670	Two Million Single-Nucleotide Polymorphisms. PLANT Physiol. 164, 412–423.					
6/1	https://doi.org/10.1104/pp.113.228213					
072 673	Dairu, IN.A., Eller, P.D., Alwoou, T.S., Currey, IVI.C., Stilver, A.L., Lewis, Z.A., Seiker, E.U., Cresko, W.A., Johnson, F.A., 2008, Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers					
674	PLoS One 3, e3376. https://doi.org/10.1371/iournal.pone.0003376					
675	Barua, U.M., Chalmers, K.J., Hackett, C.A., Thomas, W.T.B., Powell, W., Waugh, R., 1993.					
676	Identification of RAPD markers linked to a Rhynchosporium secalis resistance locus in barley					

677 using near-isogenic lines and bulked segregant analysis. Heredity (Edinb). 71, 177–184. 678 https://doi.org/10.1038/hdy.1993.122 Bauer, E., Weyen, J., Schiemann, A., Graner, A., Ordon, F., 1997. Molecular mapping of novel 679 680 resistance genes against Barley Mild Mosaic Virus (BaMMV). TAG Theor. Appl. Genet. 95, 1263-681 1269. https://doi.org/10.1007/s001220050691 682 Bayer, M.M., Rapazote-Flores, P., Ganal, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P.D., Thomas, W., Waugh, R., 2017. Development and Evaluation of a Barley 683 684 50k iSelect SNP Array. Front. Plant Sci. 8, 1–10. https://doi.org/10.3389/fpls.2017.01792 685 Bazakos, C., Hanemian, M., Trontin, C., Jiménez-Gómez, J.M., Loudet, O., 2017. New Strategies and 686 Tools in Quantitative Genetics: How to Go from the Phenotype to the Genotype. Annu. Rev. 687 Plant Biol. 68, 435–455. https://doi.org/10.1146/annurev-arplant-042916-040820 688 Becker, J., Heun, M., 1995. Barley microsatellites: allele variation and mapping. Plant Mol. Biol. 27, 689 835-845. https://doi.org/10.1007/BF00020238 690 Becker, J., Vos, P., Kuiper, M., Salamini, F., Heun, M., 1995. Combined mapping of AFLP and RFLP 691 markers in barley. Mol. Gen. Genet. 249, 65–73. https://doi.org/10.1007/BF00290237 692 Beckmann, J.S., Soller, M., 1988. Detection of linkage between marker loci and loci affecting 693 quantitative traits in crosses between segregating populations. Theor. Appl. Genet. 76, 228– 694 236. https://doi.org/10.1007/BF00257850 695 Blattner, F.R., 2009. Progress in phylogenetic analysis and a new infrageneric classification of the 696 barley genus Hordeum (Poaceae: Triticeae). Breed. Sci. 59, 471-480. 697 https://doi.org/10.1270/jsbbs.59.471 698 Bortesi, L., Fischer, R., 2015. The CRISPR/Cas9 system for plant genome editing and beyond. 699 Biotechnol. Adv. 33, 41–52. https://doi.org/10.1016/j.biotechadv.2014.12.006 700 Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in 701 man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32, 314–31. 702 Burr, B., Burr, F.A., Thompson, K.H., Albertson, M.C., Stuber, C.W., 1988. Gene mapping with 703 recombinant inbreds in maize. Genetics 118, 519–26. https://doi.org/10.1063/1.3634023 704 Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der 705 Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., Schulze-Lefert, P., 1997. 706 The Barley Mlo Gene: A Novel Control Element of Plant Pathogen Resistance. Cell 88, 695–705. 707 https://doi.org/10.1016/S0092-8674(00)81912-1 708 Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F., Waugh, R., 2004. A 709 structured mutant population for forward and reverse genetics in Barley (Hordeum vulgare L.). 710 Plant J. 40, 143-150. https://doi.org/10.1111/j.1365-313X.2004.02190.x 711 Carollo, V., Matthews, D.E., Lazo, G.R., Blake, T.K., Hummel, D.D., Lui, N., Hane, D.L., Anderson, O.D., 712 2005. GrainGenes 2.0. an improved resource for the small-grains community. Plant Physiol. 139, 713 643–51. https://doi.org/10.1104/pp.105.064485 714 Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A., Cistué, L., Corey, A., Filichkina, T., Johnson, 715 E.A., Hayes, P.M., 2011. Construction and application for QTL analysis of a Restriction Site 716 Associated DNA (RAD) linkage map in barley. BMC Genomics 12, 4. 717 https://doi.org/10.1186/1471-2164-12-4 718 Clapham, D., Edwards, J., Portlock, P., 1973. Haploid Hordeum plants from anthers in vitro. Z. 719 Pflanzenzucht 69, 142–155. 720 Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., 721 Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S., Varshney, R.K., Szűcs, P., Sato, 722 K., Hayes, P.M., Matthews, D.E., Kleinhofs, A., Muehlbauer, G.J., DeYoung, J., Marshall, D.F., 723 Madishetty, K., Fenton, R.D., Condamine, P., Graner, A., Waugh, R., 2009. Development and 724 implementation of high-throughput SNP genotyping in barley. BMC Genomics 10, 582. 725 https://doi.org/10.1186/1471-2164-10-582 726 Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., 727 Muehlbauer, G.J., Kleinhofs, A., Wise, R.P., 2004. A new resource for cereal genomics: 22K 728 barley GeneChip comes of age. Plant Physiol. 134, 960-8.

729 https://doi.org/10.1104/pp.103.034462 730 Colmsee, C., Beier, S., Himmelbach, A., Schmutzer, T., Stein, N., Scholz, U., Mascher, M., 2015. BARLEX – the Barley Draft Genome Explorer. Mol. Plant 8, 964–966. 731 732 https://doi.org/10.1016/j.molp.2015.03.009 733 Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Shaw, P., Bayer, M., Thomas, W., 734 Marshall, D., Hedley, P., Tondelli, A., Pecchioni, N., Francia, E., Korzun, V., Walther, A., Waugh, 735 R., 2012. Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring 736 growth habit and environmental adaptation in cultivated barley. Nat. Genet. 44, 1388–1392. 737 https://doi.org/10.1038/ng.2447 738 Commission on genetic ressources for food and agriculture; FAO, 2010. The state of ex situ 739 conservation, in: The Second Report on the State of the World's Plant Genetic Resources for 740 Food and Agriculture. Rome, pp. 54–90. 741 Costa-Silva, J., Domingues, D., Lopes, F.M., 2017. RNA-Seq differential expression analysis: An 742 extended review and a software tool. PLoS One 12, e0190152. 743 https://doi.org/10.1371/journal.pone.0190152 744 Damerval, C., Maurice, A., Josse, J.M., de Vienne, D., 1994. Quantitative trait loci underlying gene 745 product variation: a novel perspective for analyzing regulation of genome expression. Genetics 746 137, 289-301. Doležel, J., Vrána, J., Šafář, J., Bartoš, J., Kubaláková, M., Šimková, H., 2012. Chromosomes in the flow 747 748 to simplify genome analysis. Funct. Integr. Genomics 12, 397–416. 749 https://doi.org/10.1007/s10142-012-0293-0 750 Edwards, M.D., Stuber, C.W., Wendel, J.F., 1987. Molecular-marker-facilitated investigations of 751 quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. 752 Genetics 116, 113-25. FAOSTAT, 2018. Statistical databases. Food Agric. Organ. United Nations. 753 754 Feuk, L., Carson, A.R., Scherer, S.W., 2006. Structural variation in the human genome. Nat. Rev. 755 Genet. 7, 85–97. https://doi.org/10.1038/nrg1767 Friedt, W., Kaiser, R., Götz, R., Umbach, H., Foroughi-Wehr, B., 1987. Genetic basis of breeding for 756 757 resistance to barley yellow mosaic virus (BaYMV), in: Cooper, J.I., Asher, M.J.C. (Eds.), 758 Proceedings of a Conference at the University of St Andrews. Association of Applied Biologists, 759 Warwick, UK. 760 Gallais, A., Moreau, L., Charcosset, A., 2007. Detection of marker–QTL associations by studying 761 change in marker frequencies with selection. Theor. Appl. Genet. 114, 669–681. 762 https://doi.org/10.1007/s00122-006-0467-z 763 Giovannoni, J.J., Wing, R.A., Ganal, M.W., Tanksley, S.D., 1991. Isolation of molecular markers from 764 specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic 765 Acids Res. 19, 6553–6568. https://doi.org/10.1093/nar/19.23.6553 766 Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., Green, P., Hafeez, A., Hayta, S., Melton, R.E., Steed, A., 767 768 Sarkar, A., Carter, J., Perkins, L., Lord, J., Tester, M., Osbourn, A., Moscou, M.J., Nicholson, P., 769 Harwood, W., Martin, C., Domoney, C., Uauy, C., Hazard, B., Wulff, B.B.H., Hickey, L.T, 2018. 770 Speed breeding in growth chambers and glasshouses for crop breeding and model plant 771 research. Nature protocols, 13(12), 2944. 772 Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., Stein, N., 2009. TILLING in the two-rowed barley 773 cultivar "Barke" reveals preferred sites of functional diversity in the gene HvHox1. BMC Res. 774 Notes 2, 258. https://doi.org/10.1186/1756-0500-2-258 775 Graner, A., Bauer, E., 1993. RFLP mapping of the ym4 virus resistance gene in barley. Theor. Appl. 776 Genet. 86, 689-693. https://doi.org/10.1007/BF00222657 777 Graner, A., Bauer, E., Kellermann, A., Proeseler, Wenzel, G., Ordon, F., 1995. RFLP analysis of 778 resistance to the barley yellow mosaic virus complex. Agronomie 15, 475–479. 779 https://doi.org/10.1051/agro:19950716 780 Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G., Herrmann,

781 R.G., 1991. Construction of an RFLP map of barley. Theor. Appl. Genet. 83, 250–256. 782 https://doi.org/10.1007/BF00226259 783 Graner, A., Streng, S., Kellermann, A., Schiemann, A., Bauer, E., Waugh, R., Pellio, B., Ordon, F., 1999. 784 Molecular mapping and genetic fine-structure of the rym5 locus encoding resistance to 785 different strains of the Barley Yellow Mosaic Virus Complex. TAG Theor. Appl. Genet. 98, 285– 786 290. https://doi.org/10.1007/s001220051070 Grover, A., Sharma, P.C., 2016. Development and use of molecular markers: past and present. Crit. 787 788 Rev. Biotechnol. 36, 290-302. https://doi.org/10.3109/07388551.2014.959891 789 Gupta, P.K., Varshney, R.K., 2000. The development and use of microsatellite markers for genetic 790 analysis and plant breeding with emphasis on bread wheat. Euphytica 113, 163–185. 791 https://doi.org/10.1023/A:1003910819967 792 Gurushidze, M., Hensel, G., Hiekel, S., Schedel, S., Valkov, V., Kumlehn, J., 2014. True-Breeding 793 Targeted Gene Knock-Out in Barley Using Designer TALE-Nuclease in Haploid Cells. PLoS One 9, 794 e92046. https://doi.org/10.1371/journal.pone.0092046 795 Haley, C.S., Knott, S.A., others, 1992. A simple regression method for mapping quantitative trait loci 796 in line crosses using flanking markers. Heredity (Edinb). 69, 315-324. 797 Hansson, M., Komatsuda, T., Stein, N., Muehlbauer, G.J., 2018. Molecular Mapping and Cloning of 798 Genes and QTLs. pp. 139–154. https://doi.org/10.1007/978-3-319-92528-8_10 799 Hirsch, C. N., Hirsch, C. D., Brohammer, A. B., Bowman, M. J., Soifer, I., Barad, O., Shem-Tov, D., 800 Baruch, K., Lu, F., Hernandez, A. G., Fields, C. J., Wright, C. L., Koehler, K., Springer, N. M., 801 Buckler, E., Buell, C. R., de Leon, N., Kaeppler, S. M., Childs, K. L., Mikel, M. A., 2016. Draft 802 Assembly of Elite Inbred Line PH207 Provides Insights into Genomic and Transcriptome Diversity 803 in Maize. The Plant Cell, 28(11), 2700-2714. 804 Holland, P.M., Abramson, R.D., Watson, R., Gelfand, D.H., 1991. Detection of specific polymerase 805 chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA 806 polymerase. Proc. Natl. Acad. Sci. 88, 7276–7280. https://doi.org/10.1073/pnas.88.16.7276 807 Holme, I.B., Wendt, T., Gil-Humanes, J. et al. Plant Mol Biol (2017) 95: 111. 808 https://doi.org/10.1007/s11103-017-0640-6 809 Huang, B.E., Verbyla, K.L., Verbyla, A.P., Raghavan, C., Singh, V.K., Gaur, P., Leung, H., Varshney, R.K., 810 Cavanagh, C.R., 2015. MAGIC populations in crops: current status and future prospects. Theor. 811 Appl. Genet. 128, 999-1017. https://doi.org/10.1007/s00122-015-2506-0 812 Hübner, S., Höffken, M., Oren, E., Haseneyer, G., Stein, N., Graner, A., Schmid, K., Fridman, E., 2009. 813 Strong correlation of wild barley (Hordeum spontaneum) population structure with 814 temperature and precipitation variation. Mol. Ecol. 18, 1523-36. 815 https://doi.org/10.1111/j.1365-294X.2009.04106.x Hurgobin, B., Edwards, D., 2017. SNP Discovery Using a Pangenome: Has the Single Reference 816 817 Approach Become Obsolete? Biology (Basel). 6, 21. https://doi.org/10.3390/biology6010021 818 International Barley Genome Sequencing Consortium, 2012. A physical, genetic and functional 819 sequence assembly of the barley genome. Nature 491, 711–6. 820 https://doi.org/10.1038/nature11543 821 Jaccoud, D., Peng, K., Feinstein, D., Kilian, A., 2001. Diversity arrays: a solid state technology for 822 sequence information independent genotyping. Nucleic Acids Res. 29, E25. 823 https://doi.org/10.1371/journal.pone.0019379 824 Jiao, W.B., Schneeberger, K., 2017. The impact of third generation genomic technologies on plant 825 genome assembly. Curr. Opin. Plant Biol. 36, 64–70. https://doi.org/10.1016/j.pbi.2017.02.002 826 Jost, M., Taketa, S., Mascher, M., Himmelbach, A., Yuo, T., Shahinnia, F., Rutten, T., Druka, A., 827 Schmutzer, T., Steuernagel, B., Beier, S., Taudien, S., Scholz, U., Morgante, M., Waugh, R., Stein, N., 2016. A homolog of Blade-On-Petiole 1 and 2 (BOP1/2) controls internode length and 828 829 homeotic changes of the barley inflorescence. Plant Physiol. 171, pp.00124.2016. 830 https://doi.org/10.1104/pp.16.00124 831 Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G.J., Maclean, D., Cock, P.J., 832 Leggett, R.M., Bryan, G.J., Cardle, L., Hein, I., Jones, J.D.G., 2013. Resistance gene enrichment

833 sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant 834 genomes and rapid mapping of resistance loci in segregating populations. Plant J. 76, 530–544. 835 https://doi.org/10.1111/tpj.12307 836 Kaiser, R., Friedt, W., 1992. Gene for Resistance to Barley Mild Mosaic Virus in German Winter Barley 837 Located on Chromosome 3L. Plant Breed. 108, 169–172. https://doi.org/10.1111/j.1439-838 0523.1992.tb00116.x 839 Kaiser, R., Friedt, W., 1989. Chromosomal location of resistance to barley yellow mosaic virus in 840 German winter-barley identified by trisomic analysis. Theor. Appl. Genet. 77, 241–245. 841 https://doi.org/10.1007/BF00266193 842 Kang, H.S., Kim, S.H., Lee, S.W., Kim, S.W., Ryu, J., Kim, J.-B., Yeom, S.-I., Kang, S.-Y., Jo, Y.D., 2018. 843 Optimization of TILLING system based on capillary electrophoresis for targeted selection of 844 pepper gene mutants. Hortic. Environ. Biotechnol. 59, 447–460. 845 https://doi.org/10.1007/s13580-018-0049-4 Kanyuka, K., Druka, A., Caldwell, D.G., Tymon, A., McCallum, N., Waugh, R., Adams, M.J., 2005. 846 847 Evidence that the recessive bymovirus resistance locus rym4 in barley corresponds to the 848 eukaryotic translation initiation factor 4E gene. Mol. Plant Pathol. 6, 449-58. 849 https://doi.org/10.1111/j.1364-3703.2005.00294.x 850 Kao, C.H., Zeng, Z.B., Teasdale, R.D., 1999. Multiple interval mapping for quantitative trait loci. Genetics 152, 1203-16. https://doi.org/10.1016/0046-8177(92)90086-I 851 Kao, K.N., Saleem, M., Abrams, S., Pedras, M., Horn, D., Mallard, C., 1991. Culture conditions for 852 853 induction of green plants from barley microspores by anther culture methods. Plant Cell Rep. 9, 854 595–601. https://doi.org/10.1007/BF00231796 Kapusi, E., Corcuera-Gómez, M., Melnik, S., Stoger, E., 2017. Heritable Genomic Fragment Deletions 855 856 and Small Indels in the Putative ENGase Gene Induced by CRISPR/Cas9 in Barley. Frontiers in 857 Plant Science, 8, 540. doi:10.3389/fpls.2017.00540 858 Kersey, P.J., Allen, J.E., Armean, I., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., 859 Falin, L.J., Grabmueller, C., Humphrey, J., Kerhornou, A., Khobova, J., Aranganathan, N.K., 860 Langridge, N., Lowy, E., McDowall, M.D., Maheswari, U., Nuhn, M., Ong, C.K., Overduin, B., 861 Paulini, M., Pedro, H., Perry, E., Spudich, G., Tapanari, E., Walts, B., Williams, G., Tello-Ruiz, M., 862 Stein, J., Wei, S., Ware, D., Bolser, D.M., Howe, K.L., Kulesha, E., Lawson, D., Maslen, G., Staines, 863 D.M., 2016. Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res. 44, 864 D574–D580. https://doi.org/10.1093/nar/gkv1209 Kleinhofs, A., Kilian, A., Saghai Maroof, M.A., Biyashev, R.M., Hayes, P., Chen, F.Q., Lapitan, N., 865 866 Fenwick, A., Blake, T.K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J.D., Hoffman, D., Skadsen, R., Steffenson, B.J., 867 868 1993. A molecular, isozyme and morphological map of the barley (Hordeum vulgare) genome. 869 Theor. Appl. Genet. 86, 705–712. https://doi.org/10.1007/BF00222660 870 Knüpffer, H., 2009. Triticeae Genetic Resources in ex situ Genebank Collections, in: Muehlbauer, G.J., Feuillet, C. (Eds.), Genetics and Genomics of the Triticeae. Springer US, New York, NY, pp. 31-871 872 79. https://doi.org/10.1007/978-0-387-77489-3_2 873 Knüpffer, H., van Hintum, T.J.L., 2003. Summarised diversity - the Barley Core Collection, in: von 874 Bothmer, R., van Hintum, T., Knüpffer, H., Sato, K. (Eds.), Diversity in Barley (Hordeum Vulgare). 875 Elsevier Science, pp. 259–267. 876 Konieczny, A., Ausubel, F.M., 1993. A procedure for mapping Arabidopsis mutations using codominant ecotype-specific PCR-based markers. Plant J. 4, 403–410. 877 878 https://doi.org/10.1046/j.1365-313X.1993.04020403.x 879 Konishi, T., Ban, T., Iida, Y., Yoshimi, R., 1997. Genetic analysis of disease resistance to all strains of BaYMV in a Chinese barley landrace, Mokusekko 3. TAG Theor. Appl. Genet. 94, 871–877. 880 881 https://doi.org/10.1007/s001220050489 882 Krasileva, K. V., Vasquez-Gross, H.A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., 883 Ramirez-Gonzalez, R.H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A.L., Uauy, C., 884 Dubcovsky, J., 2017. Uncovering hidden variation in polyploid wheat. Proc. Natl. Acad. Sci. 114,

885 E913-E921. https://doi.org/10.1073/pnas.1619268114 886 Kumar, N., Galli, M., Ordon, J., Stuttmann, J., Kogel, K., Imani, J., 2018. Further analysis of barley MORC1 using a highly efficient RNA-guided Cas9 gene-editing system. Plant Biotechnol J, 16: 887 888 1892-1903. doi:10.1111/pbi.12924 Kurowska, M., Labocha-Pawłowska, A., Gnizda, D., Maluszynski, M., Szarejko, I., 2012. Molecular 889 890 analysis of point mutations in a barley genome exposed to MNU and gamma rays. Mutat. Res. 891 Mol. Mech. Mutagen. 738-739, 52-70. https://doi.org/10.1016/j.mrfmmm.2012.08.008 Lababidi, S., Mejlhede, N., Rasmussen, S.K., Backes, G., Al-Said, W., Baum, M., Jahoor, A., 2009. 892 893 Identification of barley mutants in the cultivar 'Lux' at the Dhn loci through TILLING. Plant 894 Breed. 128, 332–336. https://doi.org/10.1111/j.1439-0523.2009.01640.x 895 Lander, E.S., Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP 896 linkage maps. Genetics 121, 185–99. https://doi.org/10.1038/hdy.2014.4 897 Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., Uauy, C., Harwood, W., 898 2015. Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-899 guided Cas9 nuclease. Genome Biol. 16, 258. https://doi.org/10.1186/s13059-015-0826-7 900 Li, M., Liu, X., Bradbury, P., Yu, J., Zhang, Y.-M., Todhunter, R.J., Buckler, E.S., Zhang, Z., 2014. 901 Enrichment of statistical power for genome-wide association studies. BMC Biol. 12, 73. 902 https://doi.org/10.1186/s12915-014-0073-5 903 Linde-Laursen, I., Heslop-Harrison, J.S., Shepherd, K.W., Taketa, S., 1997. The barley Genome and its 904 Relationship with the Wheat Genomes. A Survey with an Internationally Agreed 905 Recommendation for Barley Chromosome Nomenclature. Hereditas 126, 1–16. 906 https://doi.org/10.1111/j.1601-5223.1997.00001.x 907 Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., Ecker, J.R., 2008. 908 Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis. Cell 133, 523-909 536. https://doi.org/10.1016/j.cell.2008.03.029 910 Litt, M., Luty, J.A., 1989. A hypervariable microsatellite revealed by in vitro amplification of a 911 dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. 44, 397–401. 912 Lu, F., Romay, M. C., Glaubitz, J. C., Bradbury, P. J., Elshire, R. J., Wang, T., Li, Y., Li, Y., Semagn, K., 913 Zhang, X., Hernandez, A. G., Mikel, M. A., Soifer, I., Barad, O., Buckler, E. S., 2015. High-914 resolution genetic mapping of maize pan-genome sequence anchors. Nature Communications, 915 6, 6914. https://doi.org/10.1038/ncomms7914 916 Lüpken, T., Stein, N., Perovic, D., Habekuß, A., Krämer, I., Hähnel, U., Steuernagel, B., Scholz, U., 917 Zhou, R., Ariyadasa, R., Taudien, S., Platzer, M., Martis, M., Mayer, K., Friedt, W., Ordon, F., 918 2013. Genomics-based high-resolution mapping of the BaMMV/BaYMV resistance gene rym11 919 in barley (Hordeum vulgare L.). Theor. Appl. Genet. 126, 1201–1212. 920 https://doi.org/10.1007/s00122-013-2047-3 Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radchuk, V., 921 922 Dockter, C., Hedley, P.E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X.-Q., 923 Zhang, Q., Barrero, R.A., Li, L., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., 924 Staňková, H., Vrána, J., Chan, S., Muñoz-Amatriaín, M., Ounit, R., Wanamaker, S., Bolser, D., 925 Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Chailyan, A., Sampath, 926 D., Heavens, D., Clissold, L., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, 927 J.K., Tan, C., Wang, P., Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard, M.I., Borisjuk, L., 928 Houben, A., Doležel, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P., Muehlbauer, G.J., Clark, 929 M.D., Caccamo, M., Schulman, A.H., Mayer, K.F.X., Platzer, M., Close, T.J., Scholz, U., Hansson, 930 M., Zhang, G., Braumann, I., Spannagl, M., Li, C., Waugh, R., Stein, N., 2017. A chromosome 931 conformation capture ordered sequence of the barley genome. Nature 544, 427–433. 932 https://doi.org/10.1038/nature22043 933 Mascher, M., Jost, M., Kuon, J.-E., Himmelbach, A., Aßfalg, A., Beier, S., Scholz, U., Graner, A., Stein, N., 2014. Mapping-by-sequencing accelerates forward genetics in barley. Genome Biol. 15, R78. 934 935 https://doi.org/10.1186/gb-2014-15-6-r78

936 Mascher, M., Richmond, T.A., Gerhardt, D.J., Himmelbach, A., Clissold, L., Sampath, D., Ayling, S.,

937 Steuernagel, B., Pfeifer, M., D'Ascenzo, M., Akhunov, E.D., Hedley, P.E., Gonzales, A.M., Morrell, 938 P.L., Kilian, B., Blattner, F.R., Scholz, U., Mayer, K.F.X., Flavell, A.J., Muehlbauer, G.J., Waugh, R., 939 Jeddeloh, J.A., Stein, N., 2013. Barley whole exome capture: a tool for genomic research in the 940 genus Hordeum and beyond. Plant J. 76, 494–505. https://doi.org/10.1111/tpj.12294 941 Maurer, A., Draba, V., Jiang, Y., Schnaithmann, F., Sharma, R., Schumann, E., Kilian, B., Reif, J.C., 942 Pillen, K., 2015. Modelling the genetic architecture of flowering time control in barley through 943 nested association mapping. BMC Genomics 16, 290. https://doi.org/10.1186/s12864-015-944 1459-7 Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., 945 946 Roessner, S., Gundlach, H., Kubaláková, M., Suchánková, P., Murat, F., Felder, M., Nussbaumer, 947 T., Graner, A., Salse, J., Endo, T., Sakai, H., Tanaka, T., Itoh, T., Sato, K., Platzer, M., Matsumoto, 948 T., Scholz, U., Doležel, J., Waugh, R., Stein, N., 2011. Unlocking the Barley Genome by 949 Chromosomal and Comparative Genomics. Plant Cell 23, 1249–1263. 950 https://doi.org/10.1105/tpc.110.082537 951 McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S., 2000. Targeting Induced LocalLesions IN 952 Genomes (TILLING) for Plant Functional Genomics. Plant Physiol. 123, 439-442. 953 https://doi.org/10.1104/pp.123.2.439 954 Michalek, W., Künzel, G., Graner, a, 1999. Sequence analysis and gene identification in a set of mapped RFLP markers in barley (Hordeum vulgare). Genome 42, 849-53. 955 956 Michelmore, R.W., Paran, I., Kesseli, R. V, 1991. Identification of markers linked to disease-resistance 957 genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. 88, 9828–9832. 958 959 https://doi.org/10.1073/pnas.88.21.9828 960 Milner, S., Jost, M., Taketa, S., Mazón, E., Himmelbach, A Oppermann, M., Weise, S., Knüpffer, H., 961 Basterrechea, M., König, P., Schüler, D., Sharma, R., Pasam, R., Rutten, T., Guo, G., Xu, D., Zhang, 962 J., Herren, G., Müller, T., Krattinger, S., Keller, B., Jiang, Y., Gonzalez, M., Zhao, Y., Habekuß, A., 963 Färber, S., Ordon, F., Lange, M., Börner, A., Graner, A., Reif, J., Scholz, U., Mascher, M., Stein, N., 964 2018. Genebank genomics highlights the diversity of a global barley collection. Nature genetics. 965 https://doi.org/10.1038/s41588-018-0266-x 966 Monat, C., Schreiber, M., Stein, N., Mascher, M., 2018. Prospects of pan-genomics in barley. 967 Theoretical and Applied Genetics, 1-12. 968 Muñoz-Amatriaín, M., Cuesta-Marcos, A., Endelman, J.B., Comadran, J., Bonman, J.M., Bockelman, 969 H.E., Chao, S., Russell, J., Waugh, R., Hayes, P.M., Muehlbauer, G.J., 2014. The USDA Barley Core 970 Collection: Genetic Diversity, Population Structure, and Potential for Genome-Wide Association 971 Studies. PLoS One 9, e94688. https://doi.org/10.1371/journal.pone.0094688 972 Muñoz-Amatriaín, M., Moscou, M.J., Bhat, P.R., Svensson, J.T., Bartoš, J., Suchánková, P., Šimková, H., 973 Endo, T.R., Fenton, R.D., Lonardi, S., Castillo, A.M., Chao, S., Cistué, L., Cuesta-Marcos, A., 974 Forrest, K.L., Hayden, M.J., Hayes, P.M., Horsley, R.D., Makoto, K., Moody, D., Sato, K., Vallés, 975 M.P., Wulff, B.B.H., Muehlbauer, G.J., Doležel, J., Close, T.J., 2011. An Improved Consensus 976 Linkage Map of Barley Based on Flow-Sorted Chromosomes and Single Nucleotide 977 Polymorphism Markers. Plant Genome J. 4, 238. 978 https://doi.org/10.3835/plantgenome2011.08.0023 979 Nice, L.M., Steffenson, B.J., Brown-Guedira, G.L., Akhunov, E.D., Liu, C., Kono, T.J.Y., Morrell, P.L., 980 Blake, T.K., Horsley, R.D., Smith, K.P., Muehlbauer, G.J., 2016. Development and Genetic 981 Characterization of an Advanced Backcross-Nested Association Mapping (AB-NAM) Population 982 of Wild x Cultivated Barley. Genetics 203, 1453–1467. 983 https://doi.org/10.1534/genetics.116.190736 Nissan-Azzouz, F., Graner, A., Friedt, W., Ordon, F., 2005. Fine-mapping of the BaMMV, BaYMV-1 and 984 985 BaYMV-2 resistance of barley (Hordeum vulgare) accession PI1963. Theor. Appl. Genet. 110, 986 212-218. https://doi.org/10.1007/s00122-004-1802-x 987 Olson, M., Hood, L., Cantor, C., Botstein, D., 1989. A common language for physical mapping of the 988 human genome. Science (80-.). 245, 1434–1435. https://doi.org/10.1126/science.2781285

989 Ordon, F., Bauer, E., Friedt, Graner, A., 1995. Marker-based selection for the ym4 BaMMV-resistance 990 gene in barley using RAPDs. Agronomie 15, 481–485. https://doi.org/10.1051/agro:19950717 991 Pankin, A., Campoli, C., Dong, X., Kilian, B., Sharma, R., Himmelbach, A., Saini, R., Davis, S.J., Stein, N., 992 Schneeberger, K., von Korff, M., 2014. Mapping-by-Sequencing Identifies HvPHYTOCHROME C 993 as a Candidate Gene for the early maturity 5 Locus Modulating the Circadian Clock and 994 Photoperiodic Flowering in Barley. Genetics 198, 383–396. 995 https://doi.org/10.1534/genetics.114.165613 996 Pasam, R.K., Sharma, R., Walther, A., Özkan, H., Graner, A., Kilian, B., 2014. Genetic Diversity and 997 Population Structure in a Legacy Collection of Spring Barley Landraces Adapted to a Wide Range 998 of Climates. PLoS One 9, e116164. https://doi.org/10.1371/journal.pone.0116164 999 Pellio, B., Streng, S., Bauer, E., Stein, N., Perovic, D., Schiemann, A., Friedt, W., Ordon, F., Graner, A., 1000 2005. High-resolution mapping of the Rym4/Rym5 locus conferring resistance to the barley 1001 yellow mosaic virus complex (BaMMV, BaYMV, BaYMV-2) in barley (Hordeum vulgare ssp. 1002 vulgare L.). Theor. Appl. Genet. 110, 283–293. https://doi.org/10.1007/s00122-004-1832-4 1003 Poland, J.A., Brown, P.J., Sorrells, M.E., Jannink, J.-L., 2012. Development of High-Density Genetic 1004 Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. 1005 PLoS One 7, e32253. https://doi.org/10.1371/journal.pone.0032253 1006 Potokina, E., Druka, A., Luo, Z., Wise, R., Waugh, R., Kearsey, M., 2008. Gene expression quantitative 1007 trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide 1008 transcriptional regulation. Plant J. 53, 90-101. https://doi.org/10.1111/j.1365-1009 313X.2007.03315.x Putnam, N.H., O'Connell, B.L., Stites, J.C., Rice, B.J., Blanchette, M., Calef, R., Troll, C.J., Fields, A., 1010 1011 Hartley, P.D., Sugnet, C.W., Haussler, D., Rokhsar, D.S., Green, R.E., 2016. Chromosome-scale 1012 shotgun assembly using an in vitro method for long-range linkage. Genome Res. 26, 342–50. 1013 https://doi.org/10.1101/gr.193474.115 1014 Rifkin, S.A., 2012. Quantitative Trait Loci (QTL), Methods in Molecular Biology. Humana Press, 1015 Totowa, NJ. https://doi.org/10.1007/978-1-61779-785-9 1016 Saghai Maroof, M.A., Biyashev, R.M., Yang, G.P., Zhang, Q., Allard, R.W., 1994. Extraordinarily 1017 polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and 1018 population dynamics. Proc. Natl. Acad. Sci. 91, 5466-5470. 1019 https://doi.org/10.1073/pnas.91.12.5466 1020 Saiki, R., Scharf, S., Faloona, F., Mullis, K., Horn, G., Erlich, H., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of 1021 1022 sickle cell anemia. Science (80-.). 230, 1350–1354. https://doi.org/10.1126/science.2999980 1023 Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., Vrána, J., 1024 Kubaláková, M., Krattinger, S.G., Wicker, T., Doležel, J., Keller, B., Wulff, B.B.H., 2016. Rapid 1025 gene isolation in barley and wheat by mutant chromosome sequencing. Genome Biol. 17, 221. 1026 https://doi.org/10.1186/s13059-016-1082-1 Sannemann, W., Huang, B.E., Mathew, B., Léon, J., 2015. Multi-parent advanced generation inter-1027 1028 cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of 1029 concept. Mol. Breed. 35, 86. https://doi.org/10.1007/s11032-015-0284-7 1030 Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A.H., Nielsen, K.L., Jørgensen, J.-E., Weigel, 1031 D., Andersen, S.U., 2009. SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat. Methods 6, 550–551. https://doi.org/10.1038/nmeth0809-550 1032 1033 Schulte, D., Ariyadasa, R., Shi, B., Fleury, D., Saski, C., Atkins, M., DeJong, P., Wu, C.-C., Graner, A., 1034 Langridge, P., Stein, N., 2011. BAC library resources for map-based cloning and physical map 1035 construction in barley (Hordeum vulgare L.). BMC Genomics 12, 247. 1036 https://doi.org/10.1186/1471-2164-12-247 1037 Semagn, K., Babu, R., Hearne, S., Olsen, M., 2014. Single nucleotide polymorphism genotyping using 1038 Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop 1039 improvement. Molecular Breeding, 33(1), 1–14. doi:10.1007/s11032-013-9917-x 1040 Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., Nordborg, M., 2012. An efficient

1041 multi-locus mixed-model approach for genome-wide association studies in structured 1042 populations. Nat. Genet. 44, 825-830. https://doi.org/10.1038/ng.2314 1043 Spannagl, M., Nussbaumer, T., Bader, K.C., Martis, M.M., Seidel, M., Kugler, K.G., Gundlach, H., 1044 Mayer, K.F.X., 2016. PGSB PlantsDB: updates to the database framework for comparative plant 1045 genome research. Nucleic Acids Res. 44, D1141–D1147. https://doi.org/10.1093/nar/gkv1130 Spielman, R.S., McGinnis, R.E., Ewens, W.J., 1993. Transmission test for linkage disequilibrium: the 1046 1047 insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. 52, 1048 506. Springer, N. M., Anderson, S. N., Andorf, C. M., Ahern, K. R., Bai, F., Barad, O., Barbazuk, W B., Bass, H. 1049 1050 W., Baruch, K., Ben-Zvi, G., Buckler, E. S., Bukowski, R., Campbell, M. S., Cannon, E. K. S., 1051 Chomet, P., Dawe, R. K., Davenport, R., Dooner, H. K., Du, L. H., Du, C., Easterling, K. A., Gault, 1052 C., Guan, J. -C., Hunter, C. T., Jander, G., Jiao, Y., Koch, K. E., Kol, G., Köllner, T. G., Kudo, T., Li, 1053 Q., Lu, F., Mayfield-Jones, D., Mei, W., McCarty, D. R., Noshay, J. M., Portwood, J. L., Ronen, G., Settles, A. M., Shem-Tov, D., Shi, J., Soifer, I., Stein, J. C., Stitzer, M. C., Suzuki, M., Vera, D. L., 1054 1055 Vollbrecht, E., Vrebalov, J. T., Ware, D., Wei, S., Wimalanathan, K., Woodhouse, M. R., Xiong, 1056 W., Brutnell, T. P., 2018. The maize W22 genome provides a foundation for functional genomics 1057 and transposon biology. Nature Genetics, 50(9), 1282–1288. doi:10.1038/s41588-018-0158-0 1058 Steffenson, B.J., Olivera, P., Roy, J.K., Jin, Y., Smith, K.P., Muehlbauer, G.J., 2007. A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. Aust. J. Agric. Res. 58, 1059 1060 532. https://doi.org/10.1071/AR07123 1061 Stein, N., Muehlbauer, G.J., 2018. The Barley Genome, Compendium of Plant Genomes. Springer 1062 International Publishing. https://doi.org/10.1007/978-3-319-92528-8 1063 Stein, N., Perovic, D., Kumlehn, J., Pellio, B., Stracke, S., Streng, S., Ordon, F., Graner, A., 2005. The 1064 eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in Hordeum vulgare (L.). Plant J. 42, 912–922. https://doi.org/10.1111/j.1365-313X.2005.02424.x 1065 1066 Steuernagel, B., Periyannan, S.K., Hernández-Pinzón, I., Witek, K., Rouse, M.N., Yu, G., Hatta, A., 1067 Ayliffe, M., Bariana, H., Jones, J.D.G., Lagudah, E.S., Wulff, B.B.H., 2016. Rapid cloning of 1068 disease-resistance genes in plants using mutagenesis and sequence capture. Nat. Biotechnol. 1069 34, 652-655. https://doi.org/10.1038/nbt.3543 1070 Sun, Y., Wang, J., Crouch, J.H., Xu, Y., 2010. Efficiency of selective genotyping for genetic analysis of 1071 complex traits and potential applications in crop improvement. Mol. Breed. 26, 493–511. 1072 https://doi.org/10.1007/s11032-010-9390-8 Sun, S., Zhou, Y., Chen, J., Shi, J., Zhao, H., Zhao, H., Song, W., Zhang, M., Cui, Y., Dong, X., Liu, H., Ma, 1073 1074 X., Jiao, Y., Wang, B., Wei, X., Stein, J. C., Glaubitz, J. C., Lu, F., Yu, G., Liang, C., Fengler, K., Li, B., 1075 Rafalski, A., Schnable, P. S., Ware, D. H., Buckler, E. S., Lai, J. (2018). Extensive intraspecific gene 1076 order and gene structural variations between Mo17 and other maize genomes. Nature 1077 Genetics, 50(9), 1289–1295. doi:10.1038/s41588-018-0182-0 1078 Szarejko, I., Szurman-Zubrzycka, M., Nawrot, M., Marzec, M., Gruszka, D., Kurowska, M., Chmielewska, B., Zbieszczyk, J., Jelonek, J., Maluszynski, M., 2017. Creation of a TILLING 1079 1080 Population in Barley After Chemical Mutagenesis with Sodium Azide and MNU, in: Jankowicz-1081 Cieslak, J., Tai, T.H., Kumlehn, J., Till, B.J. (Eds.), Biotechnologies for Plant Mutation Breeding. Springer International Publishing, Cham, pp. 91–111. https://doi.org/10.1007/978-3-319-45021-1082 1083 66 1084 Szurman-Zubrzycka, M.E., Zbieszczyk, J., Marzec, M., Jelonek, J., Chmielewska, B., Kurowska, M.M., 1085 Krok, M., Daszkowska-Golec, A., Guzy-Wrobelska, J., Gruszka, D., Gajecka, M., Gajewska, P., 1086 Stolarek, M., Tylec, P., Sega, P., Lip, S., Kudełko, M., Lorek, M., Gorniak-Walas, M., Malolepszy, A., Podsiadlo, N., Szyrajew, K.P., Keisa, A., Mbambo, Z., Todorowska, E., Gaj, M., Nita, Z., 1087 Orlowska-Job, W., Maluszynski, M., Szarejko, I., 2018. HorTILLUS—A Rich and Renewable Source 1088 1089 of Induced Mutations for Forward/Reverse Genetics and Pre-breeding Programs in Barley 1090 (Hordeum vulgare L.). Front. Plant Sci. 9, 1–16. https://doi.org/10.3389/fpls.2018.00216 1091 Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A., Mitsuoka, C., Utsushi, H., Natsume, S., 1092 Kanzaki, H., Matsumura, H., Saitoh, H., Yoshida, K., Cano, L.M., Kamoun, S., Terauchi, R., 2013.

Page **25** of **31** Achieving sustainable cultivation of barley (ed. Dr Glen Fox and Prof Chengdao Li)

1093 MutMap-Gap: whole-genome resequencing of mutant F2 progeny bulk combined with de novo 1094 assembly of gap regions identifies the rice blast resistance gene Pii. New Phytol. 200, 276–283. 1095 https://doi.org/10.1111/nph.12369 1096 Talamè, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U., Salvi, S., 2008. TILLMore, a 1097 resource for the discovery of chemically induced mutants in barley. Plant Biotechnol. J. 6, 477– 1098 485. https://doi.org/10.1111/j.1467-7652.2008.00341.x 1099 Tanksley, S.D., Young, N.D., Paterson, A.H., Bonierbale, M.W., 1989. RFLP Mapping in Plant Breeding: 1100 New Tools for an Old Science. Nat. Biotechnol. 7, 257–264. https://doi.org/10.1038/nbt0389-1101 257 1102 Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., Bennett, M., 2017. Plant Phenomics, From Sensors to 1103 Knowledge. Curr. Biol. 27, R770–R783. https://doi.org/10.1016/j.cub.2017.05.055 1104 Tello-Ruiz, M.K., Stein, J., Wei, S., Preece, J., Olson, A., Naithani, S., Amarasinghe, V., 1105 Dharmawardhana, P., Jiao, Y., Mulvaney, J., Kumari, S., Chougule, K., Elser, J., Wang, B., 1106 Thomason, J., Bolser, D.M., Kerhornou, A., Walts, B., Fonseca, N.A., Huerta, L., Keays, M., Tang, 1107 Y.A., Parkinson, H., Fabregat, A., McKay, S., Weiser, J., D'Eustachio, P., Stein, L., Petryszak, R., 1108 Kersey, P.J., Jaiswal, P., Ware, D., 2016. Gramene 2016: comparative plant genomics and 1109 pathway resources. Nucleic Acids Res. 44, D1133–D1140. https://doi.org/10.1093/nar/gkv1179 1110 The International Wheat Genome Sequencing Consortium, 2014. Ancient hybridizations among the ancestral genomes of bread wheat. Science (80-.). 345, 1250092–1250092. 1111 https://doi.org/10.1126/science.1250092 1112 1113 Thind, A.K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J., 1114 Krattinger, S.G., 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-1115 range chromosome assembly. Nat. Biotechnol. 35, 793–796. https://doi.org/10.1038/nbt.3877 Thoday, J.M., 1961. Location of Polygenes. Nature 191, 368–370. https://doi.org/10.1038/191368a0 1116 1117 Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., 1118 Young, K., Taylor, N.E., Henikoff, J.G., Comai, L., Henikoff, S., 2003. Large-scale discovery of 1119 induced point mutations with high-throughput TILLING. Genome Res. 13, 524-30. 1120 https://doi.org/10.1101/gr.977903 1121 Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., Khan, A.A., Filkov, V., Tai, T.H., Dubcovsky, J., Comai, L., 2011. Discovery of Rare 1122 1123 Mutations in Populations: TILLING by Sequencing. PLANT Physiol. 156, 1257–1268. 1124 https://doi.org/10.1104/pp.110.169748 Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. Van De, Hornes, M., Friters, A., Pot, J., Paleman, J., 1125 1126 Kuiper, M., Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 1127 23, 4407-4414. https://doi.org/10.1093/nar/23.21.4407 Watson, A., Ghosh, S., Williams, M.J., Cuddy, W.S., Simmonds, J., Rey, M.-D., Asyraf Md Hatta, M., 1128 1129 Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N.M., Breakspear, A., Korolev, A., Rayner, T., 1130 Dixon, L.E., Riaz, A., Martin, W., Ryan, M., Edwards, D., Batley, J., Raman, H., Carter, J., Rogers, C., Domoney, C., Moore, G., Harwood, W., Nicholson, P., Dieters, M.J., DeLacy, I.H., Zhou, J., 1131 1132 Uauy, C., Boden, S.A., Park, R.F., Wulff, B.B.H., Hickey, L.T., 2018. Speed breeding is a powerful 1133 tool to accelerate crop research and breeding. Nat. Plants 4, 23–29. 1134 https://doi.org/10.1038/s41477-017-0083-8 1135 Wendt, T., Holm, P.B., Starker, C.G., Christian, M., Voytas, D.F., Brinch-Pedersen, H., Holme, I.B., 2013. TAL effector nucleases induce mutations at a pre-selected location in the genome of 1136 primary barley transformants. Plant Mol. Biol. 83, 279–285. https://doi.org/10.1007/s11103-1137 1138 013-0078-4 Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A., Kilian, A., 2004. Diversity 1139 1140 Arrays Technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. 101, 1141 9915–9920. https://doi.org/10.1073/pnas.0401076101 1142 Wenzl, P., Li, H., Carling, J., Zhou, M., Raman, H., Paul, E., Hearnden, P., Maier, C., Xia, L., Caig, V., 1143 Ovesná, J., Cakir, M., Poulsen, D., Wang, J., Raman, R., Smith, K.P., Muehlbauer, G.J., Chalmers, 1144 K.J., Kleinhofs, A., Huttner, E., Kilian, A., 2006. A high-density consensus map of barley linking

1145 DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics 7, 206. 1146 https://doi.org/10.1186/1471-2164-7-206 Wicker, T., Schulman, A.H., Tanskanen, J., Spannagl, M., Twardziok, S., Mascher, M., Springer, N.M., 1147 1148 Li, Q., Waugh, R., Li, C., Zhang, G., Stein, N., Mayer, K.F.X., Gundlach, H., 2017. The repetitive landscape of the 5100 Mbp barley genome. Mob. DNA 8, 1–16. 1149 https://doi.org/10.1186/s13100-017-0102-3 1150 Wicker, T., Zimmermann, W., Perovic, D., Paterson, A.H., Ganal, M., Graner, A., Stein, N., 2005. A 1151 detailed look at 7 million years of genome evolution in a 439 kb contiguous sequence at the 1152 1153 barley Hv-eIF4E locus: recombination, rearrangements and repeats. Plant J. 41, 184–94. 1154 https://doi.org/10.1111/j.1365-313X.2004.02285.x 1155 Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S. V., 1990. DNA polymorphisms 1156 amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535. 1157 https://doi.org/10.1093/nar/18.22.6531 Wise, R.P., Caldo, R.A., Hong, L., Shen, L., Cannon, E., Dickerson, J.A., 2007. BarleyBase/PLEXdb, in: 1158 1159 Edwards, D. (Ed.), Plant Bioinformatics. Humana Press, Totowa, NJ, pp. 347–363. https://doi.org/10.1007/978-1-59745-535-0_17 1160 1161 Worley, K.C., Richards, S., Rogers, J., 2017. The value of new genome references. Exp. Cell Res. 358, 1162 433-438. https://doi.org/10.1016/j.yexcr.2016.12.014 Yang, N., Xu, X.-W., Wang, R.-R., Peng, W.-L., Cai, L., Song, J.-M., Li, W., Luo, X., Niu, L., Wang, Y., Jin, 1163 M., Chen, L., Luo, J., Deng, M., Wang, L., Pan, Q., Liu, F., Jackson, D., Yang, X., Chen, L.-L., Yan, J. 1164 1165 (2017). Contributions of Zea mays subspecies mexicana haplotypes to modern maize. Nature 1166 Communications, 8(1), 1874. doi:10.1038/s41467-017-02063-5 1167 Yang, P., Habekuß, A., Hofinger, B.J., Kanyuka, K., Kilian, B., Graner, A., Ordon, F., Stein, N., 2017. 1168 Sequence diversification in recessive alleles of two host factor genes suggests adaptive selection for bymovirus resistance in cultivated barley from East Asia. Theor. Appl. Genet. 130, 1169 1170 331-344. https://doi.org/10.1007/s00122-016-2814-z 1171 Yang, P., Habekuß, A., Ordon, F., Stein, N., 2014a. Analysis of bymovirus resistance genes on proximal barley chromosome 4HL provides the basis for precision breeding for BaMMV/BaYMV 1172 1173 resistance. Theor. Appl. Genet. 127, 1625-1634. https://doi.org/10.1007/s00122-014-2324-9 1174 Yang, P., Lüpken, T., Habekuss, A., Hensel, G., Steuernagel, B., Kilian, B., Ariyadasa, R., Himmelbach, A., Kumlehn, J., Scholz, U., Ordon, F., Stein, N., 2014b. PROTEIN DISULFIDE ISOMERASE LIKE 5-1 1175 1176 is a susceptibility factor to plant viruses. Proc. Natl. Acad. Sci. U. S. A. 111, 2104–9. https://doi.org/10.1073/pnas.1320362111 1177 1178 Yin, Z., Lan, H., Tan, G., Lu, M., Vasilakos, A. V., Liu, W., 2017. Computing Platforms for Big Biological 1179 Data Analytics: Perspectives and Challenges. Comput. Struct. Biotechnol. J. 15, 403–411. 1180 https://doi.org/10.1016/j.csbj.2017.07.004 Yu, J., Holland, J.B., McMullen, M.D., Buckler, E.S., 2008. Genetic design and statistical power of 1181 1182 nested association mapping in maize. Genetics 178, 539–51. 1183 https://doi.org/10.1534/genetics.107.074245 1184 Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., 1185 Nielsen, D., Holland, J.B., Kresovich, S., Buckler, E.S., 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38, 203-8. 1186 1187 https://doi.org/10.1038/ng1702 Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Brueggeman, R.S., Muehlbauer, 1188 G.J., Wise, R.P., Wing, R.A., 2000. A bacterial artificial chromosome library for barley (Hordeum 1189 1190 vulgare L.) and the identification of clones containing putative resistance genes. TAG Theor. 1191 Appl. Genet. 101, 1093-1099. https://doi.org/10.1007/s001220051584 Zeng, Z.-B., 1993. Theoretical basis for separation of multiple linked gene effects in mapping 1192 1193 quantitative trait loci. Proc. Natl. Acad. Sci. 90, 10972–10976. 1194 Zhang, H., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., 1195 Schweizer, P., Zierold, U., Langridge, P., Varshney, R.K., Wobus, U., Graner, A., 2004. Large-scale 1196 analysis of the barley transcriptome based on expressed sequence tags. Plant J. 40, 276–290.

- 1197 https://doi.org/10.1111/j.1365-313X.2004.02209.x
- Zhang, Z., Ersoz, E.S., Lai, C.-Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett,
 D.K., Ordovas, J.M., Buckler, E.S., 2010. Mixed linear model approach adapted for genome-wide
 association studies. Nat. Genet. 42, 355–60. https://doi.org/10.1038/ng.546
- 1201 Zhu, C., Gore, M., Buckler, E.S., Yu, J., 2008. Status and Prospects of Association Mapping in Plants.
- 1202 Plant Genome J. 1, 5. https://doi.org/10.3835/plantgenome2008.02.0089
- 1203
- 1204







Marker	RFLP	AFLP	RAPD	SSR	DArT	SNP array	GBS
Co-dominant/Dominant	Co-dominant	Dominant	Dominant	Co-dominant	Dominant	Co-dominant	Co-dominant
Genome abundance	High	Very high	Very high	Medium	Very high	Very high	Very high
Reproducibility	High	High	Low	High	High	High	High
Number of marker scored per assay	Low	Moderate	Low	Low	Moderate	High	High
Technical requirement	High	Moderate	Low	Moderate	High	Low	Moderate
Cost per data point	High	Moderate	Low	Low	Very low	Low	Very low
Automation	Difficult	Moderate	Difficult	Moderate	Easy	Easy	Easy
Prior sequence knowledge	No	No	No	Yes	No	Yes	Preferable
Marker development	Fast	Moderate	Long	Long	Long	Long	Moderate
Ease of analysis	Difficult	Difficult	Difficult	Easy	Easy	Easy	Easy
Quantity of DNA required (µg)	10000	500 to 1000	20	50	50 to 100	200	20
Equipment cost	Moderate	Moderate	Low	Low	High	High	High
Throughput	Low	Moderate	Moderate	Moderate	High	High	High

Table 1: Characteristics of the most used markers in barley, currently or in the past.