Genetic Dissection of Barley Morphology and Development[1][W][OA]

Arnis Druka, Jerome Franckowiak, Udda Lundqvist, Nicola Bonar, Jill Alexander, Kelly Houston, Slobodanka Radovic, Fahimeh Shahinnia, Vera Vendramin, Michele Morgante, Nils Stein, and Robbie Waugh*

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, United Kingdom (A.D., N.B., J.A., K.H., R.W.); Department of Employment, Economic Development, and Innovation, Agric-Science Queensland, Hermitage Research Station, Warwick, Queensland 4370, Australia (J.F.); Nordic Genetic Resource Center, Smedjevägen 3, SE–230 53, Alnarp, Sweden (U.L.); The Applied Genomics Institute, University of Udine, Udine 33100, Italy (S.R., V.V., M.M.); and Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany (F.S., N.S.)

Since the early 20th century, barley (*Hordeum vulgare*) has been a model for investigating the effects of physical and chemical mutagens and for exploring the potential of mutation breeding in crop improvement. As a consequence, extensive and well-characterized collections of morphological and developmental mutants have been assembled that represent a valuable resource for exploring a wide range of complex and fundamental biological processes. We constructed a collection of 881 backcrossed lines containing mutant alleles that induce a majority of the morphological and developmental variation described in this species. After genotyping these lines with up to 3,072 single nucleotide polymorphisms, comparison to their recurrent parent defined the genetic location of 426 mutant alleles to chromosomal segments, each representing on average <3% of the barley genetic map. We show how the gene content in these segments can be predicted through conservation of synteny with model cereal genomes, providing a route to rapid gene identification.

In 1928, the eminent American geneticist and plant breeder L.J. Stadler published a manuscript in *Science* demonstrating that ionizing radiation could increase the mutation frequency in barley (*Hordeum vulgare*) and that the induced mutations were transmitted to subsequent generations (Stadler, 1928). His observations ignited the field of plant mutation research that, over the last 80 years, has both explored the potential of mutation breeding in crop improvement (Harten, 1998; Ahloowalia, 2004; Lundqvist, 2009) and generated considerable basic understanding about fundamental processes of plant morphology, physiology, and development. The barley research community actively exploited the simple diploid genetics of the crop, particularly during the heyday of mutation research in the 1950s to 1970s, ultimately incorporating characterized mutant lines into collections that grew to contain thousands of accessions. Of particular note was the Scandanavian mutation research program established by the Swedish geneticists H. Nilsson-Ehle and A. Gustafsson that provided a legacy of over 10,000 different characterized mutants that remain today expertly stored in the NordGen genebank (http://www.nordgen.org/).

Barley mutants have been used to isolate or validate genes such as nitrate reductase (Somers et al., 1983), key genes in the anthocyanin pathway (von Wettstein, 2007), a gene responsible for the floral bract phenotype *HOODED* (Muller et al., 1995), the row-type gene *SIX-ROWED SPIKE1* (*VRS1*; Komatsuda et al., 2007), the hull adhesion gene *NAKED CARYOPSIS* (*NUD*; Taketa et al., 2008), plant height genes *UZU DWARF* (*UZU*; Chono et al., 2003) and *SLENDER1* (Chandler et al., 2002), several endosperm development genes (Felker et al., 1983; Morell et al., 2003; Röder et al., 2006; Clarke et al., 2008), and two disease lesion mimic mutations *NECROTIC1* (Rostoks et al., 2006) and *NECROTIC.S1* (Zhang et al., 2009) among others. However, despite these individual achievements, with few exceptions (Pozzi et al., 2003; Rossini et al., 2006), the barley mutant resources have not yet been systematically explored using the tools of modern genetics. Unlike Arabidopsis

[1] This work was supported by the European Research Area Networks in Plant Genomics (project no. ERAPGFP/06.046A; Genomics-Assisted Dissection of Barley Morphology and Development awarded to R.W., N.S., and M.M. via their respective funding agencies). Development of the Bowman Near Isogenic Lines was possible through long-term funding to J.F. from the American Malting Barley Association, Inc.

* Corresponding author; e-mail robbie.waugh@scri.ac.uk.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Arnis Druka (arnis.druka@scri.ac.uk).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.110.166249

Figure 1. (*Legend appears on following page.*)

(*Arabidopsis thaliana*), the large size of the barley genome and lack of a genome sequence have precluded community-wide positional cloning efforts.

As primary barley mutants were induced or discovered in a range of accessions, which after mutagenesis also contained a different spectrum of background mutations, direct comparisons between lines are compromised, especially for subtle, environmentally sensitive phenotypes. To overcome these issues, repeated backcross-based transfer of a mutant locus into the genetic background of a common recurrent parent (with phenotypic selection at each cycle) has been applied. The outcome of this effort for each mutant is a pair of Near Isogenic Lines (NILs) that are characterized by having a relatively small genetic interval originating from the donor that contains the mutated locus, embedded in the genome of the recurrent parent. The theory of backcrossing and various types of genetic analyses involving backcrossing experiments in a wide range of species have been documented extensively in the literature (Hospital, 2005). The development of NILs can be accelerated considerably by molecular marker-assisted selection (Frisch and Melchinger, 2005), and today marker-assisted backcross breeding is commonly used in crop improvement.

A series of NILs in the same recurrent parent background can also be used to dissect traits at a genome-wide scale. In Arabidopsis, six developmental traits with different heritabilities were analyzed using a recombinant inbred line population in parallel with a genome-wide NIL population. The NILs had greater power than the recombinant inbred line population to detect small-effect quantitative trait loci, at the expense of local resolution (Keurentjes et al., 2007). In maize (*Zea mays*), a set of 89 NILs was created using marker-assisted selection to analyze flowering-time traits (Szalma et al., 2007). In soybean (*Glycine max*), an iron-inefficient NIL and differential seed protein content NIL were analyzed using several existing and emerging methodologies for genetic introgression mapping: single-feature polymorphism analysis, Illumina GoldenGate single nucleotide polymorphism (SNP) genotyping, and de novo SNP discovery via RNA-Seq analysis of next-generation sequence data, with the latter being most informative (Severin et al., 2010). In barley, an extensive backcrossing program was initiated in the mid-1980s to introgress mutated loci from the worldwide collection of morphological and developmental mutants into a common genetic background, the cultivar Bowman (Franckowiak et al., 1985; Fig. 1).

We recently developed two high-throughput SNP-based genotyping arrays based on Illumina's oligo pool assays (OPAs) that each allow simultaneous genotyping of 1,536 SNPs across the barley genome (Close et al., 2009). Using this platform we genotyped several barley reference populations allowing the development of a gene-based consensus map containing 2,943 gene-based SNPs and covering a genetic distance of 1,636 centimorgans (cM; Close et al., 2009). This gene-based map provides a template for the analysis of wide range of barley genetic stocks with the markers themselves, providing a homology bridge to the fully sequenced genomes of rice (*Oryza sativa*), maize, *Brachypodium*, and sorghum (*Sorghum bicolor*). As a result, conserved synteny can provide a glimpse of the putative gene content of any region of the barley genome. Here, we report the development and genotypic analysis of an extended collection of NILs containing mutant alleles that induce a majority of the morphological and developmental variation described in barley.

RESULTS

Characteristics of the Bowman NIL Collection

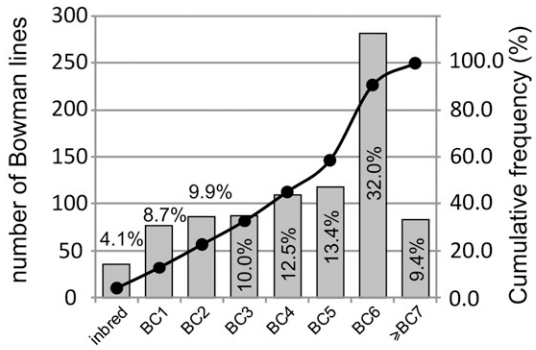
We produced a series of independent barley backcross-derived lines containing mutant alleles by recurrent backcrossing followed by selfing, with phenotypic selection in each cycle. Photographs of some mutant groups are shown in Figure 1. For all lines, the two-rowed U.S. spring-type cultivar Bowman (Franckowiak et al., 1985) was used as the recurrent parent. A summary of the basic characterization data is presented according to the level of backcrossing in Figure 2 and Supplemental Table S1.

The final collection of 881 lines varied in genetic complexity from F1 inbred lines to BC_{10} inbred lines (Fig. 2A). Of these, 768 lines were BC_2 inbred or greater and as such can more accurately be considered NILs. The original mutation containing donor genotypes were assembled, characterized, and documented by the barley mutation research community over a 50-year period in the mid-20th century. They were represented by over 300 different barley accessions. However, approximately one-half of the mutations were identified in one of eight cultivated barley genotypes (cultivars; Fig. 2B). The remaining donors were represented in less than 20 NILs each. For simplicity, we refer to all 881 lines as the Bowman NILs.

Two hundred ninety-eight (34%) of the introgressed phenotypic variants were spontaneous mutants while

Figure 1. Phenotypes of some barley backcross-derived lines carrying mutations affecting morphological and developmental processes such as tillering and plant stature (A–E), lateral floret development (F–L), inflorescence branching (M–O), rachis internode length (P–S), awn development (T–Y), color of the floral organs (Z–AH), chlorophyll biosynthesis in leaf blades and culms (AI–AO), development of the ligular region (AP–AV), early maturity (AW and AX), and necrosis (AY–BD). Phenotypes of the recurrent parent of these lines cv Bowman is shown in A, F, G, M, AG, and AP. For the full description, see Supplemental Text S1. Abbreviations: in G and H: ltFLO,lateral floret; cOGL, central outer glumes; cFLO, central floret; ltOGL, lateral outer glumes. In section AP: LIG, ligule; AUR, auricles.

A



B

	Total	inbred	BC1	BC2	BC3	BC4	BC5	BC6	>=BC7
Bonus	185	5	10	19	28	27	25	59	12
Foma	60		6	13	10	9	7	14	1
Betzes	41	1	2	3	2	6	16	10	1
Akashinriki	34				3	5	3	15	8
Morex	32	6	6	4	5	7	1	2	1
Steptoe	31	3	7	5	7	4	2	3	
Volla	22		2	2	3	4	4	7	
Birgitta	21		1	1	2	2	3	9	3

C

	Total	inbred	BC1	BC2	BC3	BC4	BC5	BC6	>=BC7
spontaneous	298	11	25	31	17	27	52	95	40
X-ray	125	2	6	7	8	20	16	55	11
fast neutron	109	14	13	12	20	17	8	21	4
sodium azide	73		3	8	11	8	9	25	9
ethyl methanesulfonate	61	1	4	9	7	10	7	17	6
ethylene imine	52	1	6	5	11	6	8	12	3
γ-ray	39	1	5	4	3	3	4	18	1

D

	Total	inbred	BC1	BC2	BC3	BC4	BC5	BC6	>=BC7
Male sterile	105		11	20	4	6	21	35	8
Eceriferum	92		2	8	11	16	7	40	8
Semidwarf	57		9	7	11	13	8	7	2
Laxatum	48	2	7	10	10	5	6	8	
Dense spike	47		2	4	5	9	8	16	3
Chlorina	43	6	5	2	4	4	3	12	7
Short awn	38		1	4	2	5	7	12	7
Brachytic	31		2		1	2	6	17	3
Erectoides	30	1		2	2	2	5	13	5
Necrotic leaf spot	29	7	6	1	5	4	3	3	
Desynapsis	24		3	2	2	6	10		
Early maturity	22		3	1		5	2	7	4

Figure 2. Basic features of the Bowman NIL population. A, Population-wide distribution of backcrossing levels. Absolute (histogram, left *y* axis) and cumulative (line graph, right *y* axis) frequencies are plotted. Relative frequencies of the individual lines are shown above or within the histograms as percentages. B and D, Number of NILs at different backcross generations according to the original parent (B), the mutagen used (C), and the phenotypic classification (D). Only groups represented by a total of >20 NILs are shown. Cell shading (white to black) is based on the number of NILs (shown within the cells) within each individual group. Groups are ordered according to the total number of lines per class (highest on top).

583 (56%) were induced by approximately 30 different mutagenic agents (Fig. 2C). The Bowman NILs are mostly affected in barley flower (236 lines), leaf (164 lines), and culm development (166 lines; Figs. 1 and 2D).

Genotypes of the Bowman NILs

Genotypic analysis of the most common donor lines with up to 3,072 SNPs (BOPA1 and BOPA2; Close et al., 2009) allowed us to predict that, in the majority of cases, mutant alleles originating from these donors and retained in the Bowman NILs would be carried on genomic introgressions with different regional SNP-marker haplotypes from those of the recurrent parent. Consequently, in lines that retain only a small donor genome segment, we make the assumption that any observed SNP polymorphisms are most likely to be genetically linked to the mutant alleles. After geno-

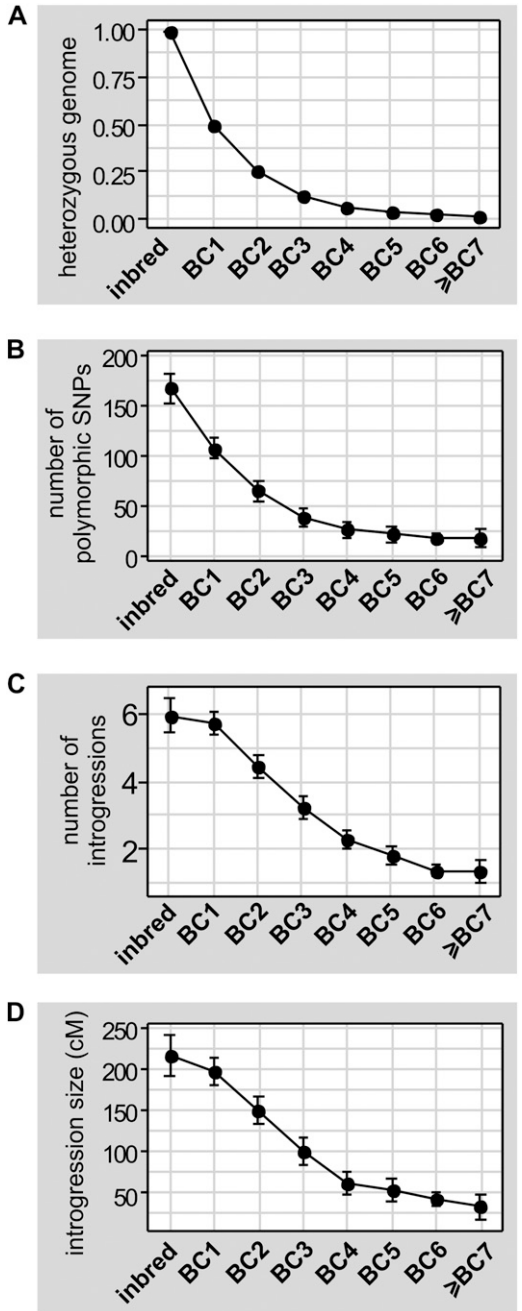


Figure 3. Reduction in genome complexity by backcrossing. Individual value plots show predicted or theoretical fraction of the heterozygous recurrent parent genome after each backcrossing event (A), calculated using the following formula $r = (2^m - 1)/2^m$ (where m is the number of backcrosses); average values of the polymorphic SNP frequencies (B); number of introgressions (C); and the interval size (D). Vertical error bars show 95% Bonferroni corrected confidence intervals for the mean with errors pooled across the groups (inbred: $n = 36$, BC1: $n = 77$, BC2: $n = 87$, BC3: $n = 88$, BC4: $n = 110$, BC5: $n = 118$, BC6: $n = 282$, and >=BC7: $n = 83$).

typing all 881 Bowman NILs, we observed, as expected, that different backcross generations revealed considerable variation in genetic complexity (Fig. 3; Supplemental Table S2). In terms of the frequency of polymorphic SNPs between the NIL and the cv Bowman, the complexity reduction correlated well with the backcross generation following the established model (Briggs, 1938; Fig. 3, A and B). Only 37 Bowman NILs revealed no polymorphisms after this analysis. In these lines we conclude that either the effective mutation was induced in a region of the genome with the same haplotype as Bowman or the introgression is too small to be detected at the current marker density. In some cases, where the seed stocks were available as a mixture of homozygous dominant and heterozygous genotypes (for example, lines carrying *male sterile genetic* [*msg*] mutations) we may have selected for genotyping the homozygous dominant line, resulting in genotypes indistinguishable from the cv Bowman.

Size and Chromosomal Distribution of the Introgressed Regions

The majority of the SNPs used to genotype the Bowman isolines were also used to construct a barley consensus map (Close et al., 2009). This map context enables a straightforward estimation of the number of introgressions present in individual NILs and their size. For both of these parameters we found concordance with theoretical predictions (Briggs, 1938; Fig. 3, A, C, and D). Low resolution linkage assignments based on a combination of isozyme data and testcrosses with classical genetic stocks have been reported previously for 461 Bowman NILs, and 295 of these (those with one or two introgressions) can be compared to our mapping results (Supplemental Tables S1 and S2). At the chromosomal level, there is 78% correspondence between the classical and the SNP mapping data. The noncorresponding fraction can likely be explained by less-powerful classical mapping techniques that possibly led to more misplacement, by genetic stock mix ups during generation of the backcross-derived lines or by residual introgressions that, while retained, are unlinked to the trait.

The introgressed segments in the Bowman NILs reveal a relatively even distribution across the barley genome (Fig. 4). However, on average 3 times higher introgression frequencies were found at the genetic centromeres (crudely defined as the central regions of each barley chromosome where the frequency of SNP markers on the genetic map is significantly increased). These represent recombinationally inactive but physically expansive regions that have been estimated to contain between 30% and 50% of mapped barley genes. An increased frequency of mutations (and hence introgressions) contained within these specific genetic intervals is therefore not unexpected.

We considered those NILs that contain very small and defined introgressions (<10 cM or <0.6% of barley genetic map) an important class. Small introgressions provide an opportunity to explore rapid candidate gene identification and isolation, especially

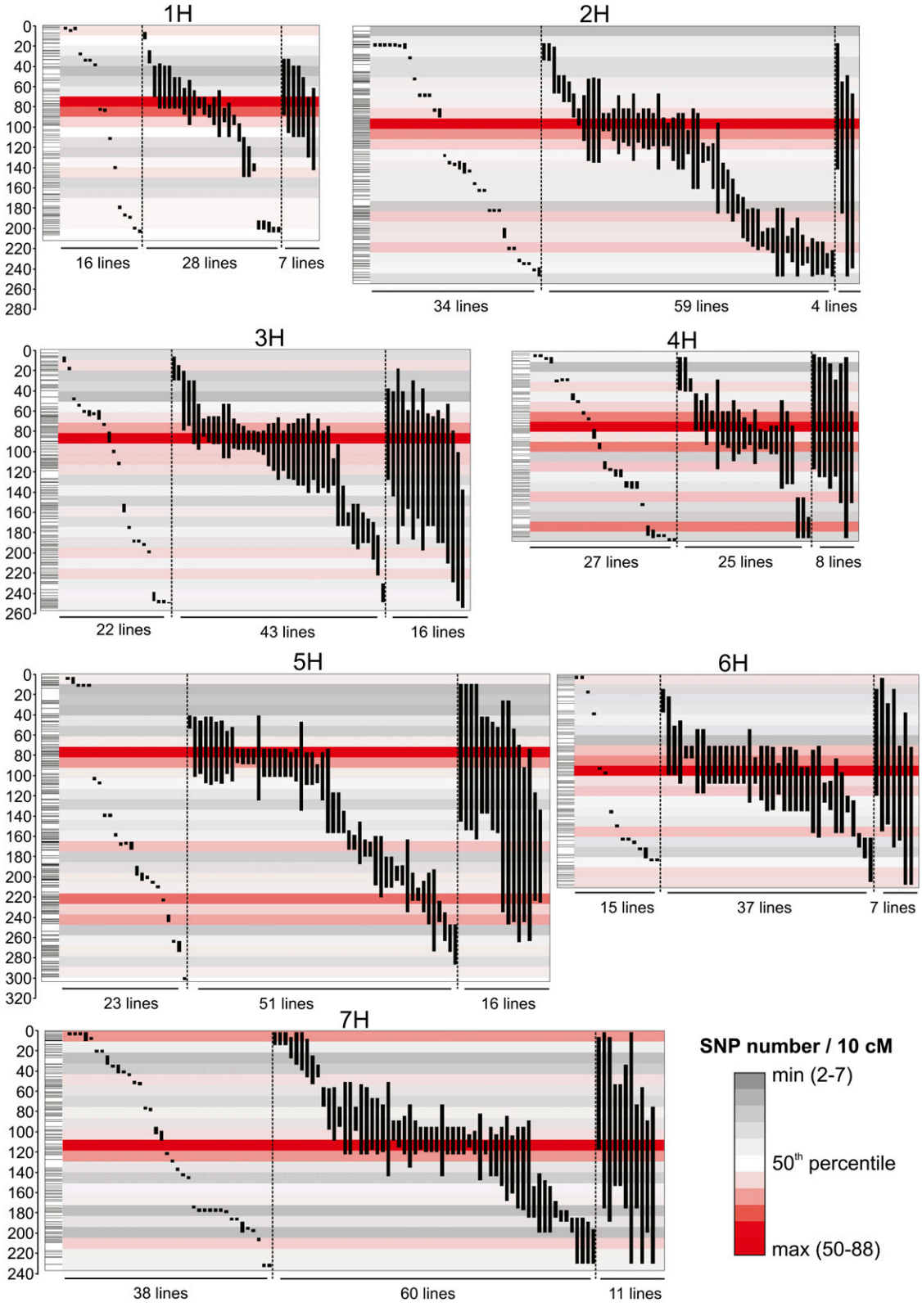


Figure 4. Chromosomal distribution of introgressed segments containing mutant alleles. Four hundred thirty lines that have one or two donor introgressions into Bowman were partitioned into three groups based on introgression size, small (<10 cM), medium (10–60 cM), and large (>60 cM; separated by a dotted vertical line), and ordered based on the average marker positions. Numbers below the chromosome graphs show the number of lines within the groups. Genetic map scale (in centimorgans) and distribution of the loci in the consensus map are shown on the left for all chromosomes. Coloring from white to gray to red is based on the SNP frequency within the 10-cM intervals.

when the mutant locus in the Bowman NILs is supported by a range of additional allelic variants. Of the 881 Bowman NILs, we classified 176 as small introgressions of <10 cM. These represent a range of phenotypes and chromosomal locations (Fig. 4; Supplemental Tables S1 and S2). The average introgression size for this group is 3.7 cM, ranging from 0 cM (32 lines) to 9 to 10 cM (11 lines).

The presence of multiple mutant alleles is valuable for both confirming gene location and eliminating spurious background introgressions. After genotypic alignment genes represented by multiple alleles form groups of complementary overlapping introgressions (COIs) that better resolve the genetic interval containing a mutant locus and eliminate unlinked introgressions. Based on phenotypic and genotypic information, 414 lines form 98 groups (genes) of COIs altogether representing 35 different phenotypic classes. In the majority of cases these effectively eliminate spurious introgressions from further genetic analysis. However, for about one-half of the groups, there is no gain in resolution from using COIs, as the line containing the smallest introgression is completely contained within the introgressions present in the other lines. However, for 18 groups a 10% to 90% reduction of the putative gene-containing region can be inferred (see example in Supplemental Fig. S1).

Exploiting Synteny for Gene Prediction in Barley

To explore the feasibility of mutant gene identification using the SNP mapping data and synteny models, we analyzed five NILs where the identity of the mutant locus is known: BW069 and BW880 lines containing the *THIRD OUTER GLUME* (*HvTRD1*) gene on chromosome 1H (Whipple et al., 2010), BW898 containing *HvVRS1* (Komatsuda et al., 2007) on 2H, BW885 containing *HvUZU* (Chono et al., 2003) on 3H, and two loci on chromosome 7H, namely BW638 containing the *HvNUD1* gene (Taketa et al., 2008) and BW905 containing *HvWAXY* (Rohde et al., 1988). Based on the assumption that synteny is conserved with rice and that major structural rearrangements are absent, we estimated the number of genes contained within the introgressed segments.

For *HvTRD1* there is no barley EST available; hence no SNP assay or probe set has been designed (Fig. 5A). The rice ortholog of *HvTRD1* is *NECK LEAF1* (*OsNL1*). Based on our barley-rice synteny model, the inferred position of *OsNL1* or *HvTRD1* is 132.5 cM on chromosome 1H. This is within the 6.7-cM introgression defined by four polymorphic SNPs of the *HvTRD1* carrying line BW880; however, only two of them have rice homologs in the syntenic position. Even so, the inferred (model-based) position of the *HvTRD1* gene is less than 100 rice gene models away (Fig. 5A). By considering additional markers inferred from comparative rice-barley gene expression analyses the interval can be further reduced to six rice genes (K. Houston, unpublished data).

A comparison between Bowman and BW898 (*HvVRS1*) identified six polymorphic SNPs within a 6.3-cM interval on chromosome 2H (Fig. 5B). Although the rice ortholog of *HvVRS1* is not at a conserved syntenic position, five of the six SNPs identify rice orthologs that span a 2.5-Mb region on chromosome Os04 containing 303 gene models. The SNPs immediately flanking *HvVRS1* define a 2.8-cM interval, spanning 300 kb in rice and 34 gene models (Fig. 5B). In BW885 (*HvUZU*), 10 of 15 polymorphic SNPs clustered on chromosome 3H defined a 10.3-cM introgression close to the genetic centromere (Fig. 5C). A separate introgression remained on chromosome 1H. Nine of the 3H SNP-containing genes have clear rice homologs, and seven belong to a conserved block of synteny covering 2.3 Mb containing 320 gene models. The SNPs immediately flanking *HvUZU* are 5.6 cM apart, delimiting a 370-kb region containing 56 gene models (Fig. 5C). For *HvNUD1*, three cosegregating SNPs were located within a 190-kb region of the rice genome containing 27 gene models (Fig. 5D). This region contains 20 gene models (approximately 100 kb) proximal to the rice homolog of *HvNUD1*, but still within the 95% confidence boundaries of the synteny model. Finally, BW905 (*HvWAXY*) has a 17.8-cM introgression defined by eight SNPs. The genes containing these SNPs all have rice homologs, with five spread over a 1.2-Mb region on rice chromosome Os06 that harbors 169 gene models. The two *HvWAXY* flanking markers are 360 kb apart and the interval contains 50 putative genes (Fig. 5D).

DISCUSSION

A theoretical framework for estimating the proportion of donor genome likely to remain in NILs developed by backcross breeding (Briggs, 1938; Hospital, 2001) has led to the widespread use of at least six backcross generations to recover an introgressed phenotype in an otherwise recurrent parent background. Our detailed genotypic characterization of over 800 BC_1 to BC_{10} derived inbred lines provides, at an unprecedented scale, an accurate and representative practical appraisal of the reduction in genome complexity achieved by different levels of backcrossing. Using a combination of related parameters (Fig. 3) our results reveal a good concordance with the theoretical predictions (Briggs, 1938; Hospital, 2001).

Four hundred twenty-six Bowman NILs with introgressed segments of <50 cM in length (i.e. <3% of the barley genetic map) are powerful resources for both understanding the biology of the affected phenotypes and for gene discovery by forward genetics. However, as this remains laborious in large genome cereals, it is routinely enhanced by exploiting conservation of synteny with the fully sequenced rice, *Brachypodium*, sorghum, and maize genomes (Flavell, 2009). The gene-based SNP platform used to genotype the Bowman NILs facilitates this analysis as the gene sequences

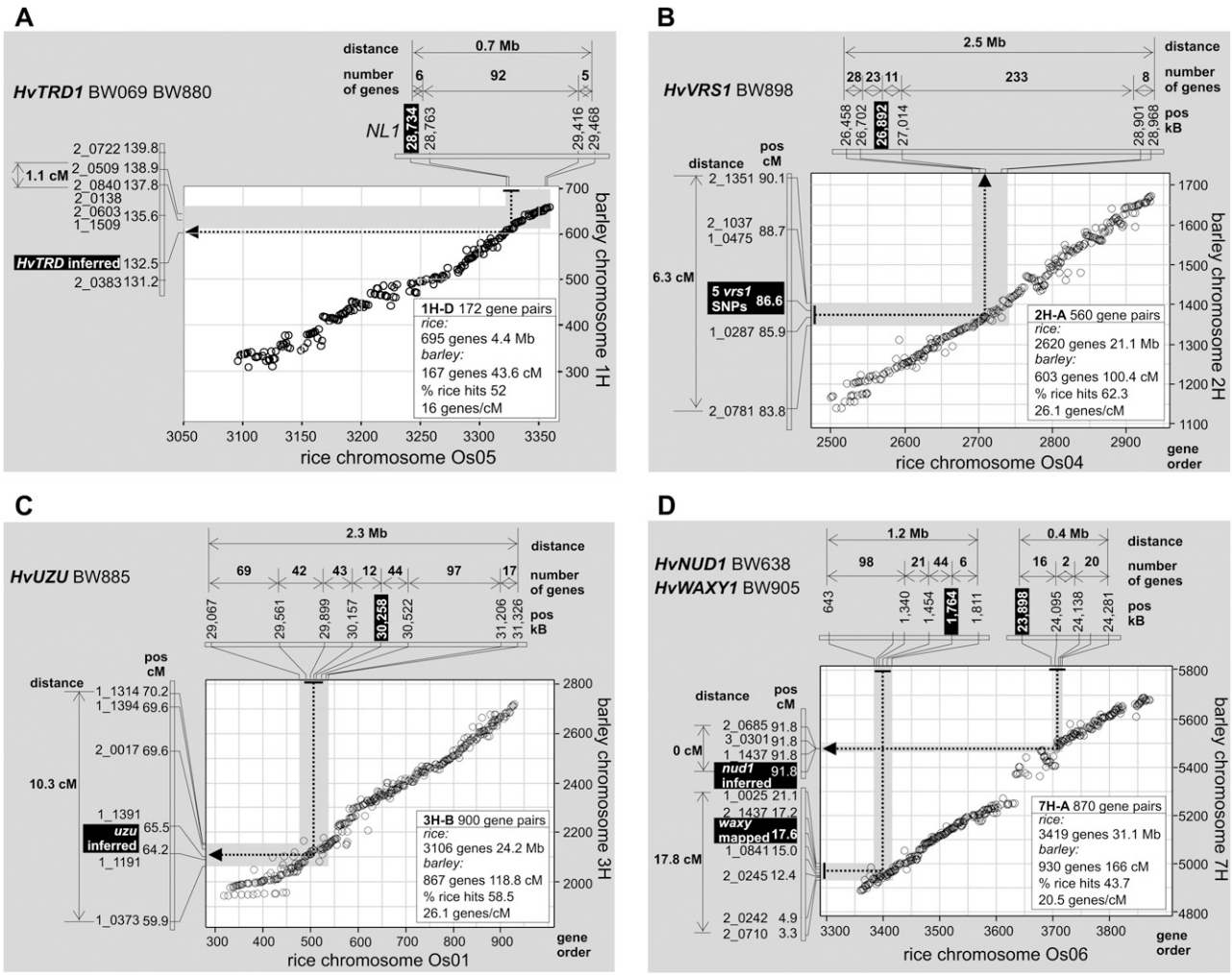


Figure 5. Predictions of the position and precision of locating *HvTRD1* (A), *HvVRS1* (B), *HvUZU* (C), and *HvNUD1* and *HvWAXY* (D) genes using Bowman NIL genotypic data, in conjunction with the current barley consensus gene map and barley-rice conservation of synteny models. Scatterplot is based on barley-rice homolog pairs, plotting genome-wide ordinal values where the *x* axis = rice physical gene order, *y* axis = barley genetic gene order. Insets show summaries of the model inputs. Gray areas within the scatterplot indicate both the barley and the inferred rice regions as defined by SNPs that are polymorphic between cv Bowman and the corresponding NIL. Expansion of these regions on the physical and genetic scales is shown as graphs on the top and left of the scatterplots depicting polymorphic SNPs and their genetic positions (barley graph) and physical positions and number of genes in the intervals (rice graph). Positions of the actual genes and/or their inferred positions based on synteny are shown as black boxed white text. The direction of the arrow (the dotted line) within the scatterplot indicates how the position of each of the genes was predicted. For *HvTRD1*, *HvUZU*, and *HvNUD1* the arrow points from rice to barley because these genes have not been located genetically on our SNP map, but have rice homologs that are assumed to be in a syntenic position and therefore can be used to predict the position of the barley homologs. The arrow points from barley to rice for *HvVRS1* because the corresponding SNPs have been mapped in barley. Note that the rice homolog is not at the syntenuous position. No arrow is shown for *HvWAXY* because both barley and rice homologs exist for this gene.

represent a common currency that reliably anchors almost 90% of barley genetic map to the sequence of these genomic models (Goff et al., 2002; Supplemental Table S3).

To explore the feasibility of mutant gene identification using the SNP mapping data and synteny models, we analyzed five NILs where the identity of the mutant locus is known: *HvTRD1* (Whipple et al., 2010), *HvVRS1* (Komatsuda et al., 2007), *HvUZU* (Chono et al., 2003), *HvNUD1* (Taketa et al., 2008), and *HvWAXY* (Rohde et al., 1988). Based on the assumption that synteny is conserved with rice and that major structural rearrangements are absent, we estimated the number of genes contained within the introgressed segments. The number of positional candidate rice gene models ranged from 22 (*HvNUD1*) to 303 (*HvVRS1*). While these examples illustrate possible outcomes from inferring the location of a mutated locus using the NIL genotypic data alone, they also provide context for the likely impact of our data on morphological and devel-

opmental biology, and more broadly on trait genetics in the Triticeae cereals. The defining SNPs are ready-to-go gene-based markers that enable highly efficient targeting for positional cloning projects. In such endeavors the emerging draft barley genome sequence will play an increasingly important role (Schulte et al., 2009), allowing individuals to focus on positionally and functionally consistent candidate genes where simple targeted resequencing of wild-type and mutant alleles would represent a rapid validation approach. For many mutants, such an off-the-shelf validation strategy is already available in the large collections of documented mutant allele series housed and maintained in germplasm collections such as NordGen (http://www.nordgen.org/) and the U.S. Department of Agriculture National Small Grains Collection (http://www.ars-grin.gov/npgs/acc/acc_queries.html). Predicted gene frequencies, especially in highly recombinogenic regions, suggest that for many loci, gene identification and validation by this route is both realistic and possible.

Despite the above successes, in the majority of cases, and especially those for mutants lacking multiple allelic variants, positional cloning using recombinant populations of outcrossed Bowman lines will remain essential to determining the identity of a mutated gene. Collecting this additional genetic data is important because there will be instances where mutations have been generated in donor lines that have identical regional haplotypes to Bowman, which could result in comparative analysis being conducted in the wrong genetic region. We have therefore crossed approximately 100 selected Bowman mutants to the barley cultivars Bowman, Morex, and Barke, generating extended F2 and/or F3 populations specifically for positional cloning applications. In this context, the Bowman NIL genotypic data gives a head start by providing mapped and polymorphic genetic markers that can be used immediately to validate the location of a mutant gene. Subsequently, through comparative genetic analysis, assembly of the putative regional barley gene content may identify strong candidate genes or even functional orthologs of characterized genes in related grasses.

Selection during backcrossing was mostly based on a single character (reflected in the locus name), but many mutations act pleiotropically, affecting multiple aspects of barley plant morphology and development (e.g. Fig. 2, AP–AV). One task for the future will be a detailed and systematic comparative phenotypic analysis of the Bowman NILs at the whole plant level at different developmental stages. When the identity of many individual genes is determined, such a phenotypic data set may facilitate the interpretation of a gene's action and interactions underlying related phenotypes.

CONCLUSION

Many trait-based quantitative trait loci analyses have been conducted in barley (and in closely related species) and by meeting the twin imperatives of positional and biological concordance, our hypothesis is that many of the mutants that we have genetically characterized represent extreme examples of natural genetic variation commonly measured as quantitative variation in natural or constructed populations (Druka et al., 2010). If this hypothesis is correct, the barley mutants described here provide a convenient and tractable Mendelian solution to the identification and characterization of genes controlling non-Mendelian or quantitative traits.

MATERIALS AND METHODS

Generation of the Bowman Lines

The generation of barley (*Hordeum vulgare*) lines with single morphological markers started in 1985 using genetic stocks with multiple morphological markers (Wolfe and Franckowiak, 1991). These included Master Dominant (GSHO 3450) and Master Recessive (GSHO 3451) stocks and other recessive phenotype stocks covering all seven barley chromosomes. Subsequently other genetic stocks from many sources were acquired and used for the backcrossing project (Supplemental Table S1). The two-rowed spring barley cultivar Bowman (PI 483237; Franckowiak et al., 1985) was selected as a recurrent female parent. Three to six F1 seeds were planted, followed by 60 to 120 F2 progenies. From each F2 progeny, two or three plants that visually exhibited the mutant phenotype were harvested. Four F3 seeds from one plant were sown and one was used for subsequent backcrosses.

Seed Availability

Seeds of the Bowman backcross-derived lines can be obtained from the Scottish Crop Research Institute (Invergowrie, Dundee, United Kingdom) by contacting Arnis Druka (arnis.druka@scri.ac.uk). Alternatively, they can be requested from the National Small Grains Collection, U.S. Department of Agriculture, Agricultural Research Service, National Small Grains Germplasm Research Facility, National Small Grains Collection (Aberdeen, Idaho) by contacting Harold E. Bockelman (harold.bockelman@ars.usda.gov), or from the Nordic Genetic Resource Center (Alnarp, Sweden) by contacting Morten Rasmussen (morten.rasmussen@nordgen.org).

Genotyping

Plants were grown in a glasshouse maintained between 16°C and 24°C with natural light and supplemented with high-pressure sodium lamps to provide 16 h day length. Automated irrigation was used to keep compost reasonably wet. Fungicidal spraying and insecticidal fumigation was applied at regular intervals. One leaf tip at Zadoks 11 developmental stage was harvested and placed into 96-deep well blocks and used for DNA extraction using a Nucleoplex plant DNA kit (product no. 33300, Tepnel Life Sciences PLC) on a Nucleoplex machine (Tepnel Life Sciences PLC) according to the manufacturer's protocols. DNA was eluted in 100 μL of 10 mM Tris (pH 8). DNA quality was assessed by running 2 μL of each sample on an ethidium bromide-stained 1% agarose gel and visualization under UV light. The DNA quantity was estimated by comparing sample band intensity to known amount of standard DNA (Hyperladder1, BIO-33025, Bioline).

SNP genotyping was carried out at the Southern California Genotyping Consortium (SCGC) at the University of California, Los Angeles. Between 500 ng and 1 μg of genomic DNA was provided in 96-well skirted v-bottom polypropylene microplates (catalogue no. MSP–9601) from Bio-Rad sealed with strip caps. Genomic DNA was quantified using a Quant-iT Picogreen dsDNA assay kit (Invitrogen no. P11496) and normalized to 50 ng per μL. A total of 250 ng of gDNA was used per reaction with 1,536 BOPA1 and BOPA2 SNPs typed in each reaction.

The multiplexed genotyping reactions (Golden Gate assay) themselves were based on well-characterized biochemical reactions but carried out under highly controlled and optimized conditions. An oligo pool containing allele-specific and locus-specific oligos tailed with IllumiCode sequences and universal primer sequences were created for 1,536 unique barley SNP loci per each

pool (BOPA1 and BOPA2; Close et al., 2009). The oligo pool was allowed to hybridize to genomic DNA, undergo primer extension and ligation such that each allele generated a unique PCR template with common primer sequences. A multiplexed PCR generated labeled products from the templates and the amplified products were hybridized to the decoded array so that the genotype information for each locus was scanned, decompiled, and recorded by an Illumina 1000 SNP genotyping system.

Genotype Inference and Identification of the Polymorphisms

BeadLab output files (.idat) corresponding to cy3 and cy5 dye intensities for each sample were uploaded from the SCGC server for analysis. They were imported in the Beadstudio v2.0 (Ilumina, Inc.) software and genotype calls were calculated using Cluster All SNPs function, resulting in 683,941 (45.5%) AA genotype scores, 690,682 (45.9%) BB genotype scores, and 65,399 (4.3%) AB genotype scores. A total of 63,722 (4.2%) cy3 and cy5 pairs did not result in reliable signal, therefore were assigned as no call (NC). The cumulative frequency genotype call distribution across the SNPs and across the lines shows that the majority of AB and NC calls are due to underperforming SNPs, rather than lines (DNA samples). Thus, only 11 lines had call rates less than 0.9; of those, two had 0.58 and 0.73, the rest had better than 0.8.

All the lines were included in the analysis, but 342 markers that had low marker score or were not informative (low level of allelic discrimination) were removed from the analysis. Further refinement of the genotypes was by visual inspection of the raw data. Thus, genotypes were reported by 1,194 SNP markers and consecutively used to identify those that are polymorphic between cv Bowman and each NIL.

After preliminary data analysis, 84 NILs that had no or a single polymorphism were genotyped with BOPA2. Summary statistics of the genotype scores for these 84 lines was similar to that obtained by the BOPA1 genotyping of 978 NILs: 54,717 (42.4%) AA calls; 68,229 (52.8%) BB calls; 1,431 (1.1%) AB calls; and 4,647 (3.6%) were assigned as NC. BOPA2 SNPs that were polymorphic between Bowman and each NIL were combined with those identified by BOPA1, aligned to the barley consensus map (Supplemental Table S2), and used for the analysis of the introgression number and size (Fig. 3).

Note that in the case of mutants that are not viable as homozygotes (e.g. *Msg* lines), the single plant chosen for genotypic analysis would have had a 1:4 chance of being wild type at the mutant locus due to segregation in an effective F2. The data from these lines should be treated with the necessary caution.

Consensus Linkage Map

Construction of the barley consensus linkage map has been reported before (Close et al., 2004). In summary, the map was assembled based on 334 segregating double-haploid lines represented by three biparental F1 doubled-haploid-based populations Steptoe × Morex, Morex × Barke, and OWB-D × OWB-R. The map consisted of 965 loci represented by 2,945 SNP markers.

Rice Physical Map—Barley Genetic Map Alignment and the Model Development

The Institute for Genomic Research rice (*Oryza sativa*) gene build v5 was used for building synteny models. The number of gene models was reduced from 56,278 to 41,129 genes by excluding repeated elements and their genes. The consensus sequences of the barley EST assemblies (Close et al., 2007) numbers 21, 32, and 35 were used for the homology search against the rice genomic sequence using the tblastX algorithm (Altschul et al., 1990) with e-value cutoff <e-05, resulting in 38,975 top hits for the assembly number 21, 31,265 for number 32, and 36,053 for number 35. The total number of unique rice hits was 17,868. Transcript-derived markers (TDMs; Potokina et al., 2008) were integrated in the SNP-based map as described previously (Druka et al., 2008). The map distances in the integrated map were adjusted by taking into account TDM mapping data (A. Druka, unpublished data). Scatterplots of the gene (SNP and TDM marker) order values from the barley consensus map and homologous gene order values from the rice physical map identified the general pattern of blocks with high conservation of synteny (Supplemental Table S3). Small inversions within the blocks were rotated, and genes that were clearly outside the synteny region were removed. These blocks were used to predict the position of barley genes in rice and vice versa.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Reduction of the gene-containing interval using groups of COIs.

Supplemental Table S1. Summary of the Bowman backcross-derived line population.

Supplemental Table S2. Bowman backcross-derived line genotyping data.

Supplemental Table S3. Rice-barley synteny regions.

Supplemental Text S1. File contains extended legend for the Figure 1, description of the column headings for all three supplemental tables, and note about naming barley genes and alleles.

ACKNOWLEDGMENTS

We thank Joe DeYoung at SCGC for performing the genotypic analysis.

Received September 18, 2010; accepted November 14, 2010; published November 18, 2010.

LITERATURE CITED

Ahloowalia BS (2004) Global impact of mutation-derived varieties. Euphytica **135:** 187–204

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Briggs F (1938) The use of the backcross in crop improvement. Am Nat **72:** 285–292

Chandler PM, Marion-Poll A, Ellis M, Gubler F (2002) Mutants at the Slender1 locus of barley cv Himalaya: molecular and physiological characterization. Plant Physiol **129:** 181–190

Chono M, Honda I, Zeniya H, Yoneyama K, Saisho D, Takeda K, Takatsuto S, Hoshino T, Watanabe Y (2003) A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. Plant Physiol **133:** 1209–1219

Clarke B, Liang R, Morell MK, Bird AR, Jenkins CL, Li Z (2008) Gene expression in a starch synthase IIa mutant of barley: changes in the level of gene transcription and grain composition. Funct Integr Genomics **8:** 211–221

Close TJ, Bhat PR, Lonardi S, Wu Y, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S, et al (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics **10:** 582

Close TJ, Wanamaker S, Roose ML, Lyon M (2007) HarvEST. Methods Mol Biol **406:** 161–177

Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA, Wing RA, Muehlbauer GJ, Kleinhofs A, Wise RP (2004) A new resource for cereal genomics: 22K barley GeneChip comes of age. Plant Physiol **134:** 960–968

Druka A, Druka I, Centeno AG, Li H, Sun Z, Thomas WT, Bonar N, Steffenson BJ, Ullrich SE, Kleinhofs A, et al (2008) Towards systems genetic analyses in barley: integration of phenotypic, expression and genotype data into GeneNetwork. BMC Genet **9:** 73

Druka A, Franckowiak J, Lundqvist U, Bonar N, Alexander J, Guzy-Wrobelska J, Ramsay L, Druka I, Grant I, Macaulay M, et al (2010) Exploiting induced variation to dissect quantitative traits in barley. Biochem Soc Trans **38:** 683–688

Felker FC, Peterson DM, Nelson OE (1983) Growth characteristics, grain filling, and assimilate transport in a shrunken endosperm mutant of barley. Plant Physiol **72:** 679–684

Flavell R (2009) Role of model plant species. Methods Mol Biol **513:** 1–18

Franckowiak JD, Foster AE, Pederson VD, Pyler RE (1985) Registration of 'Bowman' barley. Crop Sci **25:** 883

Frisch M, Melchinger AE (2005) Selection theory for marker-assisted backcrossing. Genetics **170:** 909–917

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J,

Sessions A, Oeller P, Varma H, et al (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science **296:** 92–100

Harten AM (1998) Mutation Breeding: Theory and Practical Applications. Cambridge University Press, New York

Hospital F (2001) Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. Genetics **158:** 1363–1379

Hospital F (2005) Selection in backcross programmes. Philos Trans R Soc Lond B Biol Sci **360:** 1503–1511

Keurentjes JJ, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M (2007) Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. Genetics **175:** 891–905

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, et al (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci USA **104:** 1424–1429

Lundqvist U (2009) Eighty years of Scandinavian barley mutation genetics and breeding. *In* QY Shu, ed, Induced Plant Mutations in the Genomics Era. Food and Agriculture Organization of the United Nations, Rome, pp 39–43

Morell MK, Kosar-Hashemi B, Cmiel M, Samuel MS, Chandler P, Rahman S, Buleon A, Batey IL, Li Z (2003) Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties. Plant J **34:** 173–185

Müller KJ, Romano N, Gerstner O, Garcia-Maroto F, Pozzi C, Salamini F, Rohde W (1995) The barley Hooded mutation caused by a duplication in a homeobox gene intron. Nature **374:** 727–730

Potokina E, Druka A, Luo Z, Wise R, Waugh R, Kearsey M (2008) Gene expression quantitative trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. Plant J **53:** 90–101

Pozzi C, di Pietro D, Halas G, Roig C, Salamini F (2003) Integration of a barley (Hordeum vulgare) molecular linkage map with the position of genetic loci hosting 29 developmental mutants. Heredity **90:** 390–396

Röder MS, Kaiser C, Weschke W (2006) Molecular mapping of the shrunken endosperm genes seg8 and sex1 in barley (Hordeum vulgare L.). Genome **49:** 1209–1214

Rohde W, Becker D, Salamini F (1988) Structural analysis of the waxy locus from Hordeum vulgare. Nucleic Acids Res **16:** 7185–7186

Rossini L, Vecchietti A, Nicoloso L, Stein N, Franzago S, Salamini F, Pozzi C (2006) Candidate genes for barley mutants involved in plant architecture: an in silico approach. Theor Appl Genet **112:** 1073–1085

Rostoks N, Schmierer D, Mudie S, Drader T, Brueggeman R, Caldwell DG, Waugh R, Kleinhofs A (2006) Barley necrotic locus nec1 encodes the cyclic nucleotide-gated ion channel 4 homologous to the Arabidopsis HLM1. Mol Genet Genomics **275:** 159–168

Schulte D, Close TJ, Graner A, Langridge P, Matsumoto T, Muehlbauer G, Sato K, Schulman AH, Waugh R, Wise RP, et al (2009) The international barley sequencing consortium—at the threshold of efficient access to the barley genome. Plant Physiol **149:** 142–147

Severin AJ, Peiffer GA, Xu WW, Hyten DL, Bucciarelli B, O'Rourke JA, Bolon YT, Grant D, Farmer AD, May GD, et al (2010) An integrative approach to genomic introgression mapping. Plant Physiol **154:** 3–12

Somers DA, Kuo TM, Kleinhofs A, Warner RL (1983) Nitrate reductase-deficient mutants in barley: immunoelectrophoretic characterization. Plant Physiol **71:** 145–149

Stadler LJ (1928) Mutations in barley induced by X-rays and radium. Science **68:** 186–187

Szalma SJ, Hostert BM, Ledeaux JR, Stuber CW, Holland JB (2007) QTL mapping with near-isogenic lines in maize. Theor Appl Genet **114:** 1211–1228

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

von Wettstein D (2007) From analysis of mutants to genetic engineering. Annu Rev Plant Biol **58:** 1–19

Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, Jackson DP (2010) A conserved mechanism of bract suppression in the grass family. Plant Cell **22:** 565–578

Wolfe RI, Franckowiak JD (1991) Multiple dominant and recessive genetic marker stocks in spring barley. Barley Genet Newsl **20:** 117–121

Zhang L, Lavery L, Gill U, Gill K, Steffenson B, Yan G, Chen X, Kleinhofs A (2009) A cation/proton-exchanging protein is a candidate for the barley NecS1 gene controlling necrosis and enhanced defense response to stem rust. Theor Appl Genet **118:** 385–397