

ORIGINAL ARTICLE

A high-density transcript linkage map of barley derived from a single population

K Sato, N Nankaku and K Takeda

Research Institute for Bioresources, Okayama University, Kurashiki, Japan

A high-resolution transcript linkage map of barley was created using a single doubled haploid (DH) mapping population, only 3'-end expressed sequence tags (ESTs) and only PCR-based assays. Cultivar 'Haruna Nijo' and an ancestral wild-form accession 'H602' were used as EST donors and crossing parents of the mapping population. Of the 10 366 primer sets developed from a non-redundant set of 3' EST sequences, 7700 sets generated useful amplicons and 3975 (52%) showed polymorphisms between the mapping parents. Of these, 2890 (28% of the total) were mapped by single nucleotide polymorphisms (1717), cleaved amplified polymorphic sequence (933) and INDELS (240). The present work involves an estimated 9% of the genes of barley. Of the mapped ESTs, 2689 (93%) are formatted in the Affymetrix Barley 1 GeneChip and full-length cDNA

sequences are available for 1039 (36%). Mapped ESTs show highest similarity with sequences in the wheat gene index (93%) and moderate similarity with rice (50%). Comparison of mapped EST positions and the rice pseudomolecule indicated collinear regions between two species; these are particularly conserved for the entire barley chromosome 3H and rice chromosome 1. These mapped genes, together with a systematically developed set of genetic resources, will make it easier to directly clone genes showing simple inheritance and to determine the genetic basis of complex traits. The information will contribute to the development of a framework for the physical mapping of barley, which is necessary for genome sequencing. *Heredity* (2009) **103**, 110–117; doi:10.1038/hdy.2009.57; published online 20 May 2009

Keywords: *Hordeum vulgare*; expressed sequence tag; single nucleotide polymorphism; genetic map

Introduction

Barley (*Hordeum vulgare* L.) is a self-pollinated diploid ($2n=14$) plant with numerous advantages for the development of genetic stocks and their subsequent analysis (Sato *et al.*, 2003). Barley serves as a model genome for the polyploid (for example, tetraploid and hexaploid cultivated wheat) and outcrossing (for example, rye) members of the Triticeae (Bothmer *et al.*, 2003). Owing to the large diploid genome sizes (ca. 5000 Mbp, Arumuganathan and Earle, 1991), whole genome sequencing in the Triticeae has been deferred and the focus has been on construction of increasingly dense genetic maps.

Historically, barley geneticists spent decades developing high-quality genetic maps, comprised of natural and induced mutants, by using classical three-point linkage tests (Lundqvist *et al.*, 1996). Cytogenetics subsequently contributed to the genetic mapping by physically localizing genes (Künzel *et al.*, 2000). Genetic mapping of barley accelerated with the application of molecular markers to doubled haploid (DH) populations (Chen and Hayes, 1989). These DH populations have the added advantage of allowing for simultaneous genetic map construction and repeated analysis of quantitative traits (Hayes *et al.*, 1993). DH populations were first mapped

with restriction fragment length polymorphisms (RFLPs) (Graner *et al.*, 1991; Heun *et al.*, 1991; Kleinhofs *et al.*, 1993). Owing to their transferability across species, RFLPs were essential for the early synteny studies that identified evolutionary patterns and genome similarities among cereal crop species (Moore *et al.*, 1995; Devos, 2005). Subsequently, amplified fragment length polymorphisms (AFLPs) were widely used for higher-density map development (Waugh *et al.*, 1997; Hori *et al.*, 2003; Takahashi *et al.*, 2006), but this type of marker has issues with data quality. Sequence-based PCR markers—that is, sequence-tagged sites (STSs) (Mano *et al.*, 1999) and simple sequence repeats (SSRs) (Ramsay *et al.*, 2000)—proved to be more reliable and informative and continue to serve as useful anchor loci for maps based on expressed sequence tags (ESTs). Extensive, quality-controlled barley EST information is available in the HarvEST (<http://harvest.ucr.edu/>) database. These resources were also used to develop the Affymetrix Barley 1 GeneChip (Close *et al.*, 2004), a platform for large-scale expression analysis (Druka *et al.*, 2006).

Genetic mapping of ESTs establishes gene location and provides a framework for whole genome sequencing, as shown in rice (Chen *et al.*, 2002; Wu *et al.*, 2002). A high-density EST map is the foundation for map-based genome analysis as it provides a basis for selecting bacterial artificial chromosome (BAC) clones for sequencing (Varshney *et al.*, 2006). Mapped ESTs also provide access to other genomes, particularly that of rice (*Oryza sativa* L.), based on sequence similarity. Stein *et al.* (2007) found significant micro-collinearity between the rice and

Correspondence: Dr K Sato, Research Institute for Bioresources, Okayama University, 2-20-1, Chuo, Kurashiki 710-0046, Japan.

E-mail: kzsato@rib.okayama-u.ac.jp

Received 23 July 2008; revised 12 December 2008; accepted 30 March 2009; published online 20 May 2009

barley genomes using a consensus barley map with 1032 EST-based loci assayed using a combination of marker assays (including 607 RFLPs). Comparative genomic analysis of barley ESTs and rice genome sequences will enable detection of undiscovered barley genome sequences. Methods that systematically generate markers and estimate candidate orthologs based on the rice genome have enabled isolation of genes controlling key agronomic characters in barley (Furukawa *et al.*, 2007; Taketa *et al.*, 2008).

Our approach was to develop a high-resolution barley EST map using a single DH mapping population, only 3'-end ESTs and only PCR-based assays. Two of the EST donors were parents of the DH-mapping population. 'Haruna Nijo' is an elite malting cultivar. 'H602' is an accession representing the ancestral wild form of the species. A key advantage to using a single population for mapping, as opposed to a consensus map based on multiple populations, is that every map position is based on segregation data. We used only 3'-end ESTs generated by our research group to ensure sequence quality, preclude redundancy in mapping transcripts and provide a comprehensive distribution of genes on the linkage map.

Our ESTs were derived from representative strains of barley preserved at the Barley Germplasm Center in Okayama University (<http://www.shigen.nig.ac.jp/barley/>). As these strains represent a broad sample of genetic diversity in terms of degree of domestication, geographic origin and end use, we expected to maximize the degree of polymorphism in EST sequences. We used only PCR-based assays for mapping to provide the research community with accessible tools (for example, STS, CAPS (cleaved amplified polymorphic sequence) and single nucleotide polymorphisms (SNPs)) for further research and application of the mapped EST information. Strategies for using the mapped ESTs to integrate genetic and genomic resources are discussed.

Materials and methods

Source of ESTs

Three barley strains were used to construct nine cDNA libraries. Haruna Nijo is a two-row malting cultivar grown in Japan. Akashinriki is a Japanese six-row hull-less landrace used mainly for human consumption. Both are cultivated types classified as *Hordeum vulgare* ssp. *vulgare*. H602 is an accession of *Hordeum vulgare* ssp. *spontaneum*, the wild progenitor of cultivated barley. The two subspecies are classified as *Hordeum vulgare* L. after Bothmer *et al.* (2003) as no reproductive barriers or major genome differentiation are observed. The nine cDNA libraries were constructed using standard procedures (see <http://harvest.ucr.edu/>) or by oligo capping (Maruyama and Sugano, 1994). Cycle sequencing from both the 3' and 5' ends of the cDNA clones was carried out on ABI3700 or ABI3100 DNA sequencers using BigDye-terminator (Applied Biosystems Japan Ltd, Tokyo, Japan). Each read was base-called using Phred (Ewing and Green, 1998). After trimming of vector sequence and Poly A/ Poly T stretches, high-quality sequences (quality value (QV) ≥ 20) were selected from each read. These sequences were used for clustering and multiple alignment.

Primer development

Approximately 60 000 3'ESTs were assembled by Phrap (<http://www.phrap.org/>) and two non-redundant sets of sequences were derived (8753 contigs and 6686 singlets). Primers for each singlet and for the longest high-quality sequence in each contig were designed using Primer 3 (Rozen and Skaletsky, 2000). Criteria for developing primer sets were $T_m = 60^\circ\text{C}$ (with a difference of $< 2^\circ\text{C}$ between forward and reverse primers), GC content = 50% (with a difference $< 20\%$), target size = 400 bp (actual range 150–500 bp). Sequences that did not satisfy these conditions were rejected. A total of 10 366 primer sets were developed.

Mapping population development

A cross was made between Haruna Nijo and H602. The DH mapping population of 93 lines was derived from the F_1 using microspore culture by Pajbjergfonden, Denmark (Ziauddin *et al.*, 1990).

Polymorphism detection and mapping

Polymorphisms were identified using genomic DNA of Haruna Nijo, H602, and their F_1 . The PCR amplicons were electrophoresed on 1.5% agarose gels. Primer sets that did not amplify or that generated complex banding patterns were rejected. Presence or absence (dominant) and size difference (co-dominant) of polymorphisms were used directly as STS markers. All PCR amplicons were directly cycle-sequenced and aligned using CLUSTALW (Thompson *et al.*, 1994). Restriction site polymorphisms for 50 different enzymes were used to develop CAPS markers. Where neither STS nor CAPS assays were possible, SNPs were genotyped directly by fluorescence polarization (Greene *et al.*, 2002) using a 384-plate format and an ARVOsx multilabel counter (PerkinElmer Japan Co., Ltd, Yokohama, Japan). SNP alleles were scored automatically using the software provided by the manufacturer. Polymorphic SSR and STS markers with previously-reported map positions (Mano *et al.*, 1999; Ramsay *et al.*, 2000) were used as anchor markers. Linkage between these anchor markers and EST-derived STS and CAPS markers was calculated using MAPMAKER/EXP (Lander and Botstein, 1989) with the Kosambi map function (Kosambi, 1944) and a LOD threshold of 5.0. After the development of seven linkage groups, SNP-based markers were sequentially added to each linkage group at the same LOD threshold.

Sequence similarity with other genomic resources

We determined the sequence similarity of the mapped EST sequences with the unigene probe sets of the Affymetrix Barley 1 GeneChip (HarVEST Assembly 25), published full-length cDNA sequences (<http://www.shigen.nig.ac.jp/barley/>) and the rice genome (RAP2 rice representative sequences selected in each locus, <http://rapdb.dna.affrc.go.jp/>) by BlastN (Altschul *et al.*, 1990) with a threshold of $E < 1E-30$.

Results

Degree of polymorphism

Polymorphism rates for each class of marker are shown in Figure 1 and complete marker information on each locus is available (Supplementary Table 1). Of the 10 366

primer sets that were developed, 7700 generated useful amplicons. Of the 7700 primer sets, 3975 (52%) showed polymorphisms between the mapping parents. Of these, 2890 (28% of the total) were mapped (Supplementary Table 2). Nearly half of these (1717) were mapped as SNPs; the next largest group (933) were mapped using CAPS and 240 showed INDELs (196 co-dominant and 44 dominant) that could be mapped directly using fragment mobility in agarose gels. Transition SNPs were more common than transversion SNPs (A/G: 34.6%, C/T: 25.0%). The decision to map by CAPS or fluorescence polarization was based on the relative cost of the restriction enzyme.

Mapping

Fifty-eight previously-mapped markers (40 SSR and 18 STS) were used as anchor markers to define the short and long arms of each chromosome linkage group. Addition of the 2890 ESTs created a map with seven linkage groups, with a total map length of 2136 cM (Table 1, Figure 2 and online at <http://map.lab.nig.ac.jp:8085/cmap/>) (mapped 3'EST sequences are available from Supplementary Data 1). Each linkage group is densely

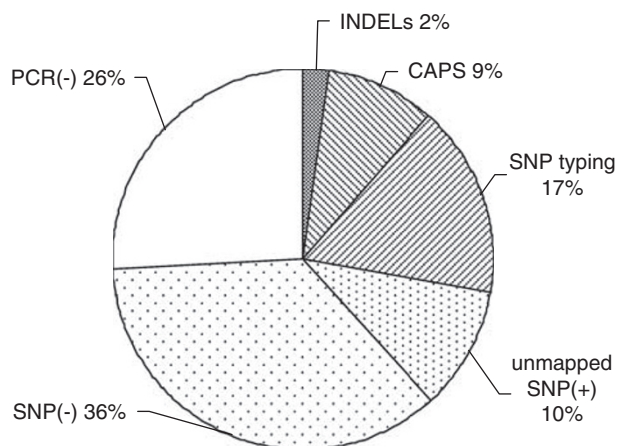


Figure 1 Categories of polymorphism for the PCR amplicons between mapping parents Haruna Nijo and H602 by the primer sets developed from the non-redundant 3'ESTs in Okayama University. PCR(-) includes both PCR failure and multiplex amplification.

populated with EST markers (a minimum of 321 on chromosome 6H and a maximum of 498 on chromosome 5H). Similar numbers of ESTs were mapped to each chromosome by the three mapping strategies (Table 1). As all EST-based markers are based on a non-redundant set of 3'EST sequences, the many clusters of co-segregating markers are attributed to the small size of the mapping population and not to 're-mapping' the same locus using different primer sets. The average percentage of missing data points per locus was 4.7%, but in some cases was as high as 43%, which introduces some ambiguity into the mapping process (Supplementary Table 1). Clusters of markers showing segregation distortion were observed on all chromosomes except 3H. The number of missing data points per locus, as well as the segregation values for each locus, are shown in Supplementary Table 1. Replacing missing data points will slightly improve map quality, whereas the unidirectional segregation distortion of alleles at each cluster suggests a biological basis, perhaps related to gametophytic selection for performance in microspore culture, which tends to produce more lethal haploids compared with the haploid induction by *Hordeum bulbosum* (Chen and Hayes, 1989). Sayed *et al.* (2002) and Muñoz-Amatriain *et al.* (2008) reported distorted marker segregation in anther culture-derived mapping populations. Overall, neither missing data nor segregation distortion affected map quality, as the LOD threshold was 5.0.

Sequence similarity

Statistics on the sequence similarity of the ESTs with the Affymetrix Barley 1 GeneChip, published full-length barley cDNA sequences, and the *Triticum aestivum* (Ta) gene index release 10 are shown in Table 2 and full data are available in Supplementary Table 3. Of the mapped ESTs, 2689 (93%) are formatted in Affymetrix Barley 1 and full-length cDNA sequences are available for 1039 (36%). Mapped ESTs show higher similarity with sequences in the Ta gene index (93%) and moderate similarity with rice locus (50%) (Supplementary Table 4).

Discussion

The present work establishes linkage map positions for an estimated 9% of the genes of barley. This estimate

Table 1 Length and marker information on the high-density barley EST map

Chromosome	Length (cM)	Marker type					Total no. of markers	
		EST (Okayama Univ)				Others		
		Length polymorphism		CAPS	SNPs			
		Codominant	Dominant			SSR		STS
1H	363	28	4	126	227	3	0	388
2H	360	28	6	154	304	4	4	500
3H	300	32	10	131	271	11	1	456
4H	259	20	6	117	198	5	2	348
5H	332	31	8	152	307	5	3	506
6H	207	17	5	126	173	3	3	327
7H	315	40	5	127	237	9	5	423
Total	2136	196	44	933	1717	40	18	2948

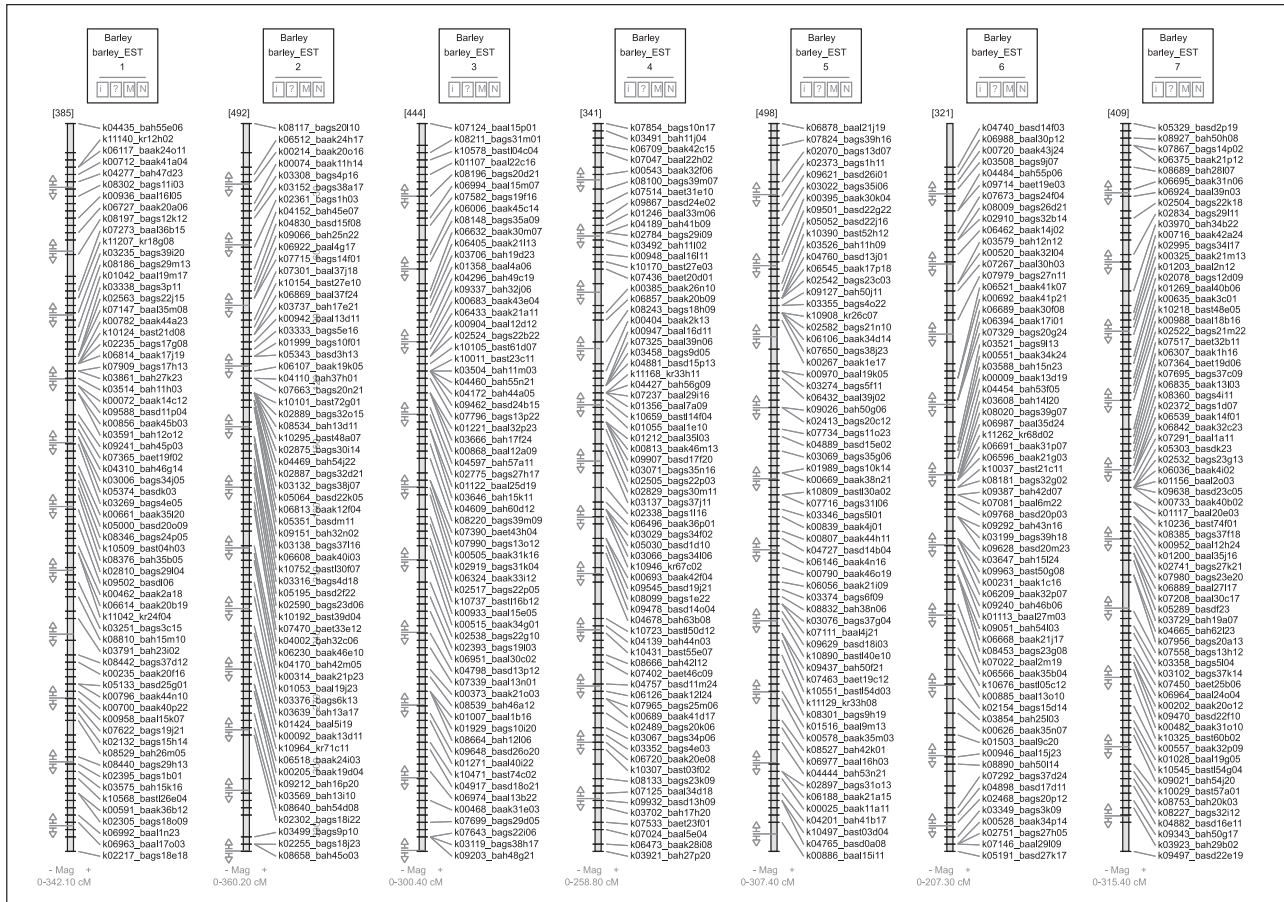


Figure 2 A CMAP of barley transcript (EST) map in the cross between Haruna Nijo and H602. SSR and STS markers were integrated as anchors for each chromosome. All the EST markers were PCR amplified by specifically developed primer pairs from the representative 3'ESTs, which were generated at Okayama University and directly accessible from the electronic version of cMAP with full EST markers (<http://map.lab.nig.ac.jp:8085/cmap/>).

Table 2 Number of accessions in the probes in Affymetrix Barley 1 GeneChip, published barley full length cDNA sequences (FLcDNA), *Triticum aestivum* gene index release 10 (TaGI10) and RAP2 rice nucleotide sequence of each locus (RAP2 nucleotide locus) showing sequence similarity by blastN search with ESTs on the map

Chromosome	GeneChip	RAP2 nucleotide locus				
		FLcDNA	TaGI10	e-30	e-3	e-5
1H	361	140	379	198	307	
2H	459	170	469	245	373	
3H	416	160	417	215	348	
4H	316	123	324	175	256	
5H	459	176	468	261	388	
6H	302	114	279	174	246	
7H	376	156	354	171	301	
Total	2689	1039	2690	1439	2219	
(%)	93	36	93	50	77	

^aThreshold for blastN search.

is based on the number of non-redundant barley transcripts (32690) by CAP3 (Huang and Madan, 1999) assembly using 3' end sequences of full-length cDNAs (Sato K, unpublished). These mapped genes, together with a systematically developed set of mapping popula-

tions, will make it easier to directly clone barley genes conferring phenotypes showing simple inheritance and to determine the genetic basis of complex traits. The information will contribute to the development of a framework for the physical mapping of barley, which is a necessary step towards genome sequencing. A high-density barley EST map will also be useful for all cereal species by way of providing additional markers and elucidating evolutionary relationships.

Contribution to the systematic development of genetic resources

Haruna Nijo and H602 are key hubs in a comprehensive set of genetic and germplasm resources. Hori *et al.* (2005) developed and used recombinant chromosome substitution lines (RCSLs) derived from these two accessions. Haruna Nijo is the recurrent parent in each of the 10 new RCSL populations developed at Okayama University (Sato, unpublished). The donor accessions for each of these new RCSL populations represent diverse key barley germplasm, including Morex, Barke, Harrington, Oregon Wolf Barley (dominant parent) and Golden Promise. These lines will facilitate access to qualitative and quantitative traits segregating between donor parents and Haruna Nijo. BAC libraries have been

constructed from both Haruna Nijo (Saisho *et al.*, 2007) and H602 (Sato, unpublished). Haruna Nijo is the basis of a new TILLING system comprised of 10 000 individuals (Sato, unpublished). Haruna Nijo and H602 are the sources of many of the ESTs (<http://www.shigen.nig.ac.jp/barley/>) that are foundation stones of contemporary barley genetic resources, including HarVEST (<http://harvest.ucr.edu/>), the US Barley Coordinated Agricultural Project (CAP) (<http://www.barleycap.org/>; Hayes and Szucs, 2006) and the Affymetrix Barley 1 GeneChip (Close *et al.*, 2004). Haruna Nijo was used to develop the first comprehensive set of barley full-length cDNAs (5006 clones) (<http://www.shigen.nig.ac.jp/barley/>). The EST map described in this report will make Haruna Nijo one of the key haplotypes for barley research.

The cross-combination of a cultivated and a wild-type barley was chosen to allow mapping genes related to domestication, to facilitate future mining of new alleles from the wild-type parent and to maximize polymorphisms at marker loci. Domestication-related genes mapped to date include brittle rachis (*btr1/btr2*), dormancy, black lemma (*Blp*) and cleistogamy (*cyl1/Cyl2*) (Hori *et al.*, 2005). H602 is a source of potentially new genes conferring resistance to fungal pathogens, including those inciting mildew (*Blumeria graminis* f. sp. *hordei*) and net blotch (*Pyrenophora teres* f. *teres*) (Sato and Takeda, 1997). The present map, together with the other genomics resources based on this germplasm (notably RCSLs and BAC libraries), can be used as a platform for cloning these genes.

Apart from the relatively small number of ESTs that we mapped directly based on INDELS, we mapped ESTs based on SNPs in 3'UTRs, exons and introns. We expected that our strategy of using 3'ESTs would provide a higher level of polymorphism than the one based on exon SNPs only (for example, Rostoks *et al.*, 2006). This expectation was based on the report of higher SNP rates in non-coding regions (for example, UTRs, introns and intergenic regions) than in exons in humans (Venter *et al.*, 2001) and soybean (Van *et al.*, 2005). Our 3'ESTs would be expected to include exons, 3'UTRs and introns. We also expected that there could be a somewhat higher rate of exon SNPs in EST sequences from wild-type and cultivated barley than in those from only cultivated barley because of mutations associated with domestication. However, our SNP rate was not as high as expected compared with the rates of 17–45% reported for three mapping populations by Stein *et al.* (2007). This difference may be due to the higher detection rate of intron SNPs by Stein *et al.* (2007) and the higher expected similarity in coding sequences between these accessions of wild-type and cultivated barley.

The present map was developed using segregation data from 93 DH lines, a small population for mapping 2890 ESTs. Although local ordering would be optimized, and some co-segregating markers separated, with a larger population, the current map provides reliable general positions for a large number of ESTs. Our strategy of using a single population differs from the consensus map approach used by Stein *et al.* (2007) and Close *et al.* (unpublished Barley CAP SNP consensus map; <http://www.harvest-web.org/hweb/bin/gmap.wc?wsize=1195x790>). Ultimately, the individual and

consensus maps should be available at a single site so that the user can view, align and compare depending on the need and interest.

The EST map as a tool for map-based cloning and dissection of complex traits

A high-density EST map will be useful for cloning genes when the mapped loci are also represented on an expression profiling array such as the Affymetrix Barley 1 GeneChip (Affymetrix Japan Co., Tokyo, Japan). As shown in Table 2, 2689 probes on the Barley 1 GeneChip are assigned genetic map positions. Addition of the 1032 ESTs reported by Stein *et al.* (2007) will further integrate the Barley 1 GeneChip with the genetic map. Furukawa *et al.* (2007) have shown that expressed probes on the Barley 1 GeneChip with genetic map positions in barley can be used to identify orthologous genes in rice. This strategy led to identifying the corresponding full-length cDNA in barley, and ultimately cloning and characterizing the aluminium-tolerance gene in barley. The availability of full-length cDNAs for 36% of the mapped ESTs will make this process even more efficient in future projects.

The present map will also be useful for mapping complex traits. Any marker can be used for mapping quantitative trait locus (QTL), but a key advantage of ESTs is that they are based on a sequence, so that the information content will not change even if the detection platform changes. By providing aligned sequence data on the two parents used in this study (Supplementary Table 5), we provide a starting point for aligning alleles from multiple germplasm accessions and a resource for designing markers for other germplasm combinations. This approach was shown by Hori *et al.* (2007a), who synchronized the positions of QTLs in eight mapping populations by using the same EST markers, all of which were extracted from the present EST map.

The EST map as a framework for physical mapping of barley

The mapped ESTs can be assigned physical coordinates using cytogenetic stocks, such as barley–wheat addition lines. For example, Nasuda *et al.* (2005) and Ashida *et al.* (2007) mapped barley EST markers on the barley chromosome deletion stocks and estimated the physical location of these ESTs. Although the resolution in segmental allocation of markers was low in wheat deletion stock mapping (Sorrells *et al.*, 2003), the approach is a useful first step in developing BAC contigs for genome sequencing. Best estimates of physical and genetic distances in the barley genome are necessary to determine the efficiency of using a genetic map as a framework for BAC clone allocation to the genome, as marker resolution is usually highest in distal regions and lowest in centromeric regions (Künzel *et al.*, 2000).

The EST map provides an integrative tool for all cereal species

As clearly shown by Devos (2005); and Moore *et al.* (1995), there is general colinearity in content of the barley and wheat genomes. Therefore, it is not surprising that 93% of wheat genes showed a high level of sequence similarity with the mapped barley ESTs (Table 2).

Together with the mapped ESTs reported by Stein *et al.* (2007) and those on HarvEST (<http://harvest.ucr.edu/>), ca. 5000 wheat genes may show collinearity with

mapped barley ESTs. This should be a significant marker resource for wheat genetic map enrichment.

As shown by Hori *et al.* (2007b), barley ESTs can be efficiently mapped on the diploid wheat genome. Homoeology is complete except for the reciprocal translocation between 4A and 5A (Figure 3, see online at <http://map.lab.nig.ac.jp:8085/cmap/>). There are examples for complementary use of wheat, and barley genetic resources can be of benefit to gene discovery for vernalization-requirement genes in both species (Yan *et al.*, 2004, 2006) based on sequence similarity.

Barley: Rice synteny is an important resource to identify barley orthologs by rice genome information (Moore *et al.*, 1995; Devos, 2005). The sequence similarity of mapped barley ESTs with rice nucleotide sequences was much lower (50%) than observed for the wheat gene index (93%). The overall genome comparison between mapped barley ESTs and rice genome sequences (IRGSP pseudomolecules Build4 (Supplementary Table 4)) identified many collinear regions between the two genomes. In particular, the entire chromosome-level collinearity was observed between barley chromosome 3H and rice chromosome 1 (Figure 4). Comparative genomic information can aid systematic marker generation even when the corresponding interval in the rice genome lacks a barley ortholog (Komatsuda *et al.*, 2007; Sutton *et al.*, 2007). This approach is particularly useful for the identification of functional homologs in both species (Inukai *et al.*, 2006) and identification of candidate orthologs in barley. This has been shown in the isolation of genes controlling key agronomic characters in barley (Furukawa *et al.*, 2007; Taketa *et al.*, 2008). Even after the release of rice genome (International Rice Genome Sequencing Project, 2005), some rice transcripts (http://rapdb.dna.affrc.go.jp/RAP2_statistics.html) were not assigned map positions: barley ESTs were used to identify rice transcript positions in the rice genome. The reverse will also be true. Rice sequence information will be useful for filling missing sequences in the barley genome.

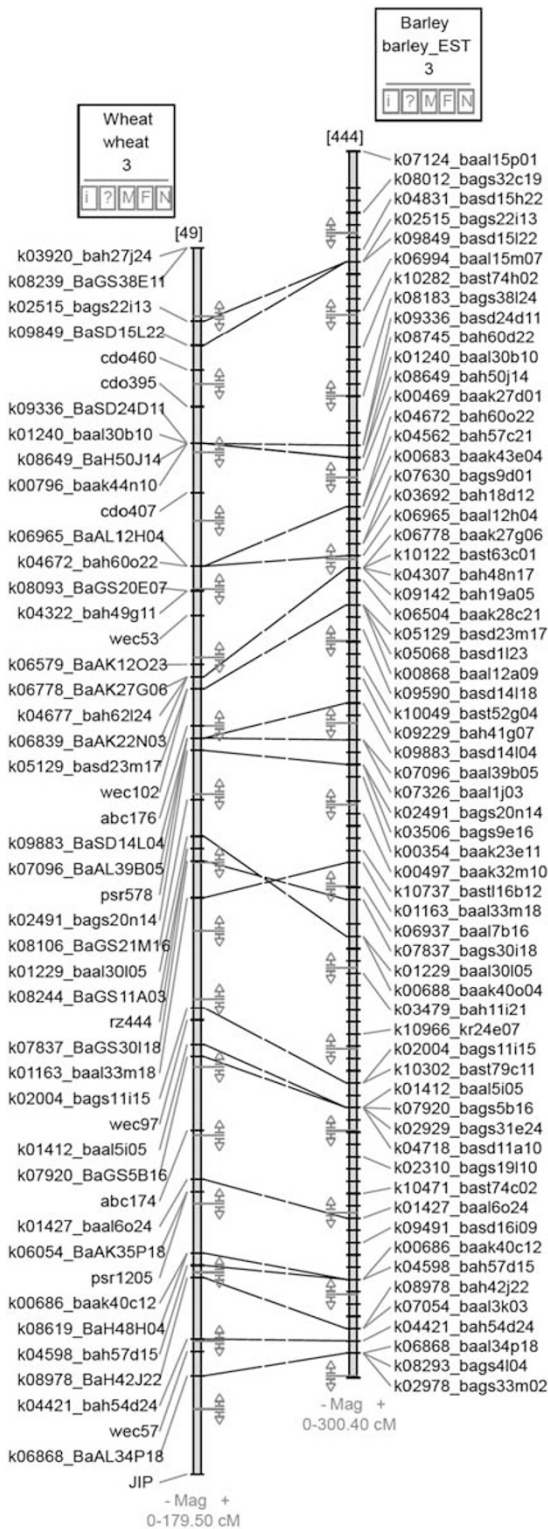


Figure 3 Comparison of barley EST markers between barley chromosome 3H and diploid wheat 3A (Hori *et al.*, 2007a,b). A total chromosome map information and EST marker sequences are accessible from the electronic version of cMAP (<http://map.lab.nig.ac.jp:8085/cmap/>).

Perspectives on EST mapping strategies

There are ~ 500 000 barley ESTs in GenBank. These were the essential basis for detecting SNPs among barley EST

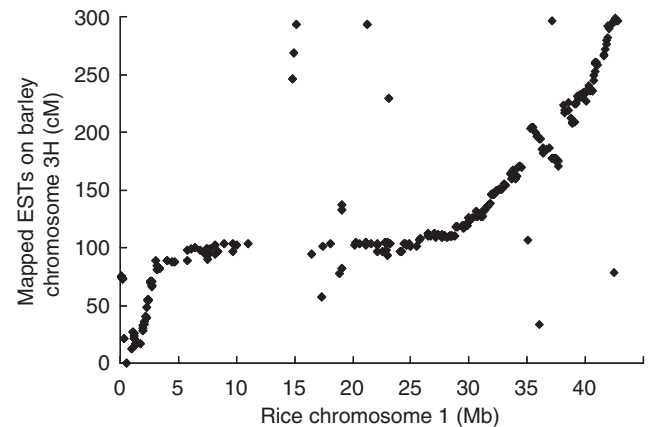


Figure 4 Complete chromosome level collinearity between mapped ESTs on barley chromosome 3H and the genome sequence of rice chromosome 1 (from IRGSP pseudomolecules Build4). A blastN (Altschul *et al.*, 1990) threshold of $E < 1E-5$ was used for the basis of sequence similarity comparison.

donors and efficient genotyping using the Oligonucleotide Pooled Assay (OPA) mapping system (Rostoks *et al.*, 2006). However, the system required redundant sets of ESTs, and the genes that are mapped may be biased to the same highly expressed genes in multiple genotypes. The present map comprised of 2890 ESTs took 3 years to generate, which is slow compared with the OPA mapping system. However, the present map was effective for mapping transcripts with lower-expression levels and may therefore represent a complementary set of mapped transcripts.

Redundancy and assay polymorphism rates will always be challenges for EST mapping projects. If we were to start EST mapping today, we would use full-length cDNAs to minimize redundancy. Polymorphism detection would also be improved with more available sequence. Polymorphism assay platforms are constantly improving in terms of cost, accuracy and throughput.

Acknowledgements

We thank Dr Patrick M Hayes (Oregon State University, USA) for his critical reading of the manuscript, Drs Hans Harudrup and Heidi Jaiser (Pajerbier Foundation, Denmark) for producing the doubled haploid population, and Dr Yukiko Yamazaki (National Institute of Genetics, Japan) for developing the cMAP resource. This project was supported by grants from Core Research for Evolutional Science and Technology and the Program of Promotion of Basic Research Activities for Innovative Biosciences.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Arumuganathan K, Earle ED (1991). Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* **9**: 208–218.
- Ashida T, Nasuda S, Sato K, Endo TR (2007). Dissection of barley chromosome 5H in common wheat. *Genes Genet Syst* **82**: 123–133.
- Bothmer R von, Sato K, Komatsuda T, Yasuda S, Fischbeck G (2003). The domestication of cultivated barley. In: Bothmer R von, Hintum ThJL van, Knüpfner H, Sato K (eds). *Diversity in Barley (Hordeum vulgare)*. Elsevier Science B.V., Amsterdam, The Netherlands, pp 9–27.
- Chen FQ, Hayes PM (1989). A comparison of *Hordeum bulbosum*-mediated haploid production efficiency in barley using invitro floret and tiller culture. *Theor Appl Genet* **77**: 701–704.
- Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B, Fang G *et al.* (2002). An integrated physical and genetic map of the rice genome. *Plant Cell* **14**: 537–545.
- Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA *et al.* (2004). A new resource for cereal genomics: 22 K barley GeneChip comes of age. *Plant Physiol* **134**: 960–968.
- Devos KM (2005). Updating the 'crop circle'. *Curr Opin Plant Biol* **8**: 155–162.
- Druka A, Muehlbauer G, Druka I, Caldo R, Baumann U, Rostoks N *et al.* (2006). An atlas of gene expression from seed to seed through barley development. *Funct Integr Genomics* **6**: 202–211.
- Ewing B, Green P (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**: 186–194.
- Furukawa J, Yamaji N, Wang H, Mitani N, Murata Y, Sato K *et al.* (2007). An aluminum-activated citrate transporter in barley. *Plant Cell Physiol* **48**: 1081–1091.
- Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G *et al.* (1991). Construction of an RFLP map of barley. *Theor Appl Genet* **83**: 250–256.
- Greene RA, DiMeo JJ, Malone ME, Swartwout S, Liu J, Buzby PR (2002). AcycloPrime, a novel method for SNP analysis using fluorescence polarization. *Proc SPIE* **4626**: 332–339.
- Hayes P, Szucs P (2006). Disequilibrium and association in barley: thinking outside the glass. *Proc Natl Acad Sci USA* **103**: 18385–18386.
- Hayes PM, Liu BH, Knapp SJ, Chen F, Jones B, Blake T *et al.* (1993). Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. *Theor Appl Genet* **87**: 392–401.
- Heun M, Kennedy A, Anderson J, Lapitan N, Sorrels M, Tanksley S (1991). Construction of restriction fragment length polymorphism map of barley (*Hordeum vulgare*). *Genome* **34**: 437–447.
- Hori K, Kobayashi T, Shimizu A, Sato K, Takeda K, Kawasaki S (2003). Efficient construction of high-density linkage map and its application to QTL analysis in barley. *Theor Appl Genet* **107**: 806–813.
- Hori K, Sato K, Nankaku N, Takeda K (2005). QTL analysis in recombinant chromosome substitution lines and doubled haploid lines derived from a cross between *Hordeum vulgare* ssp. *vulgare* and *Hordeum vulgare* ssp. *spontaneum*. *Molecular Breed* **16**: 295–311.
- Hori K, Sato K, Takeda K (2007a). Detection of seed dormancy QTL in multiple mapping populations derived from crosses involving novel barley germplasm. *Theor Appl Genet* **115**: 869–876.
- Hori K, Takehara S, Nankaku N, Sato K, T S, Takeda K (2007b). Linkage map construction and QTL detection based on barley ESTs in A genome diploid wheat. *Breeding Sci* **57**: 39–45.
- Huang X, Madan A (1999). CAP3: A DNA sequence assembly program. *Genome Res* **9**: 868–877.
- International rice genome sequencing project (2005). The map-based sequence of the rice genome. *Nature* **436**: 793–800.
- Inukai T, Vales MI, Hori K, Sato K, Hayes PM (2006). RMo1 confers blast resistance in barley and is located within the complex of resistance genes containing Mla, a powdery mildew resistance gene. *Mol Plant Microbe Interact* **19**: 1034–1041.
- Kleinohfs A, Kilian A, Saghai Maroof M, Biyashev R, Hayes P, Chen F *et al.* (1993). A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor Appl Genet* **86**: 705–713.
- Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D *et al.* (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc Natl Acad Sci USA* **104**: 1424–1429.
- Kosambi D (1944). The estimation of map distances from recombination values. *Ann Eugen* **12**: 172–175.
- Künzel G, Korzun L, Meister A (2000). Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* **154**: 397–412.
- Lander ES, Botstein D (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.
- Lundqvist U, Franckowiak J, Konishi T (1996). New and revised descriptions of barley genes. *Barley Genet Newsl* **26**: 22–43.
- Mano Y, Sayed-Tabatabaei BE, Graner A, Blake T, Takaiwa F, Oka S *et al.* (1999). Map construction of sequence-tagged sites (STSs) in barley (*Hordeum vulgare* L.). *Theor Appl Genet* **98**: 937–946.

- Maruyama D, Sugano S (1994). Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**: 171–174.
- Moore G, Devos KM, Wang Z, Gale MD (1995). Cereal genome evolution. Grasses, line up and form a circle. *Curr Biol* **5**: 737–739.
- Muñoz-Amatriain M, Castillo AM, Chen XW, Cistue L, Valles MP (2008). Identification and validation of QTLs for green plant percentage in barley (*Hordeum vulgare* L.) anther culture. *Mol Breed* **22**: 119–129.
- Nasuda S, Kikkawa Y, Ashida T, Islam AK, Sato K, Endo TR (2005). Chromosomal assignment and deletion mapping of barley EST markers. *Genes Genet Syst* **80**: 357–366.
- Ramsay L, Macaulay M, degli Ivanissevich S, MacLean K, Cardle L, Fuller J *et al.* (2000). A simple sequence repeat-based linkage map of barley. *Genetics* **156**: 1997–2005.
- Rostoks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML *et al.* (2006). Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc Natl Acad Sci USA* **103**: 18656–18661.
- Rozen S, Skaletsky H (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365–386.
- Saisho D, Myoraku E, Kawasaki S, Sato K, Takeda K (2007). Construction and characterization of a bacterial artificial chromosome (BAC) library for Japanese malting barley 'Haruna Nijo'. *Breed Sci* **57**: 29–38.
- Sato K, Bothmer R von, Hintum ThJL van, Knüpfper H (2003). Barley diversity, an outlook. In: Bothmer R von, Hintum ThJL van, Knüpfper H, Sato K (eds). *Diversity in Barley (Hordeum vulgare)*. Elsevier Science B.V.: Amsterdam, The Netherlands, pp 269–278.
- Sato K, Takeda K (1997). Net blotch resistance in the wild species of *Hordeum*. *Euphytica* **95**: 179–185.
- Sayed H, Kayyal H, Ramsay L, Ceccarelli S, Baum M (2002). Segregation distortion in doubled haploid lines of barley (*Hordeum vulgare* L.) detected by simple sequence repeat (SSR) markers. *Euphytica* **225**: 265–272.
- Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD *et al.* (2003). Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res* **13**: 1818–1827.
- Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M *et al.* (2007). A 1000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor Appl Genet* **114**: 823–839.
- Sutton T, Baumann U, Hayes J, Collins NC, Shi BJ, Schnurbusch T *et al.* (2007). Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* **318**: 1446–1449.
- Takahashi H, Akagi H, Mori K, Sato K, Takeda K (2006). Genomic distribution of MITEs in barley determined by MITE-AFLP mapping. *Genome* **49**: 1616–1620.
- Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda 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 U S A* **105**: 4062–4067.
- Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Van K, Hwang EY, Kim MY, Park HJ, Lee SH, Cregan PB (2005). Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. *J Hered* **96**: 529–535.
- Varshney RK, Grosse I, Hahnel U, Siefken R, Prasad M, Stein N *et al.* (2006). Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. *Theor Appl Genet* **113**: 239–250.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG *et al.* (2001). The sequence of the human genome. *Science* **291**: 1304–1351.
- Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P *et al.* (1997). Homology of AFLP products in three mapping populations of barley. *Mol Gen Genet* **255**: 311–321.
- Wu J, Maehara T, Shimokawa T, Yamamoto S, Harada C, Takazaki Y *et al.* (2002). A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* **14**: 525–535.
- Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M *et al.* (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc Natl Acad Sci USA* **103**: 19581–19586.
- Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P *et al.* (2004). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* **303**: 1640–1644.
- Ziauddin A, Simion E, Kasha KJ (1990). Improved plant regeneration from shed microspore culture in barley (*Hordeum vulgare* L.) cv. Igri. *Plant Cell Rep* **9**: 69–72.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)