Theor Appl Genet (2009) 119:613–619
DOI 10.1007/s00122-009-1071-9

ORIGINAL PAPER

An application of high-throughput SNP genotyping for barley genome mapping and characterization of recombinant chromosome substitution lines

Kazuhiro Sato · Kazuyoshi Takeda

Received: 21 August 2008 / Accepted: 13 May 2009 / Published online: 2 June 2009
© Springer-Verlag 2009

Abstract An oligo-nucleotide pooled assay (OPA) for high-throughput single nucleotide polymorphism (SNP) genotyping was used for genetic map development in order to coordinate marker information from multiple mapping resources in barley. A doubled haploid (DH) population derived from the cross between barley cultivar "Haruna Nijo" (*Hordeum vulgare* ssp. *vulgare*) and wild barley strain "H602" (*H. vulgare* ssp. *spontaneum*) was genotyped with 1,448 unigene-derived OPA-SNPs. Of these, 732 markers showed polymorphisms and 384 were cross-referenced with EST markers on our high-density transcript map. The OPA-SNP markers were well distributed on barley chromosomes as follows: 1H (93), 2H (131), 3H (123), 4H (97), 5H (108), 6H (92) and 7H (88). Using a cMAP platform, it was possible to integrate EST marker positions across high-density EST maps. The OPA-SNPs were used to genotype 99 BC_3F_5 recombinant chromosome substitution lines (RCSLs) from the same cross (Haruna Nijo/H602). These data were used to create graphical genotypes for each line and thus estimate the location, extent, and total number of introgressions from the wild barley parent. The RCSLs sampled most of the wild barley genome, with only a few missing segments. With the resources we have developed, all QTL alleles segregating in this germplasm are now potential targets for map-based cloning.

Communicated by A. Graner.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1071-9) contains supplementary material, which is available to authorized users.

K. Sato (✉) · K. Takeda
Research Institute for Bioresources,
Okayama University, Chuo, Kurashiki 710-0046, Japan
e-mail: kazsato@rib.okayama-u.ac.jp

Introduction

Expressed sequence tags (ESTs) are the most reliable and flexible resource for developing genetic markers. Single nucleotide polymorphisms (SNPs) in EST sequences have been widely used because of the high polymorphism frequency and adaptability to parallel allele detection systems. A significant number of ESTs (ca. 525K in February 2009) have been generated in barley and quality-controlled EST information is available in the HarvEST database (http://harvest.ucr.edu/). Because multiple research groups used different accessions (haplotypes) for generating these EST resources, HarvEST has functions for EST alignment, SNP identification, and SNP mapping. Three barley transcript maps, based on SNP mapping strategies, are published (Rostoks et al. 2005; Stein et al. 2007; Sato et al. 2009).

Different techniques for SNP allele discrimination have been proposed and applied in barley. The most popular technique is the cleaved amplified polymorphic sequence (CAPS) assay. An advantage of this approach is that it uses standard agarose gel electrophoresis (Sato et al. 2009). Stein et al. (2007) and Kota et al. (2008) used denatured high-performance liquid chromatography (dHPLC). Sato et al. (2009) applied fluorescently labeled acyclo terminators for genotyping SNPs and this technique allowed detection of most SNPs in the EST collection. SNPs can also be genotyped by multiplex PCR reactions, if the detection system has the sensitivity to differentiate target SNP-bearing amplicons. Massive multiplex SNP detection was first achieved by the Affymetrix GeneChip system in 1990s. Compared to this pioneering system based on a solid wafer, the more recent liquid phase bead detection systems provide greater flexibility for customizing target SNP detection. The Illumina oligonucleotide pooled assay (OPA) platform was developed for human SNP detection but it is

widely applicable to other species (Fan et al. 2003). Rostoks et al. (2006) used the OPA platform and the resources in HarvEST to develop a 1,536 SNP detection assay.

There are several key mapping populations that have been used extensively by the barley research community. The Steptoe/Morex doubled haploid (DH) population was initially mapped with RFLP markers (Kleinhofs et al. 1993) and subsequently used for marker saturation with other genotyping tools (Rostoks et al. 2005; Wenzl et al. 2006). The Oregon Wolf Barley (Costa et al. 2001) is a DH population that segregates at 12 visible mutant loci and has high sequence polymorphism between parents (Stein et al. 2007). We recently reported a high-density (2,890) transcript map based on the Haruna Nijo/H602 DH population (Sato et al. 2009).

It is important to have EST markers in common among these populations in order to map a maximum number of ESTs on the barley genome. One solution is to identify common markers via the homology of mapped EST sequences. However, this approach does not give a genetically agreeable solution when marker orders in different maps are in conflict. An alternative to map merging is to map as many markers as possible in single populations based on actual segregation data and then to use an objective alignment procedure to integrate information across populations. The use of a fast, accurate, multiplex SNP typing systems is the key to such a strategy.

Hori et al. (2005) developed a set of BC_3F_1 recombinant chromosome substitution lines (RCSLs) from the cross of Haruna Nijo/H602. They used simple sequence repeat (SSR) markers and a subset of the EST markers represented on the high-density EST map (Sato et al. 2009). However, only 85 markers were mapped in these RCSLs, and this does not provide a precise estimate of the location, number, and extent of the substituted segments. The other barley RCSL resources reported in the literature (Matus et al. 2003; Pillen et al. 2003, 2004; von Korff et al. 2004, 2005) also suffer from this lack of marker density.

In this study, we used the highly efficient Illumina OPA SNP typing system to develop and integrate high-density map information in a single mapping population (Haruna Nijo/H602). We then applied this high-density EST-based map information to define wild barley introgression in RCSLs derived from the same cross.

Materials and methods

Plant materials

A Japanese malting barley cultivar "Haruna Nijo" (*Hordeum vulgare* ssp. *vulgare*) was used as a female parent for crossing with a wild barley strain "H602" (*H. vulgare* ssp. *spontaneum*). The DH plants were developed from the F_1 of Haruna Nijo/H602 using microspore culture by Pajbjergfonden, Denmark (Ziauddin et al. 1990). The 92 plants that naturally diploidized and set seed constituted the DH mapping population.

To develop RCSLs, the F_1 was crossed with a recurrent parent "Haruna Nijo" to produce BC_1F_1 individuals. Each of 23 BC_1F_1 individuals was crossed with the recurrent parent. Each of 134 BC_2F_1 individuals was again crossed with the recurrent parent to produce BC_3F_1 RCSLs. Each BC_3F_1 individual was self-pollinated to develop a set of 134 BC_3F_2 lines. A subset of 99 RCSLs was randomly selected from the 134 and advanced to the BC_3F_5 generation at the facilities of Okayama University, Kurashiki, Japan (34°35′N and 133°46′E).

DNA isolation

Plants were grown in the greenhouse in Okayama University and 200–300 mg of seedling leaf tissue was harvested from each plant. Leaf samples were frozen in liquid nitrogen and crushed into fine powder using a multi-bead shocker (Yasui Kikai Co). Qiagen DNeasy Plant mini kits (QIAGEN Co) were used to isolate DNA from each sample. DNA samples were eluted with TE, passing the same 60 µl of sample over each column twice. Concentration was determined by nanodrop (Thermo Fisher Scientific) and adjusted to 100 ng/µl. DNA samples were electrophoresed on agarose gels for quality assurance purposes. Liquid DNA samples were normalized in water or 1×TE to a standard concentration of 100–200 ng/µl.

Polymorphism detection and mapping

Frozen DNA samples were sent to the Southern California Genotyping Consortium, Illumina BeadLab at the University of California, Los Angeles (UCLA) for the OPA-SNP assay with the 1,536-plex detection platform of barley OPA 1 (BOPA1). OPA genotyping was performed on Haruna Nijo, H602, 92 DH lines and 99 RCSLs using the Illumina GoldenGate BeadArray (Fan et al. 2003). Genotype data were manually inspected to correct for excessive emphasis on heterozygote calls using GenCall software (Illumina, San Diego, CA) at the Close lab (University of California, Riverside) and only the most reliable calls were retained. The SNP loci are designated by HarvEST unigene assembly #32 numbers. Further information regarding translation of SNP locus designations to pilot OPA numbers (POPA), barley OPA numbers (BOPA) and HarvEST unigene assembly #35 numbers is provided in Supplemental Table 1.

A set of 384 core markers from the high-density EST marker map (Sato et al. 2009) was selected to provide

Table 1 Genetic map length (Kosambi cM), number of OPA-SNP markers, and number of core markers from the high density transcript map of Sato et al. (2009)

Chromosome	Length (cM)	Total no. of markers	OPA markers	SNP markers after Sato et al. (2009)	
				Core	Total
1H	173.0	143	93	50	385
2H	182.8	190	131	59	492
3H	154.5	182	123	59	444
4H	142.8	140	97	43	341
5H	198.0	179	108	71	498
6H	142.3	131	92	39	321
7H	194.0	151	88	63	409
Total	1,187.4	1,116	732	384	2,890

uniform genome coverage. Linkage between these framework EST markers and polymorphisms from the OPA analysis was calculated at Okayama University using JOINMAP ver. 3.0 (Kyazma B.V.) with the Kosambi map function (Kosambi 1944) and a LOD threshold of 5.0.

Genotype calls from the RCSLs were placed in map order as determined from the DH lines. The chromosome segments introgressed into Haruna Nijo from wild barley were estimated from the graphical haplotypes.

Consensus map development

The cMAP environment (Fang et al. 2003) was used to develop a comparative genetic map between the Haruna Nijo/H602 map reported herein and other map resources. Dynamic links from marker names to sequence and/or annotation data were made using cMAP functions.

Results

Degree of polymorphism

Of the 1,536 SNPs represented on the Illumina OPA, 1,448 SNPs gave useful information and 732 of these SNPs were polymorphic between Haruna Nijo and H602. These 732 polymorphisms between the parents were also informative with the 92 DH lines, giving an overall polymorphism rate of 51%.

Application of OPA markers

Allele calls for the 732 polymorphic SNPs were made in the DH lines based on the parental SNP alleles. These genotypic data were then added to the framework set of 384 markers derived from the high-density EST map (Sato et al. 2009) and used for linkage map construction. All 732 OPA markers were integrated into the seven barley chromosomes, as shown in Table 1. The number of markers mapped per chromosome (number OPA markers; number total markers) was as follows: 1H (93; 143), 2H (131; 190), 3H (123; 182), 4H (97; 140), 5H (108; 179), 6H (92; 131) and 7H (88; 151). The total map length was 1187.4 cM, with an average marker density of 1.06 cM. The OPA markers were well distributed, except for clustering in some centromeric regions (Fig. 1; and marker data available in Supplemental Table 1). There were no missing data from the OPA data set, but there were up to 12 missing data points per marker in the framework marker data set (Supplemental Table 1). Segregation distortion occurred in some regions (Supplemental Table 1). Clusters of co-segregating markers are assumed to have a biological basis and define areas with high physical:genetic distance ratios (e.g., centromeric and other heterochromatic regions).

cMAP development

Based on the universal map combining OPA markers with the framework set of 384 ESTs, links to the cMAPs of the high-density EST map (Sato et al. 2009) were made, as shown in Fig. 2.

RCSL genotyping and genome coverage

The OPA marker data from the 99 RCSLs were sorted in map order according to the data from the DH population. This was successful for 723 loci; nine loci had high numbers of missing data points in the RCSLs and were not included. The RCSLs were homozygous at most loci; the few cases of heterozygote calls were treated as missing data. Within families of lines derived from the same BC_2F_1 plants, there was considerable overlap in introgressed segments. But even in these cases, unique introgression segments were identified.

By arranging substituted segments in the RCSLs by chromosome (and from short arm to long arm within each chromosome), a minimum set of RCSLs was selected to represent the unique substituted segments from wild barley.

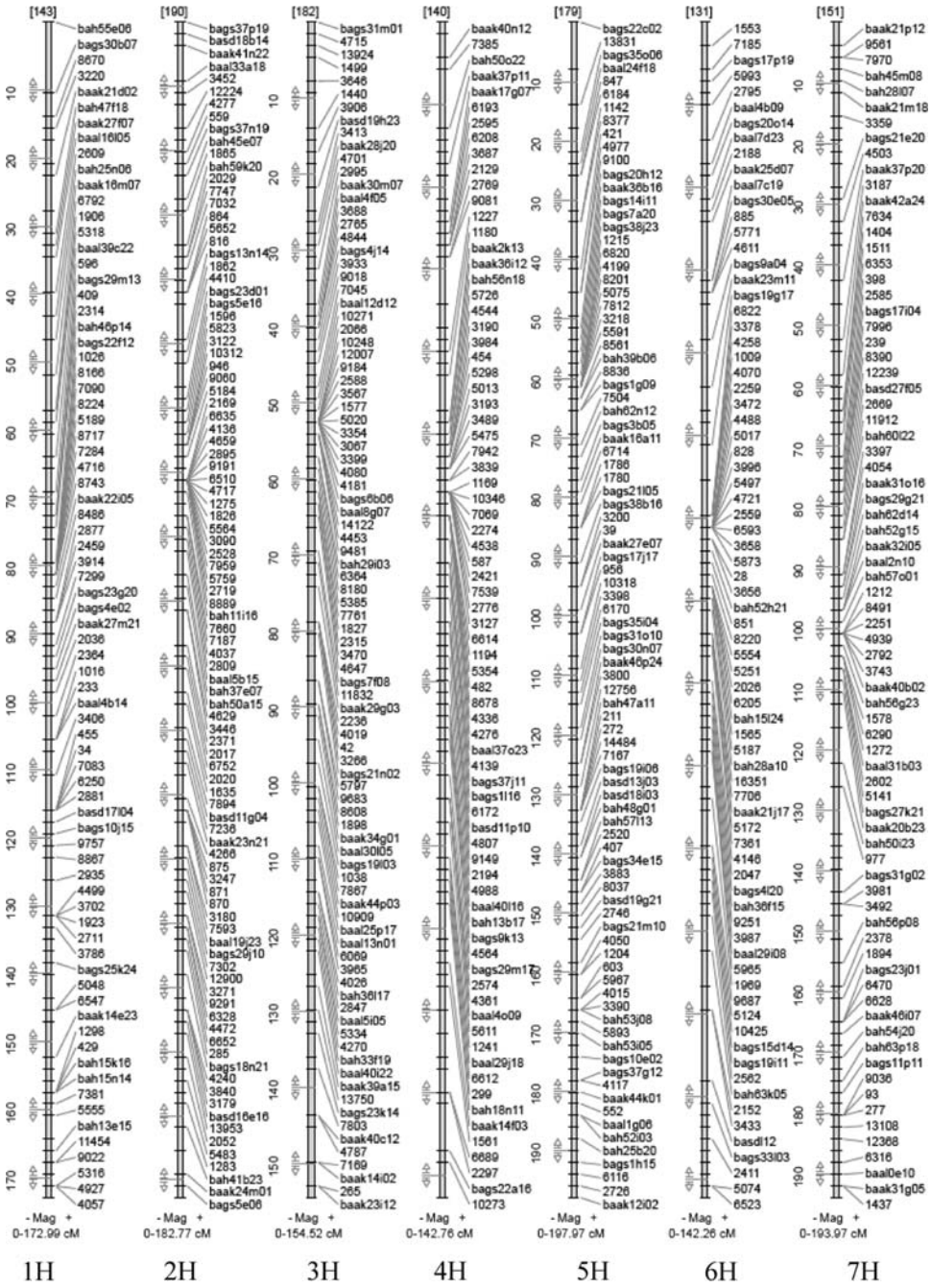


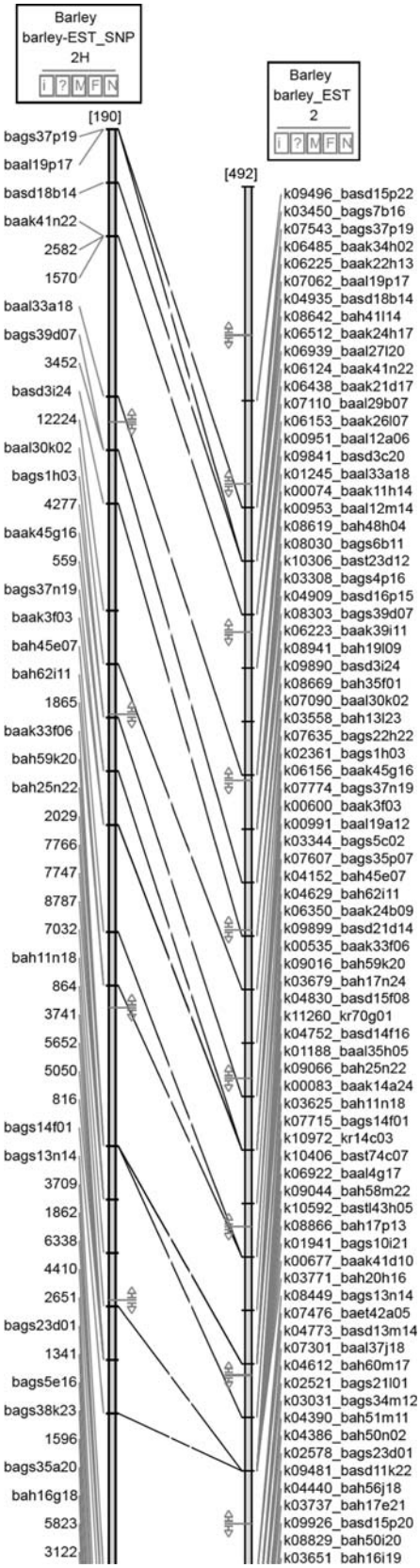
Fig. 1 A concise presentation of the Haruna Nijo/H602 linkage map comprised 732 OPA-SNP markers and 384 core markers from the 2,890 EST marker map of Sato et al. (2009). Full map information is directly accessible from the electronic version of cMAP (http://map.lab.nig.ac.jp:8085/cmap/)

Figure 3 (see also Supplemental Table 2) shows these 36 selected RCSLs, as illustrated by GGT 2.0 (van Berloo 2008). Some of the substituted segments overlap or are duplicated in order to achieve maximum genome coverage. There are some regions where introgressions were not obtained. These include the following short arm (S), long arm (L) and centromeric (C) regions: 3H, 4HC, 4HL, 5HS, 5HL and 7HL. Despite these gaps, substitutions of most of the barley genome, as defined on a genetic rather than a physical basis, were obtained.

Discussion

Marker polymorphism and mapping

One of the key advantages of using the Haruna Nijo/H602 population for this research was that the parents were primary sources of EST sequences. In HarvEST, the SNP positions between Haruna Nijo and H602 are visible and applicable for marker generation, although not all these pair-wise SNPs were integrated into the current OPA.



◀ **Fig. 2** A coordinated cMAP view of chromosome 2H using the OPA-SNP map (see also Fig .1) and the 2,890 transcript (EST) map (Sato et al. 2009). Map information is directly accessible from the electronic version of cMAP (http://map.lab.nig.ac.jp:8085/cmap/)

Sato et al. (2009) detected 3,975 polymorphisms between EST amplicons of Haruna Nijo and H602. These amplicons were obtained using 7,700 PCR primer sets developed from non-redundant 3′ ESTs. Of these SNP polymorphisms, 72% could be mapped (2,890/3,975). These markers included polymorphisms in introns. The present OPA-SNP system used only SNPs in exons and had an average polymorphism rate of 51%. This higher polymorphism rate may be the consequence of the focused approach taken to OPA design, which was based on SNPs identified among multiple EST donors, including Haruna Nijo and H602.

As shown in Fig. 1 (see also the expanded cMAP web image at http://map.lab.nig.ac.jp:8085/cmap/), the OPA-SNP markers tended to cluster at the centromeric regions of each chromosome. This may happen with any set of randomly generated markers, whereas the 384 framework markers from the high-density transcript map (Sato et al. 2009) were selected based on genetic distance and coverage. The relatively high number of markers on chromosomes 2H and 3H (Table 1) led to proportionally more markers clustered in these linkage groups (Table 1). However, co-segregating OPA-SNP markers were developed from different unigenes and therefore represent unique transcript information.

Marker coordination between high-density transcript maps

Consensus EST maps of barley are available, including the resource at HarvEST and the published reports of Stein et al. (2007), Rostoks et al. (2005) and Kota et al. (2008). There is overlap between the populations used in these consensus maps and shared markers among the populations can be used to establish consensus locus orders and to map population-specific makers. An alternative strategy that we have used successfully (Sato et al. 2009) is to use a single population for high-density mapping. In this report, we extend this single population approach to map novel OPA markers and selected prior markers in a single reference population based on actual segregation data.

Detection, alignment and selection of alien segments using RCSLs

The current application of OPA-SNP markers to RCSLs has allowed us to achieve the most precise mapping of

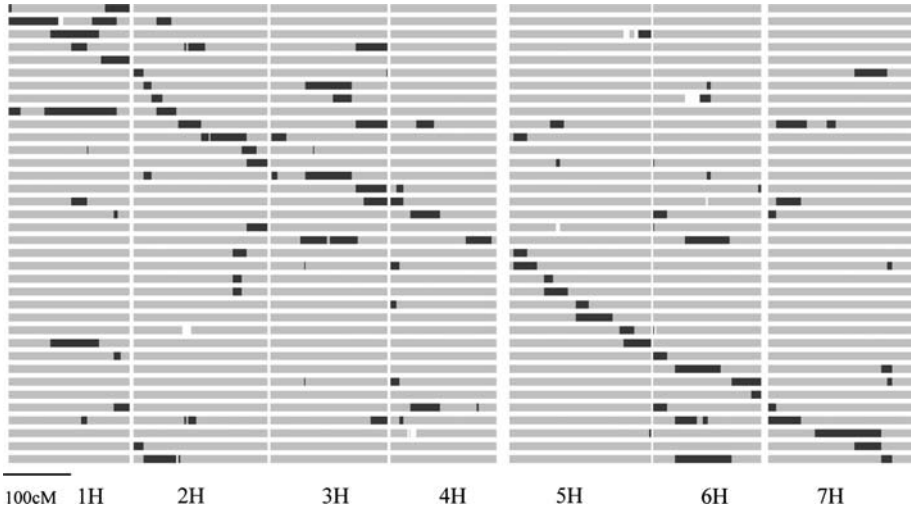


Fig. 3 Graphical genotypes of a minimum set of 36 RCSLs representing H602 wild barley substituted segments (*black*) on the background of cv. Haruna Nijo (*grey*). Missing and heterozygous markers were indicated by *blanks*. Each chromosome is oriented with short arms to the left

substituted segments in barley to date. This is a significant advance over the initial work of Hori et al. (2005) in the same germplasm. These authors identified heterozygous substituted segments in BC_3F_1 individuals using 85 SSR and EST markers and concluded that substituted segments from 19 RCSLs could represent the complete barley genome. However, a larger number of lines proved necessary, probably due to the loss of some segments during generation advancement and the lack of complete marker coverage at the BC_3F_1 generation.

The alignment of 99 BC_3F_5 RCSLs using high-resolution genotyping detected some missing segments in 3HC, 4HC, 4HL, 5HS, 5HL and 7HL. There is also apparent introgression lacking at the terminal of 1HS. However, this region may contain substituted segments, because the H602 allele is saturated at the terminus of this segment. Because the application of RCSLs will depend on the genomic region of interest, it may not be necessary to have a complete set of RCSLs. If such a set is deemed necessary, it could be developed by application of high-throughput genotyping in an earlier generation of RCSL development.

We found that a set of 36 RCSLs represents most of the alleles of a wild barley accession in the background of a cultivated barley accession. The total length of the OPA-SNP map is 1187.4 cM, whereas the map length of the RCSLs is 81.4% of the OPA-SNP (966.4 cM). The average marker density in the RCSLs is 1.33 cM/locus, although there were several dense clusters of markers. Pillen et al. (2003, 2004) reported total linkage maps in lengths in BC_2F_2 populations of 852 and 850 cM. von Korff et al. (2004) reported total map lengths in two BC_2 DH populations of 1,089 and 990 cM, with coverage of 82% of the barley genome. Therefore, it is likely that the Haruna Nijo/H602 RCSLs provide the same level of genome coverage with a high degree of marker saturation.

Efficiency of the OPA-SNP marker system for map coordination

It is a challenge to integrate marker formation from multiple maps. We demonstrate an efficient approach to EST marker coordination using OPA-SNP markers. Further applications of the OPA-SNP system to different EST-based maps will accelerate comprehensive mapping of all ESTs in barley.

All the EST markers mapped on the substituted segments in the RCSLs can be utilized to analyze any gene of interest that is segregating between Haruna Nijo and H602. This method will be especially useful for QTLs detected in this germplasm by Hori et al. (2005). Because BAC libraries are available for Haruna Nijo (Saisho et al. 2007) and H602 (Sato unpublished), these QTLs are targets for map-based isolation.

Acknowledgments We would like to thank Dr. Joe Deyong, Southern California Genotyping Consortium, Illumina BeadLab, UCLA for the OPA-SNP assay; Drs. Timothy J. Close, Prasanna Bhat, Department of Botany & Plant Sciences, University of California, Riverside for basecalling of OPA assay; Drs. Hans Harudrup and Heidi Jaiser (Pajbjergfonden, Denmark) for producing the doubled haploid population; and Dr. Yukiko Yamazaki (National Institute of Genetics, Japan) for developing the cMAP resource. The project was supported by a grant from the Program of Promotion of Basic Research Activities for Innovative Biosciences.

References

Costa JM, Corey A, Hayes P, Jobet C, Kleinhofs A, Kopisch-Obusch A, Kramer SF, Kudrna D, Li M, Riera-Lizarazu O, Sato K, Szucs P, Toojinda T, Vales MI, Wolfe RI (2001) Molecular mapping of the Oregon Wolfe barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103:415–424

Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hansen M, Steemers F, Butler SL, Deloukas P, Galver L, Hunt S, McBride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D,

Chang W, Campbell D, Zhang B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J, Chee MS (2003) Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Biol 68:69–78

Fang Z, Polacco M, Chen S, Schroeder S, Hancock D, Sanchez H, Coe E (2003) cMap: the comparative genetic map viewer. Bioinformatics 19:416–417

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*. Mol Breed 16:295–311

Kleinhofs A, Kilian A, Saghai Maroof M, Biyashev R, Hayes P, Chen F, Lapitan N, Fenwick A, Blake T, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp S, Liu B, Sorrells M, Heun M, Franckowiak J, Hoffman D, Skadsen R, Steffenson B (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–713

Kosambi D (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Kota R, Varshney RK, Prasad M, Zhang H, Stein N, Graner A (2008) EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. Funct Integr Genomics 8:223–233

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

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

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

Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF, Waugh R (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Mol Genet Genomics 274:515–527

Rostoks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall DF, Graner A, Close TJ, Waugh R (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. Proc Natl Acad Sci USA 103:18656–18661

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, Nankaku N, Takeda K (2009) A high density transcript linkage map of barley derived from a single population. Heredity (in press)

Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, Kota R, Varshney RK, Perovic D, Grosse I, Graner A (2007) A 1, 000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theor Appl Genet 114:823–839

van Berloo R (2008) GGT 2.0: versatile software for visualization and analysis of genetic data. J Hered 99:232–236

von Korff M, Wang H, Leon J, Pillen K (2004) Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare* ssp. *spontaneum*) as donor. Theor Appl Genet 109:1736–1745

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

Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesna J, Cakir M, Poulsen D, Wang J, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics 7:206

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