Mol Breeding (2006) 18:143–156
DOI 10.1007/s11032-006-9023-4

Haplotype diversity in the endosperm specific β-amylase gene *Bmy1* of cultivated barley (*Hordeum vulgare* L.)

L. V. Malysheva-Otto · M. S. Röder

Received: 28 September 2005 / Accepted: 29 May 2006 / Published online: 17 August 2006

Abstract Five single nucleotide polymorphism (SNP) sites corresponding to substitutions in the protein sequence of the β-amylase gene *Bmy1* at amino acid (AA) positions 115, 165, 233, 347 and 430 were genotyped in 493 cultivated barley accessions by Pyrosequencing and a CAPS assay. A total of 6 different haplotypes for the *Bmy1* gene were discovered of which 4 haplotypes were identified as previously described alleles *Bmy1*-Sd1, *Bmy1*-Sd2L, *Bmy1*-Sd2H and *Bmy1*-Sd3, while 2 haplotypes were new. A broad spectrum of haplotypes was found in spring barleys, while the winter barleys were dominated by the newly described haplotype *Bmy1*-Sd4. Individual haplotype frequencies varied between the geographic regions.

Three pairs of SNP loci within the gene showed highly significant ($P < 0.0001$) elevated values of linkage disequilibrium (LD) with $r^2 > 0.6$. In the European and Asian subpopulations different loci were in linkage disequilibrium due to the differences in haplotype frequency distributions. By applying LD data to select haplotype tagging SNPs, three SNP sites corresponding to AA positions 115, 233 and 347 were identified that allowed to discriminate 4 haplotypes and to capture 91.6% of the available diversity by distinguishing 452 out of 493 accessions. In a subset of 2-rowed German spring barley varieties 4 SNPs and 2 haplotypes had a significant association with the malting quality parameter final attenuation limit which is related to the total amylolytic enzymatic activity.

Keywords β-Amylase · Barley · Linkage disequilibrium · Malting quality · Pyrosequencing · Haplotype tagging SNPs

L. V. Malysheva-Otto (✉) · M. S. Röder
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben, Germany
e-mail: lmalysheva@yahoo.de

Introduction

The amylolytic enzyme β-amylase (1,4-α-D-glucan maltohydrolase; EC 3.2.1.2) catalyses the release of maltose from starch which is an important biochemical pathway during germination and the malting process. In cultivated barley (*Hordeum vulgare* L.) two loci of β-amylase were mapped on the long arm of chromosome 4H and the short arm of chromosome 2H, however, only the first locus *Bmy1* was reported to be expressed in seed tissue (Kreis et al. 1988).

The first report of full length cDNA and deduced primary structure of endosperm specific barley β-amylase for the variety Hiproly dates to Kreis et al. (1987). Yoshigi et al. (1994) cloned cDNA from the variety Haruna Nijo, and found out three amino acids (AA) substitutions at

positions 233, 347 and 527, resulting from three corresponding SNPs, as compared to variety Hiproly. The *Bmy1* gene is 3825 bp in length without promoter region and contains seven exons and six introns, the deduced protein contains approximately 550 AAs (Yoshigi et al. 1995). Following these pioneering publications several polymorphic forms of the structural gene have been reported for varieties Adorra (Erkkilä et al. 1998), Harrington (Li et al. 2000), Franklin and Schooner (Ma et al. 2001), HA52 (Erkkilä and Ahokas 2001), Steptoe and Morex (Clark et al. 2003) and wild barley (*Hordeum vulgare* subsp. *spontaneum*) (Erkkilä et al. 1998). Four alleles of the β-amylase gene giving rise to proteins with different thermostability and electrophoretic characteristics, *Bmy1*-Sd1, *Bmy1*-Sd2L *Bmy1*-Sd2H and *Bmy1*-Sd3, were identified in cultivated and wild barley (Allison and Swanston 1974; Eglington et al. 1998; Paris et al. 2002). A more recent comprehensive survey revealed a total of 14 β-amylase phenotypes differentiated by isoelectric focusing (Zhang et al. 2004).

Ma et al. (2001) investigated the effects of several specific point mutations causing AA exchanges on the thermostability and kinetic properties of the enzyme. They demonstrated that the properties of the three allelic barley β-amylase forms *Bmy1*-Sd2L, *Bmy*-Sd1 and *Bmy*-Sd2H, characterized by low, intermediate and high thermostability, are due to the combination of three residues in the protein sequence at AA positions 115, 233 and 347. While the form *Bmy1*-Sd1 only had an intermediate thermostability, it expressed favorable kinetic properties and high affinity for substrate. This property was related to the preferred residue cysteine at AA position 115. On the other hand alanine at AA position 233 caused the highest thermostability for allele *Bmy1*-Sd2H and the later described allele *Bmy1*-Sd3 (Paris et al. 2002). The use of a CAPS marker for distinguishing the presence of a codon for valine (in *Bmy1*-Sd1 and *Bmy1*-Sd2L) or alanine (in *Bmy1*-Sd2H and *Bmy1*-Sd3) allowed the investigation of a worldwide collection of barley varieties (Malysheva et al. 2004). The thermostable alleles *Bmy1*-Sd2H and *Bmy1*-Sd3 were predominantly discovered in 6-rowed Asian spring barleys with the highest frequency in Japanese varieties. Only low frequencies of *Bmy1*-Sd2H were found in European 2-rowed spring barleys and the origin of this allele was traced back to the cross Binder × Gull (Malysheva et al. 2004; Sjakste and Röder 2004).

In the present study we have extended our survey to four further SNPs causing AA exchanges at positions 115, 165, 347 and 430 of the protein sequence. The aim of the study was to draw a comprehensive picture of the natural variation in the previously described functional polymorphisms in the *Bmy1*-gene (Ma et al. 2001), to discover novel haplotypes and to analyse their distribution and significance.

Materials and methods

Plant material and DNA extraction

The investigated set included 493 barley accessions with origins from Europe (350 accessions), Asia (53 accessions from East and Central Asian countries), North America (62 accessions), Near East (19 accessions), and Africa (9 accessions). For details see http://www.pgrc.ipk-gatersleben.de/haplotypes. The seeds were either supplied by various European breeding companies or obtained from the genebank of the Leibniz Institute of Plant Genetics and Crop Plant Research at Gatersleben. The information concerning geographical origin, growth habits and row number of accessions was extracted from http://www.stmlf.bayern.de/lbp/forsch/pz/gerstenstamm.pdf or http://www.barley.ipk-gatersleben.de/ebdb.php3. Genomic DNA was extracted from pooled seedlings (5–10 plants per accession) according to Plaschke et al. (1995), and used as a template for PCR amplification.

SNPs identification and pyrosequencing assay design

The endosperm specific barley β-amylase gene *Bmy1* is composed of 7 exons and 6 introns (Yoshigi et al. 1995). For this study an alignment of the 9 publicly available sequences of the *Bmy1* gene (GenBank Nos AF061203,

AF300800, D49999, AF061204, AF414081, AB048949, AF300799, AF414082, AJ301645) was generated using MegAlign/ClustalW (LaserGene, DNASTAR Inc., Madison, WI, USA). SNPs corresponding to the reported substitutions in the deduced protein sequence at AA positions 115, 165, 347 and 430 (Kaneko et al. 2000), were identified in the genomic DNA sequence of cv. Haruna Nijo (GenBank D49999) at positions 613, 1137, 2856 and 3281 bp, respectively. These four SNPs selected for genotyping by pyrosequencing were located in exons 2, 3, 5 and 6.

To amplify the regions containing SNPs PCR primers were designed with the program Primer 0.5, kindly provided by E. Lander (Massachusetts Institute of Technology). As a template genomic sequence of the cv. Haruna Nijo was used (GenBank No. D49999). Then, sequencing primers were developed using SNP Primer Design software (version 1.01 available at http://www.techsupport.pyrosequencing.com/) from Biotage AB (Uppsala, Sweden). Details of all assays are given in Table 1. Since SNPs were dispersed over the gene, and the template for sequencing should not be too long, three fragments carrying loci SNP115, SNP165 and SNPs 347 + 430 were amplified. In all assays the reverse PCR primer was biotinylated and the sequencing primer was oriented in forward direction. Sequencing primers were placed from one to three basepairs upstream the respective SNP.

PCR amplification and SNP-genotyping by pyrosequencing

The PCR-protocol was as described by Röder et al. (1998). PCR was performed in 25 µL volume of PCR buffer (0.01 M Tris, 0.05 M KCl, 1.5 mM $MgCl_2$, 0.01% gelatine) and contained approximately 100 ng of genomic DNA, 0.2 mM of dCTP, dGTP, dTTP, dATP, 0.2 µM of each primer and 1 U of *Taq* polymerase. After 3 min at 94°C, 45 cycles were performed with 1 min at 94°C, 1 min at 55°C, 2 min at 72°C and a final extension step of 10 min at 72°C.

Pyrosequencing was carried out according to the manufacturer's standard protocols using

Table 1 Pyrosequencing assays design for genotyping 4 SNPs in the *Bmy1* gene causing substitutions in AA at positions 115,165, 347 and 430

	SNP position in the reference sequence of Haruna Nijo (GenBank D49999), bp	PCR primers for amplification of the region containing SNP	Sequencing primer for SNP detection by Pyrosequencing
SNP115	611-CGC/TGTGATTC-620 613 bp	FW: 5′-TACAAGCAGTTGTTTGAGCT-3′ REV: 5′-biotin-CAGAGAACTAGAGTGGTGGT-3′ Product size: 273 bp	FW: 5′-GCGGGACGTCGGCA-3′
SNP165	1131-GAAAGAC/GTTC-1140 1137 bp	FW: 5′-TGGCTGTGACAGATGTATGC -3′ REV: 5′-biotin-CTTCAATGTCGACGATAACAC-3′ Product size: 91 bp	FW: 5′-TCAGGGAGAACATGAAA-3′
SNP347	2851-GGATTC/TGGAG-2860 2856 bp	FW: 5′-ACCGTGCTAGCATTAACTTC-3′ REV: 5′-biotin-ATGCATTCTGTCGACAAA-3′ Product size: 499 bp	FW: 5′-GCGGAGATGAGGGAT-3′
SNP430	3275-ACTATGT/CCAA-3284 3281 bp	FW: 5′-ACCGTGCTAGCATTAACTTC-3′ REV: 5′-biotin-ATGCATTCTGTCGACAAA-3′ Product size: 499 bp	FW: 5′-GGTGGAGGGACAAAACT-3′

a Pyrosequencer PSQ HS 96 from Biotage AB (Uppsala, Sweden), and pyrograms were scored with manufacturer's software. Ten µl of the PCR mixture were used for binding to streptavidine coated sepharose beads (Amersham Biosciences GmbH, Freiburg, Germany). Accessions Alexis and Jura which carry according to previously published data different β-amylase alleles (Malysheva et al. 2004) were sampled in every pyrosequencing plate as references. The reference varieties produced identical readings in all experiments indicating a high reliability of the pyrosequencing procedure.

Data analysis

Clustering of the accessions to uncover the groups carrying identical *Bmy1* haplotypes as evaluated by five SNPs was performed by constructing a cladogram (Neighbor joining tree) with software TASSEL (available at http://www.maizegenetics.net).

Nei's genetic distances (Nei and Li 1979) between accessions carrying specific *Bmy1* haplotypes were calculated with the programme GeneFlow V.6 (developed by GENEFLOW Inc., http://www.geneflowinc.com/) using the previously published genotyping data of barley accessions at 48 genome-wide SSR loci (Malysheva-Otto et al. 2006).

Clustering of the accessions carrying specific *Bmy1* haplotypes against the scatterplot of molecular diversity of worldwide barley accessions was conducted using Principal Coordinate Analysis provided by NTSYS (Rohlf 1998) based on the genotyping data of 48 SSRs (Malysheva-Otto et al. 2006). The first two PCOA axes accounted for 8.9% and 5.9% of the total variance.

Evaluation of intragenic linkage disequilibrium (LD) estimated as squared allele frequency correlations, r^2, and tests for associations were performed using the software package TASSEL. Association tests with two traits, malt extract and final attenuation limit, were performed with the General Linear Model (GLM) option on two subsets of varieties, 2-rowed winter + spring and 2-rowed spring only, to account for the population structure.

Results

Identification of *Bmy1* haplotypes

The complete genotyping data of 493 barley accessions at four SNPs corresponding to substitutions in the protein sequence at AA positions 115, 165, 347 and 430 are available at http://www.pgrc.ipk-gatersleben.de/haplotypes. For haplotype analysis within the *Bmy1* gene previously published data of the SNP site at 2311 bp corresponding to AA position 233 in exon 4 obtained with a CAPS marker (Malysheva et al. 2004) were also considered. Clustering of accessions based on five SNPs was performed by constructing a cladogram with software TASSEL. The cladogram allowed to classify 6 distinct haplotypes in 493 accessions originating from all over the world (Fig. 1). While four haplotypes were unambiguously identified as the previously described alleles *Bmy1*-Sd1, *Bmy1*-Sd2L, *Bmy1*-Sd2H and *Bmy1*-Sd3 (Ma et al. 2001; Paris et al. 2002), 2 haplotypes were new. One of the newly described haplotypes clustered very closely to the previously detected allele *Bmy1*-Sd1. It was designated as haplotype *Bmy1*-Sd1b, while to the previously described allele the name *Bmy1*-Sd1a was assigned. The second new haplotype clustered separately, it was named *Bmy1*-Sd4 (Fig. 1). Table 2 summarizes SNP patterns for the dissected haplotypes, haplotype frequencies, and genetic distances within the accessions forming one subcluster evaluated by 48 genome-wide SSR loci. In a worldwide survey of 493 accessions the frequencies of the haplotypes representing previously described alleles varied from 0.02 for

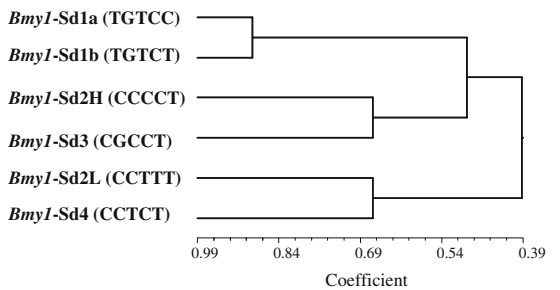

Fig. 1 Cladogram of *Bmy1* haplotypes based on five SNPs within the gene

Table 2 Natural diversity of the *Bmy1* haplotypes as evaluated with five SNP loci

SNP (*AA*) at protein sequence position:					Allele assignment**	Haplotype frequency (accs. no. per haplotype)	Mean value (range) of Nei's distance within the group***
snp115	snp165	snp233	snp347	snp430			
T (C)*	G (E)	T (V)	C (S)	C (A)	*Bmy1*-Sd1a[1]	0.20 (97)	0.56 (0.042–0.96)
T (C)	G (E)	T (V)	C (S)	T (V)	*Bmy1*-Sd1b	0.07 (34)	
C (R)	C (D)	T (V)	T (L)	T (V)	*Bmy1*-Sd2La[1]	0.14 (71)	0.67 (0.083–0.958)
C (R)	C (D)	C (A)	C (S)	T (V)	*Bmy1*-Sd2H[1]	0.12 (58)	0.72 (0.104–0.979)
C (R)	G (E)	C (A)	C (S)	T (V)	*Bmy1*-Sd3[2]	0.02 (7)	0.75 (0.438–0.917)
C (R)	C (D)	T (V)	C (S)	T (V)	*Bmy1*-Sd4	0.45 (226)	0.62 (0.021–0.979)

*Amino acids present in the protein due to sequence polymorphisms are shown in brackets. C – cysteine, R – arginine, E – glutamic acid, D – aspartic acid, V – valine, A – alanine, S – serine, L – leucine

**Previously described alleles were assigned to specific SNP patterns according to [1]Ma et al. (2001) and [2]Paris et al. (2002). Two new haplotypes were called *Bmy1*-Sd1b and *Bmy1*-Sd4

***The mean value was calculated with GeneFlow using the data of 48 SSR loci, the range of variation is given in brackets

Bmy1-Sd3, occurring in 7 accessions, to 0.20 for *Bmy1*-Sd1, occurring in 97 accessions. The newly described haplotype *Bmy1*-Sd1b, occurring in 34 accessions had a frequency of 0.07, while the other, *Bmy1*-Sd4, occurring in 226 accessions had a relatively high frequency of 0.45. Genetic distances within the group of accessions carrying the same haplotype or forming one subcluster calculated based on 48 genome-wide SSRs indicated clustering of very diverse accessions within all *Bmy1* haplotype groups. Nei's distances between accessions carrying specific *Bmy1* haplotypes varied in the interval from 0.021 to 0.979 with the mean values higher than 0.5 (Table 2).

Distribution of *Bmy1* haplotypes in cultivated barley germplasm

The distribution of *Bmy1* haplotypes in accessions originating from various geographic regions indicated that a wide spectrum of haplotypes was detected in Europe, Asia, North America and Near East (Fig. 2). However, haplotype profiles within the continents were different. Accessions from Europe and Asia exhibited specific profiles with few dominating haplotypes, *Bmy1*-Sd2H in Asia, *Bmy1*-Sd1a, *Bmy1*-Sd2L and *Bmy1*-Sd4 in Europe. In North America and Near East different haplotypes were approximately equally represented except for the absence of *Bmy1*-Sd1a in the Near East. For the African continent the number of accessions included in the analysis (9 accessions) was too low to make safe conclusions.

Two haplotypes belonging to the *Bmy1*-Sd1 cluster, *Bmy1*-Sd1a and *Bmy1*-Sd1b, revealed reverse geographic distributions on the Eurasian continent. In the Near East *Bmy1*-Sd1b was detected with the highest frequency (42.1%) compared to Asia, with a frequency of *Bmy1*-Sd1b of 9.4%, and Europe, with a very low frequency of *Bmy1*-Sd1b of 2.2%. On the other hand, no *Bmy1*-Sd1a was found among the accessions from the Near East, in Asia *Bmy1*-Sd1a appeared with low frequency (3.8%), while *Bmy1*-Sd1a was most

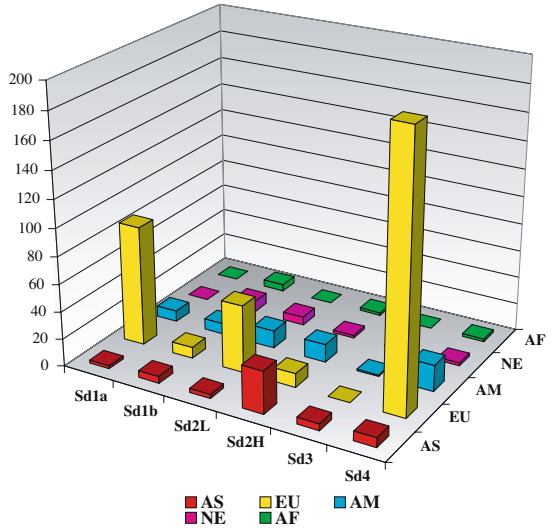

Fig. 2 Haplotype frequencies of *Bmy1* in barley accessions originating from various geographic regions worldwide. EU – Europe, AS – Asia, AF – Africa, NE – Near East, AM – America

abundant in Europe (24.8%). In North America both haplotypes were equally represented. The haplotypes *Bmy1*-Sd2L, *Bmy1*-Sd2H and *Bmy1*-Sd4 were dominating in one continent, *Bmy1*-Sd2L and *Bmy1*-Sd4 in Europe with the frequencies of 14.0% and 56.0%, respectively, *Bmy1*-Sd2H in Asia with the frequency of 53.4%. All haplotypes occurred with low frequencies on the other continents as well. Among 7 varieties carrying haplotype *Bmy1*-Sd3 5 originated from Asia.

Distinct differences in the occurrence of *Bmy1* haplotypes in spring versus winter accessions, and in 2-rowed versus 6-rowed accessions were observed (Fig. 3). Whereas spring accessions carried practically the whole spectrum of *Bmy1* haplotypes with the prevalence of *Bmy1*-Sd1a, *Bmy1*-Sd2L, and *Bmy1*-Sd2H, in the selected subset of winter accessions haplotype *Bmy1*-Sd4 occurred ten times more often than the others and the haplotypes *Bmy1*-Sd1a, *Bmy1*-Sd2L and *Bmy1*-Sd3 were represented in only one or two accessions. Both, 2-rowed and 6-rowed accessions carried the complete spectrum of haplotypes. However, the majority of the 2-rowed accessions possessed haplotypes *Bmy1*-Sd1a and *Bmy1*-Sd2L, while in the 6-rowed accessions haplotype *Bmy1*-Sd4 was most frequent.

To check the interference of *Bmy1* haplotypes clusters with the population structure, the data of the haplotypes' occurrences in barley accessions were overlayed on the two-dimensional scatterplots presenting clustering of the global barley

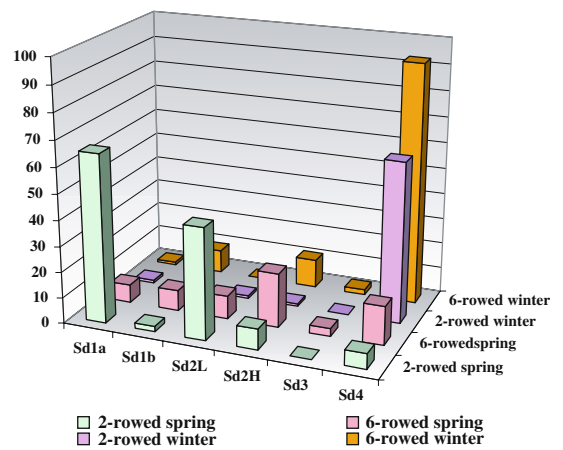

Fig. 3 Haplotype frequencies of *Bmy1* in barley accessions with different agronomic traits

population as evaluated with 48 SSR genomic loci (Fig. 4). The analysis revealed that varieties carrying specific haplotypes clustered adjacently on the scatterplot of the genome-wide molecular diversity. However, neither of the identified haplotypes was localized exclusively within the detected population substructures. On the contrary, specific *Bmy1* haplotypes were dispersed over various continents and over accessions possessing different growth types and row number. This is in accordance with the distribution of *Bmy1* haplotypes discussed above, and with the highly variable values of genetic distances between accessions belonging to a specific group of haplotypes (Table 2).

Linkage disequilibrium within the *Bmy1* gene and haplotype tagging SNPs

Linkage disequilibrium (LD) for five intragenic SNPs was assessed as squared allele frequency correlations, r^2, in the complete set of 493 accessions and in three subsets extracted with regard to geographic origin and agronomic traits, Asian, European and European 2-rowed spring accessions (Fig. 5). The LD values are indicated for all significant $r^2 > 0.1$ ($P < 0.0001$). In the world-wide population and the subpopulations of 350 European barleys, only three pairs of sites revealed highly significant elevated levels of $r^2 > 0.6$, SNP115–SNP165, SNP115–SNP430 and SNP165–SNP430, indicating an interrupted pattern of LD within the *Bmy1* gene (Fig. 5, A, C). In the subpopulation of the 112 European 2-rowed spring barleys three additional loci pairs were in LD, SNP115–SNP347, SNP165–SNP347 and SNP347–SNP430 (Fig. 5, D). LD for loci pair SNP115–SNP165 equalled $r^2 = 1.0$ in the subpopulations of European barleys (350 accessions) and European 2-rowed spring barleys (112 accessions). In the subpopulation of 53 accessions originating from Asia one loci pair from the world-wide and European populations was reproduced, SNP115–SNP165, and the other loci pair in LD, SNP115–SNP233, was characteristic for Asia (Fig. 5, B). The differing LD patterns between Europe and Asia may be caused by differences in haplotype frequencies in the two subsets.

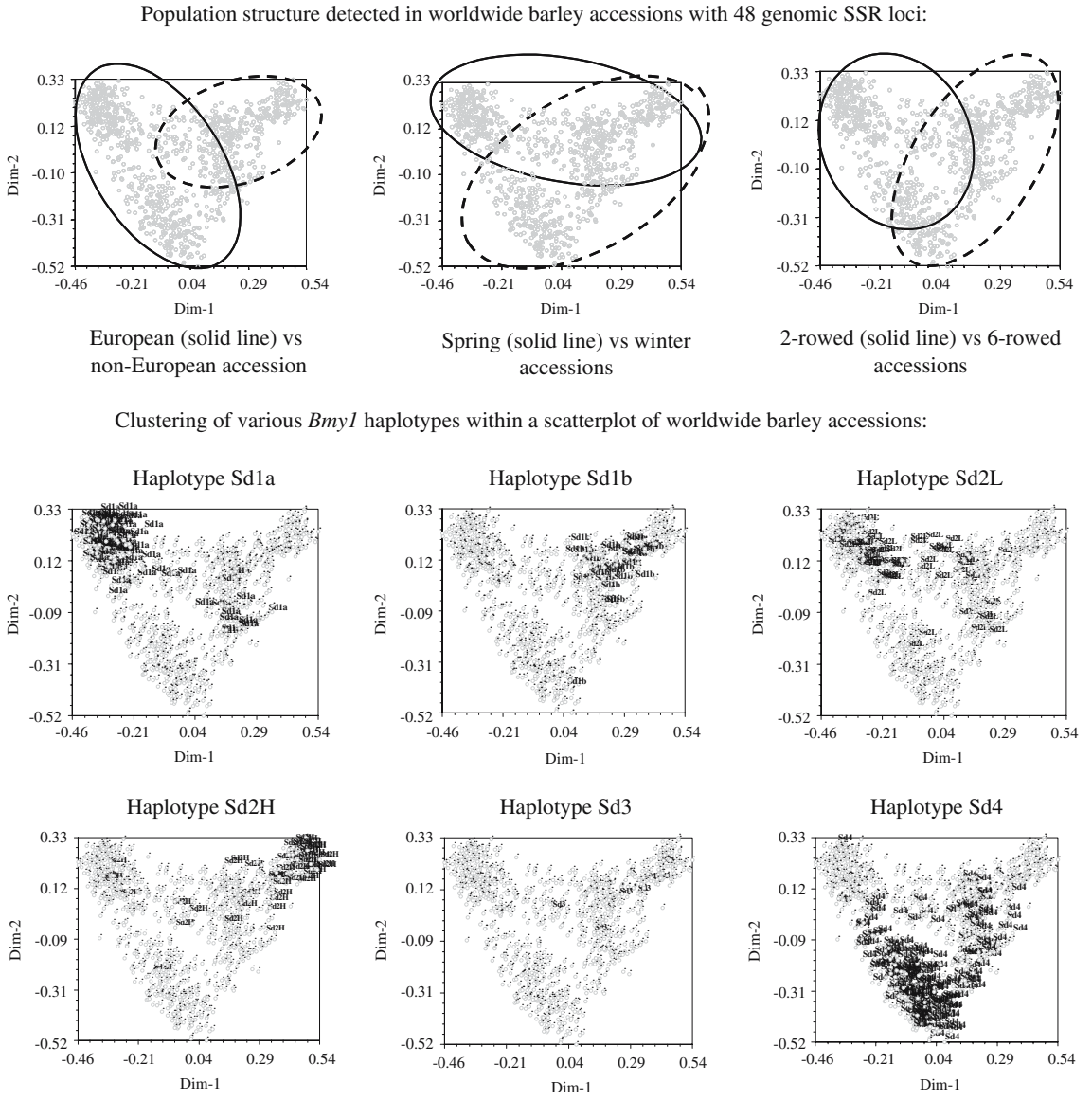

Fig. 4 Principal Coordinate Analysis of *Bmy1* haplotypes. The upper graphs depict the population structure of a worldwide set of barley cultivars based on data of 48 genome-wide SSR loci. In the following graphs various haplotypes of *Bmy1* were overlayed on the SSR data

Regarding the high level of LD within the *Bmy1* gene among loci SNP115, SNP165 and SNP430, we suggest that genotyping of SNPs at AA positions 115, 233 and 347 is sufficient for efficient identification of different *Bmy1* haplotypes. The assumption is supported by clustering of the haplotypes in the cladogram (Fig. 1). Indeed, genotyping of only these three SNPs could discriminate between the majority of the detected *Bmy1* haplotypes (Table 2). Genotyping of these three SNPs did not discriminate between the haplotypes *Bmy1*-Sd1a/*Bmy1*-Sd1b, and haplotypes *Bmy1*-Sd2H/*Bmy1*-Sd3. However, considering the very rare occurrence of the *Bmy1*-Sd1b and *Bmy1*-Sd3 haplotypes in the world and their predominant occurrence in Asia and Near East, only when accessions from these geographic regions are involved in the studies, then the genotyping marker set should also include SNP165. In summary, SNPs relating to AA

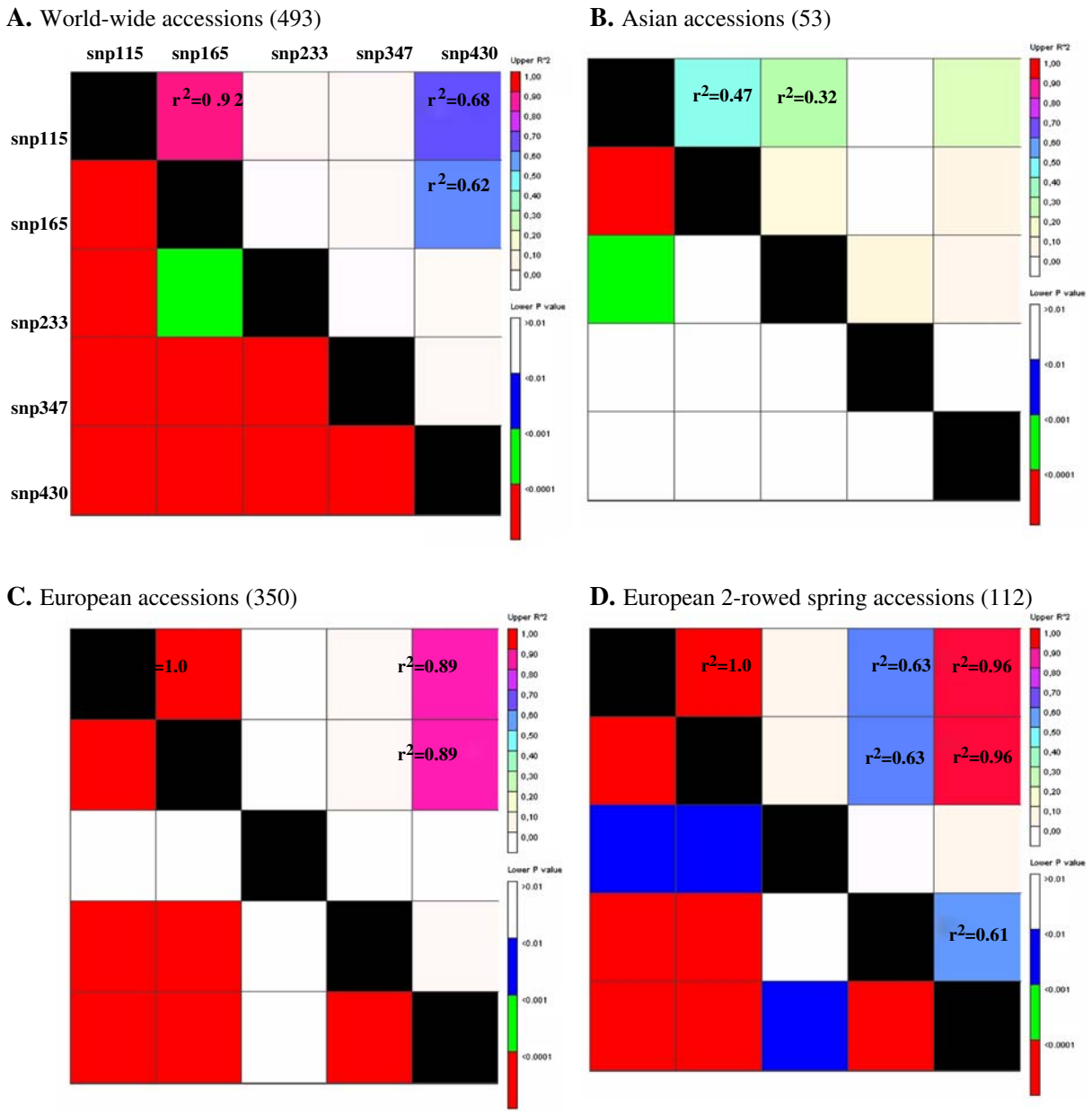

Fig. 5 Linkage disequilibrium matrices visualizing r^2 between SNP sites within the *Bmy1* gene for the worldwide set of accessions and extracted subsets

positions 115, 233 and 347 allowed to discriminate 4 haplotypes and to capture approximately 91.6% of the available diversity (452 out of 493 accessions were separated). Thus SNP155, SNP233 and SNP347 could be considered as *Bmy1* haplotype tagging SNPs for plant breeding practice.

Associations with characters important for malting

For 36 German 2-rowed spring and winter varieties included in our set quality data for malting were available from the regular official publications of the descriptions of released varieties

(Beschreibende Sortenliste, Bundessortenamt 1994–2003). These data correlated to the genotyping results produced in this study are reported in Table 3. While the 11 winter barley varieties all had the haplotype *Bmy1*-Sd4, the 25 spring barley varieties comprised the haplotypes *Bmy1*-Sd1a, *Bmy1*-Sd2L and *Bmy1*-Sd4 (Table 3). Interestingly among 56 additional German varieties from our set officially registered as ''feeding'' cultivars (according to descriptions of Bundessortenamt) 50 carried the allele *Bmy1*-Sd4 (data not shown).

Within the subset of 36 varieties four polymorphic SNPs (SNP relating to AA position 233 was monomorphic in European varieties) and three haplotypes (*Bmy1*-Sd1a, *Bmy1*-Sd2L and *Bmy1*-Sd4) were tested for associations with the two traits, malt extract and final attenuation limit. Malt extract is one of the most important properties for assessing brewing quality and describes the percentage of soluble ingredients, mostly starch and protein, in the wort. Final attenuation limit describes the sum of all fermentable products in the wort and is related to the amylolytic enzymatic activity. A high final attenuation limit is preferred for a good malting variety. To take into account the population structure we performed calculations considering the whole subset

Table 3 List of European varieties evaluated by Bundessortenamt (BSA) for malting quality traits and involved in association testing

Variety	Year of release	Haplotype	Row number and growth habit	Trait evaluation by BSA (quality range from 1 to 9)*	
				Final attenuation	Malt extract
Adonis	2002	Sd1a	2-r-S	6	7
Alexis	1986	Sd1a	2-r-S	8	8
Annabell	1999	Sd1a	2-r-S	7	8
Barke	1996	Sd1a	2-r-S	8	9
Brenda	1995	Sd1a	2-r-S	8	9
Danuta	2000	Sd1a	2-r-S	7	7
Hanka	1997	Sd1a	2-r-S	7	9
Krona	1990	Sd1a	2-r-S	8	9
Madeira	1998	Sd1a	2-r-S	8	8
Madonna	1997	Sd1a	2-r-S	8	8
Madras	1997	Sd1a	2-r-S	7	9
Minna	1992	Sd1a	2-r-S	7	8
Otis	1992	Sd1a	2-r-S	7	8
Pasadena	1998	Sd1a	2-r-S	8	8
Ria	1998	Sd1a	2-r-S	7	8
Scarlett	1995	Sd1a	2-r-S	7	9
Sissy	1990	Sd1a	2-r-S	8	8
Thuringia	1995	Sd1a	2-r-S	7	9
Libelle	1990	Sd1a	2-r-S	7	8
Apex	1983	Sd2L	2-r-S	5	7
Baronesse	1989	Sd2L	2-r-S	5	6
Gimpel	1979	Sd2L	2-r-S	5	9
Mentor	1996	Sd2L	2-r-S	4	8
Viskosa	1999	Sd2L	2-r-S	7	8
Extract	1997	Sd4	2-r-S	7	9
Angora	1990	Sd4	2-r-W	7	7
Astrid	1990	Sd4	2-r-W	5	5
Babylone	1994	Sd4	2-r-W	4	6
Jura	1995	Sd4	2-r-W	5	7
Labea	1992	Sd4	2-r-W	6	6
Leonie	2000	Sd4	2-r-W	8	7
Marinka	1985	Sd4	2-r-W	5	6
Opal	1998	Sd4	2-r-W	6	7
Regina	1995	Sd4	2-r-W	7	7
Tiffany	1996	Sd4	2-r-W	7	7
Vanessa	2000	Sd4	2-r-W	7	7

*Data for the years 1994–2003 were considered

Table 4 Significant associations of individual SNPs and specific haplotypes with two quality traits identified with General Linear Model (GLM) analysis based on 36 German malting varieties

	A set of 36 2-rowed winter and spring varieties		A subset of 25 2-rowed spring varieties	
	Quality trait		Quality trait	
	Final attenuation[1]	Malt extract	Final attenuation	Malt extract
Locus				
entire gene	**	*	**	ns
snp115	*	*	**	ns
snp165	*	*	*	ns
snp347	*	ns	***	ns
snp430	*	*	**	ns
Haplotype				
Sd4	ns	**	ns	ns
Sd1a	*	*	*	ns
Sd2L	*	ns	***	ns
Total R^2 of the model	0.486	0.423	0.629	0.166

[1]Associations are presented based on *P*-values with *, **, *** indicating $P < 0.001$, $P < 0.0001$ and $P < 0.00001$, respectively. ns = non significant ($P > 0.01$)

of 36 varieties and 25 spring varieties only (Table 4). For the whole subset associations with significance $P < 0.001$ were observed for the trait malt extract with 3 SNPs and 2 haplotypes, and for the trait final attenuation with all 4 SNPs considered and 2 haplotypes. By testing only the spring barleys no significant association was observed for the trait malt extract, while the associations for the trait final attenuation increased in significance (Table 4). Most significant ($P < 0.00001$) was the association with haplotype *Bmy1*-Sd2L; varieties carrying this β-amylase haplotype showed decreased values for final attenuation limit (Table 3). The total variation explained by the model was $R^2 = 0.629$, for the trait final attenuation.

Discussion

Haplotype diversity and its interference with population structure

The information available up to now in the literature indicated the occurrence of four different alleles of the *Bmy1* gene coding for enzymes with different thermostability (Ma et al. 2001; Paris et al. 2002). By genotyping 493 barley accessions originating from all over the world at five SNPs in the coding region of the gene we found out 6 different haplotypes, with four haplotypes corresponding to the previously described alleles and two new ones. This suggests that an enzyme with the same properties can be determined by more than one haplotype. Recently, Zhang et al. (2004) differentiated β-amylase *Bmy1* phenotypes in 8270 worldwide barley accessions according to the thermostability type (A, B or C) and based on isoelectrofocusing patterns (I, II or IV). The authors were able to classify 14 different phenotypes.

Practically the complete spectrum of haplotypes was present at all continents, except Africa where the number of investigated accessions was low, and only the frequencies of specific haplotypes varied between the continents. Dominating (occurring in >5% of accessions) were 5 haplotypes, one haplotype *Bmy1*-Sd3 was rare (occurring in 7 accessions) (Fig. 2). Zhang et al. (2004) also reported that among 14 *Bmy1* phenotypes observed in their study only 5 phenotypes occurred with high frequencies. The spectrum of frequent haplotypes was broader in spring accessions (4 haplotypes occurring in >20% of accessions), while in winter accessions only one haplotype occurred very often (*Bmy1*-Sd4), all others were rare (Fig. 3).

Recently, we reported that a world-wide population of cultivated barley was highly structured with regard to geographic origin, spring or winter growth habit and row number (Malysheva-Otto et al. 2006). As barley accessions carrying identical *Bmy1* alleles were clustered together but independently of the population structure

(Fig. 4), we suggest that their clustering can, probably, be imposed by similar malting qualities.

Analysis of linkage disequilibrium and functional significance of single SNPs

In this study we detected intragenic LD in the barley *Bmy1* gene with r^2 ranging from 0.32 to 1.0. These values are higher than in maize (*Zea mays*) with r^2 around 0.1 within 2kb (Remington et al. 2001; Palaisa et al. 2004), or aspen (*Populus tremula* L.) with $r^2 > 0.1$ within few hundreds kb (Ingvarsson 2005), but similar to the values detected for *Adh1-3* loci in wild barley with r^2 around 0.6–0.8 within 1.6 kb (Lin et al. 2002).

The LD analysis between the 5 investigated SNP sites within *Bmy1* gene revealed different patterns of LD in the worldwide set of varieties and in various subsets (Fig. 5). In the complete dataset and in the subset comprising only European varieties LD among pairs of SNP loci at AA positions 115, 165 and 430 was detected, while loci at AA positions 233 and 347 showed no LD. In the European 2-rowed spring barleys an additional locus at AA position 347 was in LD with loci 115, 165 and 430 (Fig. 5, D). In the subpopulation of Asian accessions two loci pairs were in LD, 115–165 and 115–233, the latter characteristic for Asia. In all cases, except for one loci pair in the two subpopulations of European barleys, the value of r^2 never equalled $r^2 = 1.0$, indicating recombination between any two SNP loci. Only in the sets of European barleys (350 accessions) and European 2-rowed spring barleys (112 accessions) LD for loci pair SNP115–SNP165 equalled $r^2 = 1.0$. The reason for this could be the absence of exotic varieties carrying alternative alleles at locus snp165 in European breeding programmes. The differing LD patterns and absence of a correlation between the LD and the physical distance between the SNP sites within the gene suggests that manmade selection may be a major factor determining various haplotype frequencies and the resulting patterns of LD. The ''interrupted pattern'' of LD within the gene observed in our study was also reported for *Aegilops tauschii* puroindoline genes (Massa et al. 2004), for the maize *Dwarf8* gene (Thornsberry et al. 2001) and for about 2 Mb region surrounding the maize *Y1* gene (Palaisa et al. 2004).

The functional significance of AA substitutions in five positions 115, 165, 233, 347 and 430 was investigated by site-directed mutagenesis by Ma et al. (2001). The authors showed that cysteine (C) at position 115 is responsible for improved kinetic properties of the enzyme and higher affinity to the substrate, while alanine (A) and serine (S) at positions 233 and 347, were preferred residues for enhancing thermostability of the enzyme, with the effects of the two residues being additive. Substitutions in AA positions 165 and 430 had no effect on kinetic properties and thermostability. Correlations between SNPs in the coding region of the gene (AA positions 115, 165 and 430) and biochemical properties of the enzyme were also confirmed by Clark et al. (2003). In this respect, a variety possessing the AAs cysteine (C), alanine (A) and serine (S) at positions 115, 233 and 347, was theoretically predicted to carry an enzyme combining improved thermostability with good substrate affinity (Ma et al. 2001). Unfortunately this combination of SNP alleles was not detected in our set.

To the best of our knowledge no data are available regarding the gene expression level of the *Bmy1* gene, except for the detection of different levels of β-amylase mRNA in three barley cultivars (Kreis et al. 1987). Correlation between DNA sequence polymorphisms and expression level was recently reported for another candidate gene for malting quality, serine carboxypeptidase I (*Cxp1*) (Potokina et al. 2006). For β-amylase it would be interesting to examine in the future if differentially expressed alleles of the gene exist.

Carlson et al. (2004) suggested to use a LD-based algorithm to select informative SNPs for association analysis in humans. In their experiments a threshold for $r^2 > 0.8$ allowed to resolve >80% of all haplotypes. In our experiments, if LD data were applied to select haplotype tagging SNPs, then with regard to haplotype frequencies and a threshold of $r^2 > 0.6$, three SNPs relating to AA positions 115, 233 and 347 were suggested as haplotype tagging. They allowed to discriminate 4 haplotypes and captured 91.6% of the available diversity by separating 452 out of 493 accessions.

Bmy1 haplotypes and malting quality of barley varieties

The possibility of marker assisted selection for the improvement of malting quality using β-amylase as a marker turned this enzyme into a subject of intense studies in the last years. Several groups reported the identification of different *Bmy1* alleles and their occurrence and importance in malting varieties (Eglington et al. 1998; Kihara et al. 1998; Ma et al. 2001, Paris et al. 2002; Polakova et al. 2003; Zhang et al. 2004). Recently, direct mapping of β-amylase activity QTLs demonstrated that they occured almost always in conjunction with diastatic power QTLs (Clancey et al. 2003). In Australian breeders' populations the malting quality alleles at the chromosome 4H *Bmy1* locus were associated with increased diastatic power (Coventry et al. 2003). Comparison of our data with the official evaluation implemented by German authorities (Bundessortenamt) revealed that 93% of varieties evaluated as ''feeding'' carried haplotype *Bmy1*-Sd4, and 75% of varieties evaluated as ''malting'' carried haplotype *Bmy1*-Sd1a, emphasizing the correlation between alleles of the *Bmy1* gene and malting qualities of cultivars.

The GLM (General Linear Model) analysis performed with software TASSEL revealed that DNA sequence polymorphisms in the *Bmy1* gene were associated with the malting quality parameter final attenuation limit, but not with the parameter malt extract in German spring barley varieties. Three SNP loci, at AA positions 115, 347 and 430, as well as haplotypes *Bmy1*-Sd1a and *Bmy1*-Sd2L showed an increased probability of association with the trait final attenuation in a subset of 2-rowed spring varieties. These sites represented a cystein-arginin change at AA 115, a serine-leucine change at AA 347 and alanine-valine change at AA 430. These polymorphisms determine kinetic properties (AA 115) and thermostability (AA 347) of the enzyme (Ma et al. 2001; Clark et al. 2003). Implementation of association analysis with TASSEL software was recently reported for three loci underlying QTL for synthesis of compounds involved in corn earworm resistance (Szalma et al. 2005).

The endosperm specific *Bmy1* locus was mapped on the long arm of the chromosome 4H (Kreis et al. 1988). Later, two very important genes for agronomical performance were mapped on the same arm of the chromosome 4H, *mlo* (Wettstein-Knowles 1993) and *Vrn-H2* (old nomenclature: *Sh*) (Laurie et al. 1995). Definitely, in the cultivated barley germplasm these two genes were for a long time under strong selection. However there are no real indications that selection for thermostable *Bmy1* alleles and, consequently, *Bmy1* allele frequencies were affected by linkage drag to these loci. The genetic distances between the *Bmy1* locus and the *mlo* and *Vrn-H2* loci are 5 cM and 32 cM, respectively (Frankowiak et al. 1997). For the *Vrn-H2* locus a complete correlation of the recessive allele with the spring growth habit was reported (Dubcovsky et al. 2005), whereas all individual *Bmy1* haplotypes revealed in our study occurred both in spring and winter varieties.

In a previous study (Malysheva et al. 2004) we showed that in Europe the occurrence of thermostable alleles *Bmy1*-Sd2H and *Bmy1*-Sd3 is very low. The results presented here confirmed that the European germplasm is dominated by haplotypes *Bmy1*-Sd1a and *Bmy1*-Sd4. Therefore, the pyrosequencing markers developed in this study provide a valuable tool for marker-assisted selection and breeding in order to exploit the natural occurring genetic diversity of the *Bmy1* gene.

Acknowledgement We thank Ellen Weiß for excellent technical assistance.

References

Allison MJ, Swanston JS (1974) Relationships between b-amylase polymorphisms in developing, mature and germinating grains of barley. J Inst Brew 80:285–291

Beschreibende Sortenliste. Getreide, Mais, Ölfrüchte, Leguminosen, Hackfrüchte. Bundessortenamt. 1994–2003. Landbuch Verlagsgesellschaft mbH. Hannover

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 74:106–120

Clark SE, Hayes PM, Henson CA (2003) Effects of single nucleotide polymorphisms in β-amylase1 alleles from barley on functional properties of the enzymes. Plant Physiol Biochem 41(9):798–804

Clancey JA, Han F, Ullrich SE (2003) Comparative mapping of β-amylase activity QTLs among three barley crosses. Crop Sci 43:1043–1052

Coventry SJ, Collins HM, Barr AR, Jefferies SP, Chalmers KJ, Logue SJ, Langridge P (2003) Use of putative QTLs and structural genes in marker assisted selection for diastatic power in malting barley (*Hordeum vulgare* L). Aust J Agr Res 54:1241–1250

Dubcovsky J, Chen C, Yan L (2005) Moleculat characterization of the allelic variation at the VRN-H2 vernalization locus in barley. Mol Breeding 15:395–407

Eglinton JK, Langridge P, Evans DE (1998) Thermostability variation in alleles of barley β-amylase. J Cereal Sci 28:301–309

Erkkilä MJ, Ahokas H (2001) Special barley β-amylase allele in a Finnish landrace line HA52 with high grain enzyme activity. Hereditas 134:91–95

Erkkilä MJ, Leah R, Ahokas H, Cameron-Mills V (1998) Allele-dependent barley grain β-amylase activity. Plant Physiol 117:679–685

Franckowiak J (1997) Revised linkage maps for morphological markers in barley, Hordeum vulgare. Barley Genetics Newsletter 26:9–21

Kaneko T, Kihara M, Ito K (2000) Genetic analysis of β -amylase thermostability to develop a DNA marker for malt fermentability improvement in barley, *Hordeum vulgare*. Plant Breeding 119:197–201

Kihara M, Kaneko T, Ito K (1998) Genetic variation of β-amylase thermostability among varieties of barley, *Hordeum vulgare* L., and relation to malting quality. Plant Breeding 117:425–428

Kreis M, Williamson M, Buxton B, Pywell J, Hejgaard J, Svendsen I (1987) Primary structure and differential expression of β-amylase in normal and mutant barley. Eur J Biochem 169:517–525

Kreis M, Williamson M, Shewry PR, Sharp P, Gale M (1988) Identification of a 2nd locus encoding β-amylase on chromosome-2 of barley. Genetical Res 51:13–16

Laurie DA, Pratchett N, Bezant JH, Snape JW (1995) RFLP mapping of 5 major genes and 8 quantitative trait loci controlling flowering time in a winter x spring barley (*Hordeum vulgare* L.) cross. Genome 38(3):575–585

Li C, Langridge P, Zhang X, Eckstein PE, Rossnagel BG, Lance RC, Lefol EB, Lu M, Harvey BL, Scoles GJ (2000) *Hordeum vulgare* cultivar RFLP-Harrington *Bmy1* gene for Sd1 beta-amylase, complete cds. Direct submission to GenBank AB048949

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

Ma YF, Evans DE, Logue SJ, Langridge P (2001) Mutations of barley β-amylase that improve substrate-binding affinity and thermostability. Mol Genet Genomics 266:345–52

Malysheva L, Ganal MW, Röder MS (2004) Evaluation of cultivated barley (*Hordeum vulgare* L.) germplasm for the presence of thermostable alleles of β-amylase. Plant Breeding 123:128–131

Malysheva-Otto L, Ganal MW, Röder MS (2006) Analysis of molecular diversity, population structure and linkage disequilibrium in worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). *BMC*. Genetics 7:6

Massa AN, Morris CF, Gill BS (2004) Sequence diversity of Puroindoline-a, Puroindoline-b, and the grain softness protein genes in *Aegilops tauschii* Coss. Crop Sci 44:1808–1816

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76(10):5269–5273

Palaisa K, Morgante M, Tingey S, Rafalski A (2004) Long-range pattern of diversity and linkage disequilibrium surrounding the maize gene *Y1* gene are indicative of an asymmetric selective sweep. Proc Natl Acad Sci USA 101(26):9885–90

Paris M, Jones MGK, Eglinton JK (2002) Genotyping single nucleotide polymorphisms for selection of barley β-amylase alleles. Plant Mol Biol Rep 20:149–159

Plaschke J, Ganal MW, Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91:1001–1007

Polakova K, Laurie D, Vaculova K, Ovesna J (2003) Characterization of β-amylase alleles in 79 barley varieties with Pyrosequencing. Plant Mol Biol Rep 21:439–47

Potokina E, Prasad M, Malysheva L, Röder MS, Graner A (2006) Expression genetics and haplotype analysis reveal cis regulation of serine carboxypeptidase I (*Cxp1*), a candidate gene for malting quality in barley (*Hordeum vulgare* L.). Func Integr Genomics 6: 25–35

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ESIY (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98(20):11479–84

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Rohlf FJ (1998) NTSYS-pc: numerical taxonomy and multivariate analysis system. Applied Biostatistics Inc, New York

Sjakste TG, Röder M (2004) Distribution and inheritance of β-amylase alleles in north European barley varieties. Hereditas 141:39–45

Szalma SJ, Buckler ES IV, Snook ME, McMullen MD (2005) Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. Theor Appl Genet 110:1324–1333

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielse D, Buckler IV ES (2001) Dwarf8 polymor-

phisms associate with variation in flowering time. Nature Genetics 28:286–289

von Wettstein-Knowles P (1993) Barley (Hordeum vulgare) 2N = 14. In: Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press 1993 Ed. O'Brien SJ

Yoshigi N, Okada Y, Sahara H, Koshino S (1994) PCR cloning and sequencing of the beta-amylase cDNA from barley. J Biochem (Tokyo) 115:47–51

Yoshigi N, Okada Y, Sahara H, Tamaki T (1995) A structural gene encoding beta-amylase of barley. Biosci Biotechnol Biochem 10:1991–1993

Zhang W, Kaneko T, Ishii M, Takeda K (2004) Differentiation of β-amylase phenotypes in cultivated barley. Crop Sci 44:1608–1614