Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress

Nils Rostoks, Sharon Mudie, Nicola Bonar, Linda Cardle, Joanne Russell, Luke Ramsay, David F. Marshall, Robbie Waugh

Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, Scotland, UK

Summary

Single Nucleotide Polymorphisms (SNPs) are by far the most abundant genetic markers encountered in all organisms and are particularly useful for genome-wide high throughput linkage and association mapping. We have developed SNP markers for barley genes associated with abiotic stress response by resequencing ca. 1400 gene fragments from 8 diverse barley lines.

1. A total of over 1 Mb of consensus sequences were obtained. Automated mutation detection using Polyphred identified 5372 SNPs and 361 indels in ca. 1200 unigenes. Approximately 2000 of the SNPs were manually confirmed using Mutation Surveyor software and were used for diversity assessment of the germplasm.

2. Sequence and polymorphism data is stored in a dedicated database available through the SCRI Bioinformatics website at http://bioinf.scri.ac.uk/barley_snpdb

3. A set of 333 genes was mapped in three mapping populations using mostly SNPs and indels as markers and an integrated linkage map was constructed incorporating a total of 1237 loci. Sequence-based markers have enabled us to study barley - rice gene synteny at the DNA sequence level and establish a framework synteny map.

4. A subset of 96 gene fragments was re-sequenced in an additional 16 Western European elite spring and winter barley lines. SNP diversity was studied in the whole data set, as well as the spring and winter barley subsets.

5. SNP markers have been used to design an Illumina Golden Gate OPA for linkage mapping of >1000 barley genes and for sequence diversity assessment of cultivated European barely.

Barley genes

A selection of genes that were differentially expressed in response to various abiotic stresses were identified using genome-wide transcription profiling based on the Affymetrix Barley1 GeneChip (Close et al. 2004) and by homology to stress-responsive genes in other plant species. Gene lists are available from the SCRI website:

http://bioinf.scri.sari.ac.uk/cgi-bin/barleysnp/2508_contig_list
http://bioinf.scri.sari.ac.uk/cgi-bin/barleysnp/862_contig_list

Barley genotypes

Barley line	Species	Growth habit	Ear
OWB D	*H. vulgare* ssp. *vulgare*	Spring	2 row
OWB R	*H. vulgare* ssp. *vulgare*	Spring	6 row
Steptoe	*H. vulgare* ssp. *vulgare*	Spring	6 row
Morex	*H. vulgare* ssp. *vulgare*	Spring	6 row
Lina	*H. vulgare* ssp. *vulgare*	Spring	2 row
HS92	*H. vulgare* ssp. *spontaneum*	Winter	2 row
Golden Promise	*H. vulgare* ssp. *vulgare*	Spring	2 row
Optic	*H. vulgare* ssp. *vulgare*	Spring	2 row
Barke	*H. vulgare* ssp. *vulgare*	Spring	2 row
Triumph	*H. vulgare* ssp. *vulgare*	Spring	2 row
Chariot	*H. vulgare* ssp. *vulgare*	Spring	2 row
Atem	*H. vulgare* ssp. *vulgare*	Spring	2 row
Blenheim	*H. vulgare* ssp. *vulgare*	Spring	2 row
Prisma	*H. vulgare* ssp. *vulgare*	Spring	2 row
Camargue	*H. vulgare* ssp. *vulgare*	Spring	2 row
Igri	*H. vulgare* ssp. *vulgare*	Winter	2 row
Pastoral	*H. vulgare* ssp. *vulgare*	Winter	2 row
Marinka	*H. vulgare* ssp. *vulgare*	Winter	2 row
Halcyon	*H. vulgare* ssp. *vulgare*	Winter	2 row
Regina	*H. vulgare* ssp. *vulgare*	Winter	2 row
Panda	*H. vulgare* ssp. *vulgare*	Winter	2 row
Fighter	*H. vulgare* ssp. *vulgare*	Winter	2 row
Puffin	*H. vulgare* ssp. *vulgare*	Winter	2 row
Franka	*H. vulgare* ssp. *vulgare*	Winter	6 row

Collaborations

The abiotic stress gene list was generated in collaboration with Dr. Timothy J. Close (UC Riverside, CA, USA). SNP and linkage mapping data will be integrated with the barley BAC physical map within the NSF funded project.
SCRI, UCR (Timothy J. Close) and IPK (Andreas Graner, Nils Stein) SNP data has been integrated to generate the first barley Illumina Golden Gate OPA.

References

Close et al. (2004) A new resource for cereal genomics: 22K barley GeneChip comes of age. Plant Physiol 134:960-968

Costa et al. (2001) Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103 415-424

Kleinhofs et al. (1993) A molecular, isozyme and morphological map of the barley (Hordeum vulgare) genome. Theor Appl Genet 86:705-712

Kleinhofs and Graner (2001) An integrated map of the barley genome. In: Phillips RL, Vasil IK (eds) DNA-based markers in plants, 2nd edn. Kluwer, Dordrecht, The Netherlands, pp 187-199

Ramsay et al. (2000) A simple sequence repeat-based linkage map of barley. Genetics 156:1997-2005

Rostoks et al. (2004) Single nucleotide polymorphism mapping of the barley genes involved in abiotic stresses. 9th International Barley Genetics Symposium, Brno, Czech Republic (http://www.ibgs.cz)

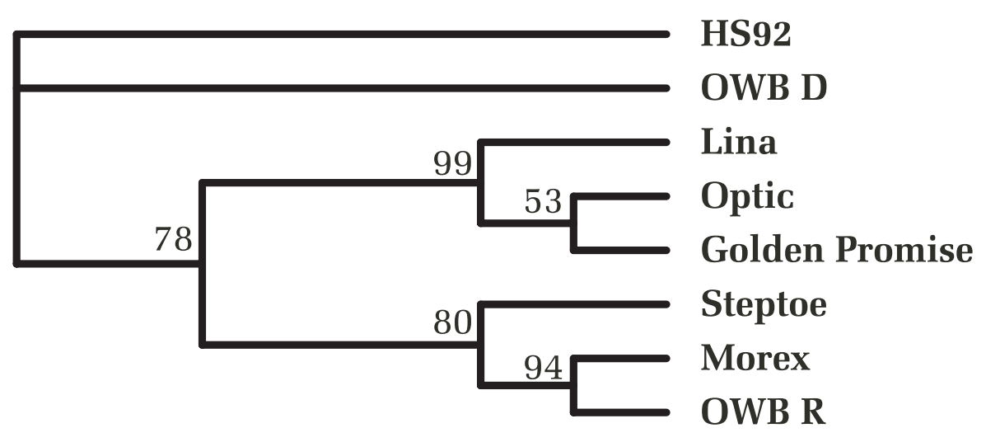
Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF, Waugh R (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Mol Genet Genomics 274: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

Acknowledgments

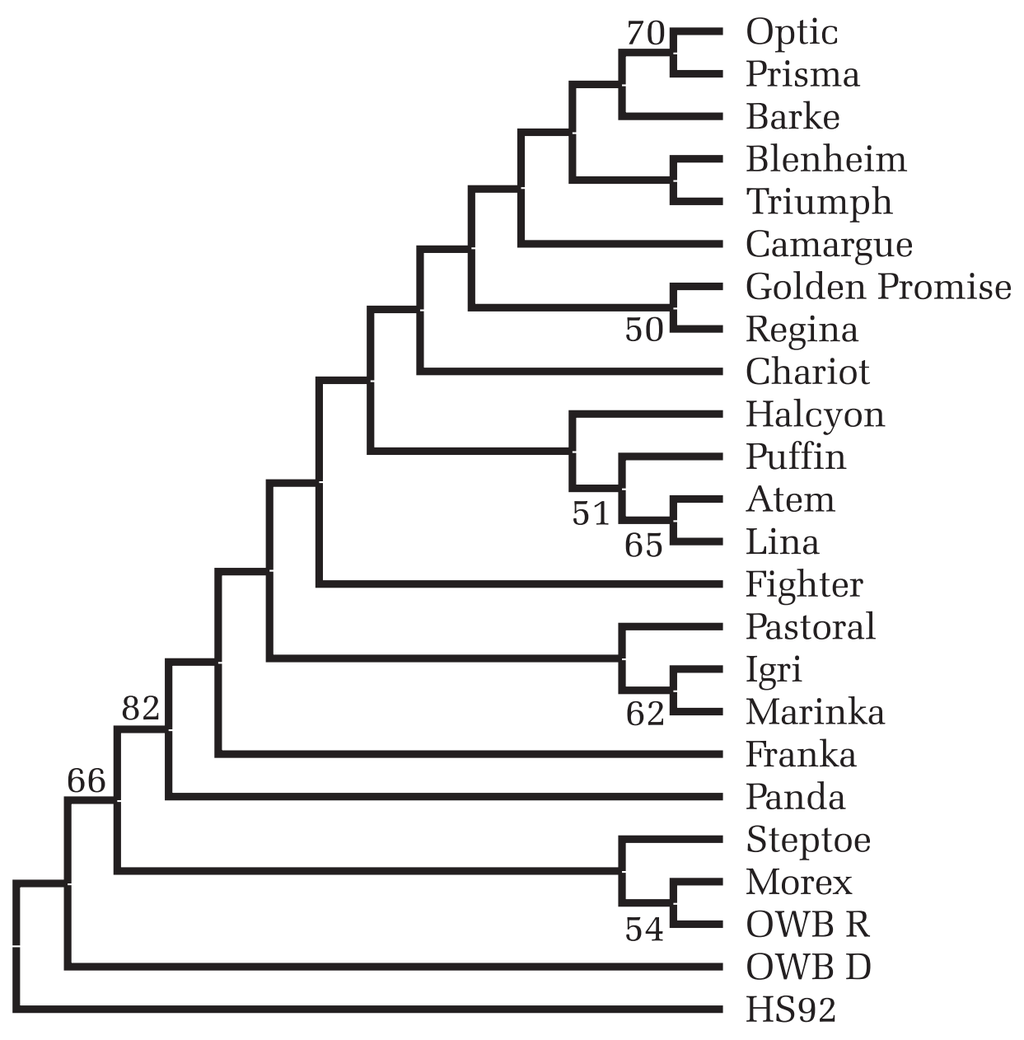
Funding is provided by BBSRC and SEERAD to RW.
Generous sharing of unpublished barley microarray data by Dr. TJ Close is gratefully acknowledged.

Diversity in barley

1. Diversity assessment of the barley germplasm used for SNP discovery. Extended majority rule consensus tree calculated using the dnapars program from the Phylip package of the 8 barley accessions derived from 2008 polymorphic sites from 602 gene fragments. Numbers on nodes indicate bootstrap values.



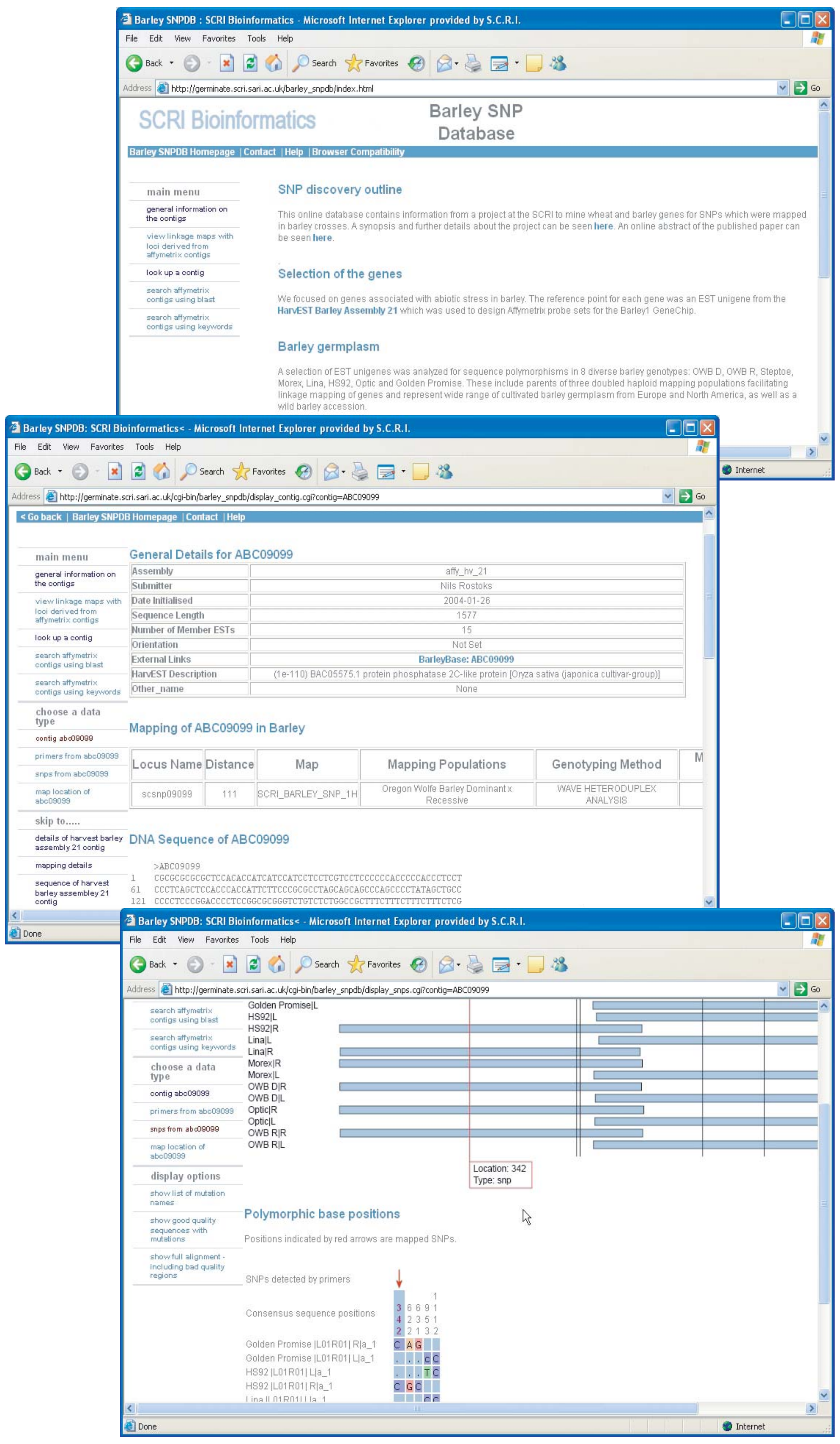
2. A subset of 96 loci were further analyzed for presence of SNP in16 European spring and winter barley cultivars. 35011 bp of high quality sequences were obtained from all 24 accessions across 88 loci. Maximum parsimony tree was calculated using Mega 3.1. Bootstrap values >50% are shown on the branches. *Hordeum vulgare* ssp. *spontaneum* line HS92 was used to root the tree.



	24 accessions	19accessions (cultivated EU barley)	EU spring barley	EU winter barley
Number of SNP	367	193	138	153
Bp per SNP (Pi)	92 bp (0.00241)	175 bp (0.00167)	249 bp (0.00149)	221 bp (0.00152)
Total number of INDEL (bp per indel)	51 (660 bp)	24 (1408 bp)	ND	ND
Average and maximum number of haplotypes (Hd)	3.2, max 9 (0.333)	2.0, max 6 (0.241)	ND	ND
Bp per SNP in introns / exons	74 bp / 104 bp	161 bp / 188 bp	ND	ND
Number of SNP in CDS (bp per SNP in CDS)	135 (117 bp)	74 (214 bp)	48 (330 bp)	57 (278 bp)
Number of amino acid substitutions	40	24	18	15

Barley SNP database

Available at http://bioinf.scri.ac.uk/barley_snpdb



http://germinate.scri.sari.ac.uk/barley_snpdb/index.html

SCRI Bioinformatics

Barley SNP Database

Barley SNPDB Homepage | Contact | Help | Browser Compatibility

main menu
general information on the contigs
view linkage maps with loci derived from affymetrix contigs
look up a contig
search affymetrix contigs using blast
search affymetrix contigs using keywords

SNP discovery outline

This online database contains information from a project at the SCRI to mine wheat and barley genes for SNPs which were mapped in barley crosses. A synopsis and further details about the project can be seen **here**. An online abstract of the published paper can be seen **here**.

Selection of the genes

We focused on genes associated with abiotic stress in barley. The reference point for each gene was an EST unigene from the **HarvEST Barley Assembly 21** which was used to design Affymetrix probe sets for the Barley1 GeneChip.

Barley germplasm

A selection of EST unigenes was analyzed for sequence polymorphisms in 8 diverse barley genotypes: OWB D, OWB R, Steptoe, Morex, Lina, HS92, Optic and Golden Promise. These include parents of three doubled haploid mapping populations facilitating linkage mapping of genes and represent wide range of cultivated barley germplasm from Europe and North America, as well as a wild barley accession.

Barley SNPDB: SCRI Bioinformatics< - Microsoft Internet Explorer provided by S.C.R.I.

http://germinate.scri.sari.ac.uk/cgi-bin/barley_snpdb/display_contig.cgi?contig=ABC09099

< Go back | Barley SNPDB Homepage | Contact | Help

main menu
general information on the contigs
view linkage maps with loci derived from affymetrix contigs
look up a contig
search affymetrix contigs using blast
search affymetrix contigs using keywords
choose a data type
contig abc09099
primers from abc09099
snps from abc09099
map location of abc09099
skip to.....
details of harvest barley assembly 21 contig
mapping details
sequence of harvest barley assembley 21 contig

General Details for ABC09099

Assembly	affy_hv_21
Submitter	Nils Rostoks
Date Initialised	2004-01-26
Sequence Length	1577
Number of Member ESTs	15
Orientation	Not Set
External Links	BarleyBase: ABC09099
HarvEST Description	(1e-110) BAC05575.1 protein phosphatase 2C-like protein [Oryza sativa (japonica cultivar-group)]
Other_name	None

Mapping of ABC09099 in Barley

Locus Name	Distance	Map	Mapping Populations	Genotyping Method
scsnp09099	111	SCRI_BARLEY_SNP_1H	Oregon Wolfe Barley Dominant x Recessive	WAVE HETERODUPLEX ANALYSIS

DNA Sequence of ABC09099

```
>ABC09099
1   CGCGCGCGCGCTCCACACCATCATCCATCCTCCTCGTCCTCCCCCCACCCCCACCCTCCT
61  CCCTCAGCTCCACCCACCATTCTTCCCGCGCCTAGCAGCAGCCCAGCCCCTATAGCTGCC
121 CCCCTCCCGGACCCCTCCGGCGCGGGTCTGTCTCTGGCCGCTTTCTTTCTTTCTTTCTCG
```

Barley SNPDB: SCRI Bioinformatics< - Microsoft Internet Explorer provided by S.C.R.I.

http://germinate.scri.sari.ac.uk/cgi-bin/barley_snpdb/display_snps.cgi?contig=ABC09099

search affymetrix contigs using blast
search affymetrix contigs using keywords
choose a data type
contig abc09099
primers from abc09099
snps from abc09099
map location of abc09099
display options
show list of mutation names
show good quality sequences with mutations
show full alignment - including bad quality regions

Golden Promise|L
HS92|L
HS92|R
Lina|L
Lina|R
Morex|R
Morex|L
OWB D|R
OWB D|L
Optic|R
Optic|L
OWB R|R
OWB R|L

Location: 342
Type: snp

Polymorphic base positions

Positions indicated by red arrows are mapped SNPs.

SNPs detected by primers

Consensus sequence positions

Golden Promise |L01R01| R|a_1
Golden Promise |L01R01| L|a_1
HS92 |L01R01| L|a_1
HS92 |L01R01| R|a_1

Integrated linkage map of 333 barley sequence-based loci (Rostoks et al. 2005). Mapping was done in Oregon Wolfe Barley Dominant x Recessive, Steptoe x Morex minimapper and Lina x HS92 DH populations (Costa et al. 2001; Kleinhofs and Graner 2001; Ramsay et al. 2000). Linkage maps for individual populations and the integrated map shown below was calculated using JoinMap 3.0. The map is based on 1237 loci. BIN markers (Kleinhofs and Graner, 2001; shown in bold) are included to facilitate comparison with other barley maps. Names of the loci consist of the HarvEST assembly #21 unigene number preceded by letters indicating type of the polymorphism used to map the locus (scsnp, scssr or scind).
Syntenous relationships with rice were explored by comparing HarvEST unigene sequences with the TIGR rice pseudomolecule 3 gene models using BLASTX. The best rice BLASTX hits and their chromosome locations in megabases are aligned along the barley chromosomes. All rice hits have E-value threshold better than e^{-20}.