Wheat Genetic Resources – How to Exploit?

A. BÖRNER[1], K. NEUMANN[1] and B. KOBILJSKI[2]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany; [2]Institute of Field and Vegetable Crops, 21000 Novi Sad, Serbia; e-mail: boerner@ipk-gatersleben.de

Abstract: It is estimated that world-wide existing germplasm collections contain about 7.5 million accessions of plant genetic resources for food and agriculture. Wheat (*Triticum* and *Aegilops*) represents the biggest group comprising 900 000 accessions. However, such a huge number of accessions is hindering a successful exploitation of the germplasm. The creation of core collections representing a wide spectrum of the genetic variation of the whole assembly may help to overcome the problem. Here we demonstrate the successful utilisation of such a core collection for the identification and molecular mapping of genes (Quantitative Trait Loci) determining the agronomic traits flowering time and grain yield, exploiting a marker-trait-association based technique. Significant marker-trait associations were obtained and are presented. The intrachromosomal location of many of these associations coincided with those of already identified major genes or quantitative trait loci, but others were detected in regions where no known genes have been located to date.

Keywords: association mapping; *ex situ* collections; flowering time; genetic resources; grain yield

World-wide existing germplasm collections for food and agriculture contain about 7.5 million accessions of which wheat represents the biggest group with nearly 900 000 samples followed by rice (~ 775 000) and barley (~ 470 000). A list of the ten world-wide largest germplasm collections by crop is given in Table 1 (FAO 2009). The wheat collections comprise 858 000 accessions of the genus *Triticum* and another 42 000 accessions of the wild ancestor *Aegilops*. Genebank collections containing > 25 000 and > 1 500 accessions of the genera *Triticum* and *Aegilops*, respectively, are given in Tables 2 and 3 (FAO 2009). Beside an accurate preservation of the germplasm the evaluation of the collections is a very important task for further utilisation (Börner 2006). It is the prerequisite for the identification of genes to be used in breeding programmes for crop improvement.

A successful exploitation of the germplasm collections is often hampered by the huge numbers of accessions stored in the seedbanks. Therefore, core collections representing the genetic variation of the whole set were created. Applying a methodology designated association mapping, largely and effectively used in human genetics, such core collections can be exploited genetically. Using that approach, a population of individual genotypes will be analysed in order to detect associations between marker patterns and trait expressions.

As an example we present results obtained from a core collection of 96 wheat accessions. Data are shown for the agronomic traits flowering time and grain yield recorded during up to six growing seasons. The wheat lines were genotyped using diversity array technology (DArT) markers in order to investigate marker-trait-associations. Homologous and homoeologous relationships of the detected loci and comparable major genes or quantitative trait loci (QTLs) already described are discussed.

Table 1. The ten largest worldwide germplasm collections by crop (FAO 2009)

Crop	Genus	Accessions
Wheat	*Triticum*	857 940
Rice	*Oryza*	773 947
Barley	*Hordeum*	470 470
Maize	*Zea*	327 931
Bean	*Phaseolus*	262 369
Sorghum	*Sorghum*	235 711
Soybean	*Glycine*	229 947
Oat	*Avena*	148 260
Groundnut	*Arachis*	128 461
Cotton	*Gossypium*	104 780

Table 2. Worldwide existing genebank collections of the genus *Triticum* comprising > 25 000 accessions (FAO 2009)

Institution	Country	No. of accessions
CIMMYT	Mexico	110 281
NCGRP	USA	57 348
ICGR-CAAS	China	43 039
NBPGR	India	35 889
ICARDA	Syria	34 951
NIAS	Japan	34 652
VIR	Russia	35 959
IDG	Italy	32 751
IPK	Germany	28 191

CIMMYT – Centro Internacional de Mejoramiento de Maíz y Trigo; NCGRP – National Center for Genetic Resources Preservation; ICGR-CAAS – Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences; NBPGR – National Bureau of Plant Genetic Resources; ICARDA – International Centre for Agricultural Research in the Dry Areas; NIAS – National Institute of Agrobiological Science; VIR – N.I. Vavilov Research Institute of Plant Industry; IDG – Instituto del Germoplasma; IPK – Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung

MATERIALS AND METHODS

A set of 96 winter wheat genotypes from altogether 21 different countries and five continents was considered for the mapping studies. These genotypes were selected from a larger core collection created at the Institute of Field and Vegetable Crops, Novi Sad, Serbia and chosen on the basis of contrasting phenotypic expression of 20 traits relevant for breeding (Kobiljski *et al*. 2002; Quarrie *et al*. 2003). The material is listed in Table 4.

The genotypes were cultivated in field plots in Novi Sad, Serbia, between 1993 and 2001. Each plot with a size of 1.2 m^2 contained 6 rows with a distance of 20 cm between the rows. Three independent plots per genotype and year were grown. The traits considered were recorded during six (flowering time) and five (grain yield) seasons. Flowering time was determined as days to flowering, when 50% of the spikes per plot flowered. Grain yield was revealed from five spikes sampled from 5 plants per plot.

Genotyping using DArT markers was performed by Triticarte Pty. Ltd. (Canberra, Australia; http://www.triticarte.com.au/), which offers this high-throughput genome profiling service. In total we received a number of 874 polymorphic DArT markers. In order to create the linkage groups we used

Table 3. Worlwide existing genebank collections of the genus *Aegilops* comprising > 1500 accessions (FAO 2009)

Institution	Country	No. of accessions
ICCI-TELAVUN	Israel	9 146
ICARDA	Syria	3 847
NPGBI-SPII	Iran	2 653
NIAS	Japan	2 433
VIR	Russia	2 248
NCGRP	USA	2 207
LPGPB	Armenia	1 827
IPK	Germany	1 526

ICCI-TELAVUN – Lieberman Germplasm Bank, Institute for Cereal Crops Improvement, Tel-Aviv University; ICARDA – International Centre for Agricultural Research in the Dry Areas; NPGBI-SPII – National Plant Gene Bank of Iran, Seed and Plant Improvement Institute; NIAS – National Institute of Agrobiological Science; VIR – N.I. Vavilov Research Institute of Plant Industry; NCGRP – National Center for Genetic Resources Preservation; LPGPB – Laboratory of Plants Gene Pool and Breeding; IPK – Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung

Table 4. Cultivar names/designations and countries (code from UN-webpage) of origin of the genotypes investigated

Acciaio – ITA	L 1A/91 – SRB	Purdue 5392 – USA
Ai-bian – CHN	Lambriego Inia – CHL	Red Coat – USA
Al-Kan-Tzao – CHN	Lr 10 – USA	Renesansa – SRB
Ana – HRV	Lr 12 – USA	Rusalka – BGR
Avalon – GBR	Magnif 41 – ARG	Siete Cerros – MEX
Bankuty 1205 – HUN	Mex. 120 – MEX	Saitama 27 – JPN
BCD 1302/83 – MDA	Mex. 17 bb – MEX	Sava – SRB
Benni multifloret – USA	Mex. 3 – MEX	Semillia Eligulata – USA
Bezostaja 1 – RUS	Min. Dwarf – AUS	Slavija – SRB
Brigant – GBR	Mina – SRB	Sofija – SRB
Cajeme 71 – MEX	Mironovska 808 – UKR	Sonalika – IND
Capelle Desprez – FRA	Nizija – SRB	Suwon 92 – IND
Centurk – USA	Norin 101 – JPN	Szegedi 768 – HUN
Ching-Chang 6 – CHN	Norin 10/Brevor14 – USA	Tibet Dwarf – CHN
Cook – AUS	Novosadska Crvena – SRB	Timson – AUS
Donska polupat. – RUS	Nova banatka – SRB	TJB 990-15 – GBR
Durin – FRA	NS 22/92 – SRB	Tom Thumb – CHN
F 4 4687 – ROM	NS 33/90 – SRB	Tr. Compactum – LVA
Florida – USA	NS 46/90 – SRB	Tr. Sphaerococcum – USA
Gala – ARG	NS 55-25 – SRB	Triple Dirk B – AUS
Hays 2 – USA	NS 559 – SRB	Triple Dirk B (bulk) – AUS
Helios – USA	NS 602 – SRB	Triple Dirk S – AUS
Highbury – GBR	NS 63-24 – SRB	UC 65680 – USA
Hira – IND	NS 66/92 – SRB	UPI 301 – IND
Holly E – USA	NS 74/95 – SRB	Vel – USA
Hope – USA	NS 79/90 – SRB	Vireo"S" – MEX
Inia 66 – MEX	Peking 11 – CHN	WWMCB 2 – USA
INTRO 615 – USA	Phoenix – USA	ZG 1011 – HRV
Ivanka – SER	PKB Krupna – SRB	ZG 987/3 – HRV
Kite – AUS	Pobeda – SRB	ZG K 3/82 – HRV
L-1 – HUN	Purdue/Loras – USA	ZG K 238/82 – HRV
L 1/91 – SRB	Purdue 39120 – USA	ZG K T 159/82 - HRV

the mapping information provided by Crossa *et al*. (2007). For estimating the population structure of the material under investigation, a subset of 219 randomly distributed markers was used to run the software STRUCTURE (Pritchard *et al*. 2000). Two subpopulations were identified in our core set. The calculation of testing for an association between markers and traits were done with the software programme TASSEL 2.01 (Bradbury *et al*. 2007). The general linear model (GLM) with including the Q-Matrix from STRUCTURE as correction for population structure was used.

In addition, with the newer version TASSEL 2.1 the mixed linear model (MLM) was implemented using Q-Matrix and the kinship-Matrix (Yu & Buckler 2006). Marker-trait-associations (MTAs) significant in both models and with $P < 0.05$ in four out of six and three out of five years for flowering time and grain yield, respectively, were considered only.

RESULTS AND DISCUSSION

Details for genetic map, population structure, linkage disequilibrium, the comparison of the GLM and MLM models as well as for a range of further characters investigated using the given core collection are presented by Neumann *et al.* (2011). For the traits flowering time and grain yield 13 and 10 MTAs, respectively, were detected (Figures 1 and 2).

The flowering time MTAs are located on chromosomes 1B, 1D (2 markers), 2B, 2D, 4B, 5B, 5D, 6A (2 markers), 6B and 7A (2 close markers). Flowering time is determined by genes controlling vernalisation response (*Vrn*), photoperiod response (*Ppd*) and earliness *per se* (*Eps*). Because all entries of the core collection investigated were pure or facultative winter wheats, no *Vrn* gene loci should be detectable. In fact none of the homoeologous group 5L major loci (McIntosh *et al*. 2008) was detected. Although one MTA was discovered on the long arm of chromosome 5B, this locus maps about 10 cM distal to the centromere whereas *Vrn-B1* was described to be located at a distance of ~ 100 cM from the centromere on 5BL (Leonova *et al*. 2003).

Major photoperiod response genes map to the short arms of the homoeologous group 2 chromosomes (McIntosh *et al*. 2008). In the present study an MTA was detected only on chromosome 2DS. However, considering the GLM only there are MTAs on the short arms of 2A and 2B but with MLM they are not fulfilling the significance criteria (details see Neumann *et al.* 2011). Although the germplasm investigated was assumed to be photoperiod insensitive the MTAs detected may reflect allelic variation at the *Ppd-1* genes. Another QTL associated with photoperiod response and described by Kuchel *et al*. (2006) is located in the centromere region of chromosome 7AS but seem not to correspond to the major locus detected on 7AS in the present study.

The flowering time MTAs mapped close to the centromeres of chromosomes 1B and 1D may reflect variation at *Eps* genes, since such a gene has been identified both on chromosome 1BL (Tóth *et al*. 2003) as well as in a syntenic region (1HL) of barley (Laurie *et al*. 1995). A further *Eps* locus maps to chromosome arm 3AL (Miura *et al*. 1999, Börner *et al*. 2002) as well as to barley chromosome arm 3HL (Laurie *et al*. 1995). In the present study, however, a significant association on 3AL was identified with the GLM only. With the MLM this MTA is significant only at the 0.1 *P*-level (Neumann *et al.* 2011).

MTAs for the trait grain yield were spread over chromosomes 1A, 3A (2 markers), 3B, 4A, 4B, 5B, 6B, 7A and 7B. For this trait a direct comparison with a genome-wide association mapping study performed by Crossa *et al*. (2007) was possible. The authors exploited also the set of DArT

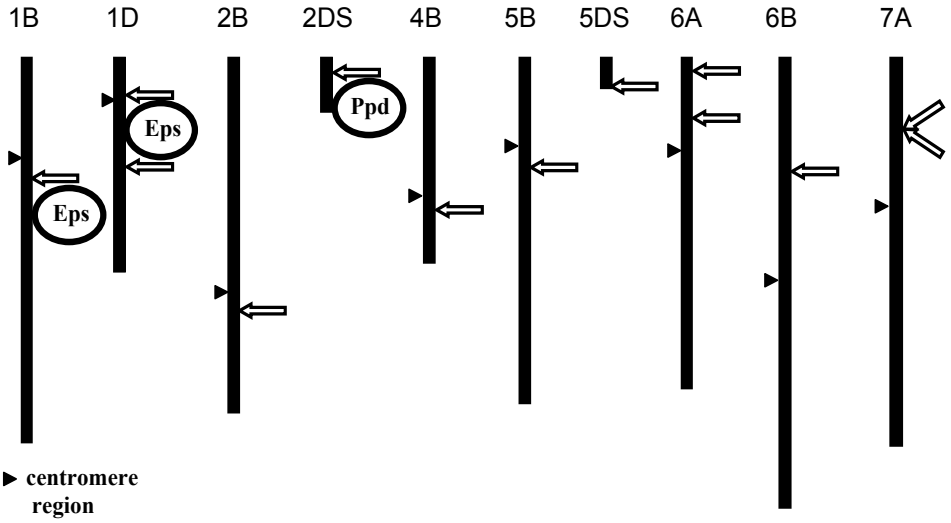

Figure 1. MTAs for flowering time marked by arrows. Positions of comparable genes/QTLs described earlier are indicated below arrows (elipses)

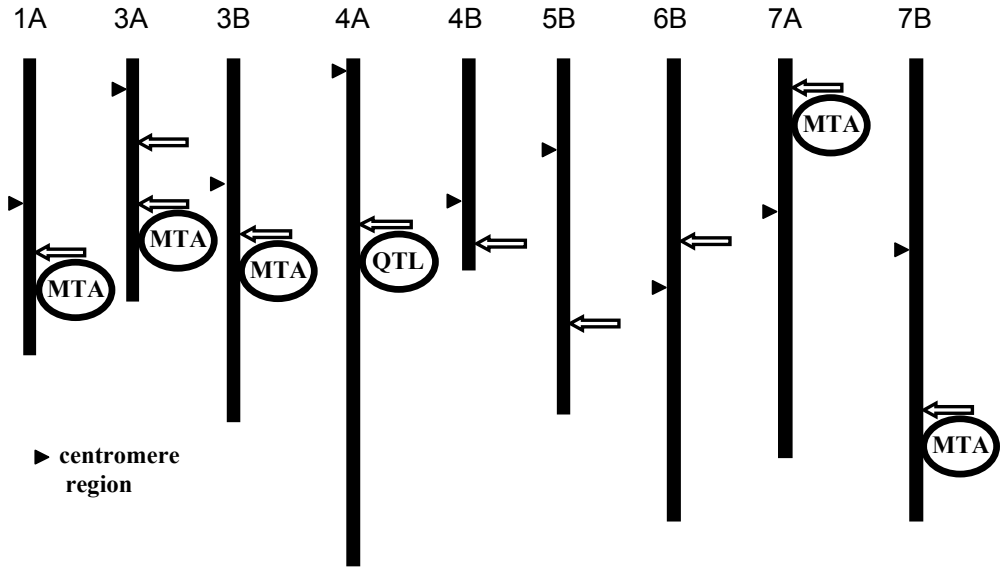

Figure 2. MTAs for grain yield marked by arrows. Positions of comparable MTAs/QTLs described earlier are indicated below arrows (elipses)

markers provided by Triticarte Pty. Ltd. for the analysis of historical multi-locational field trials considering grain yield and resistance to various foliar diseases.

Grain yield MTAs associated with identical DArT markers were detected on chromosomes 1AL, 3AL and 7BL. In addition, we identified grain yield MTAs associated with markers on chromosomes 3BL and 7AL not included in Crossa's study which however were in close distances to loci detected by Crossa *et al.* (2007). This large-scale coincidence was somewhat unexpected, given that the two studies had involved very different sets of germplasm. Grain yield has also been targeted by a number of conventional QTL mapping experiments, allowing for a comparison between MTA and known QTL locations. Thus, the location of *QYld.crc-4A* (4AL) (McCartney *et al.* 2005) coincide with the grain yield MTA detected here, whereas others do not.

CONCLUSIONS

Association-based mapping approaches in cereals appeared just recently and are rare in wheat. In contrast to bi-parental segregation-based mapping methods, considering two genotypes contrasting for the trait(s) of interest, here many individuals are analysed to detect marker-trait-associations. As expected numerous loci for the traits flowering time and grain yield were detected, frequently matching with already described major genes/QTLs or MTAs. Additionally, potential novel loci were identified that may help to better understand the architecture of complex genetic traits. The novel loci provide opportunities for further improvement of wheat, based on a marker approach. Plant breeders but also germplasm managers are challenged to un-look and process 'historical' data gathered from extensive trials.

Acknowledgements. The authors thank S. Dreisigacker, CIMMYT, Mexico, for providing the mapping data as well as I. Matthies, IPK Gatersleben, Germany, for her constructive discussions during the course of study.

References

Börner A. (2006): Preservation of plant genetic resources in the biotechnology era. Biotechnology Journal, **1**: 1393–1404.

Börner A., Schumann E., Fürste A., Cöster H., Leithold B., Röder M.S., Weber W.E. (2002): Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics, **105**: 921–936.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007): TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, **23**: 2633–2635.

Crossa J., Burgueno J., Dreisigacker S., Vargas M., Herrera-Foessel S.A., Lillemo M., Singh R.P., Trethowan R., Warburton M., Franco J., Rey-

nolds M., Crouch J.H., Ortiz R. (2007): Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics, **177**: 1889–1913.

FAO (2009): Draft second report on the state of the world's plant genetic resources for food and agriculture. Commission on Genetic Resources for Food and Agriculture. CGRFA-12/09/Inf.7 Rev.1, FAO, Rome.

Kobiljski B., Quarrie S.A., Denčić S., Kirby J., Iveges M. (2002): Genetic diversity of the Novi Sad wheat core collection revealed by microsatellites. Cell & Molecular Biology Letters, **7**: 685–694.

Kuchel H., Hollambi G., Langridge P., Williams K., Jefferies S.P. (2006): Identification of genetic loci associated with ear emergence in bread wheat. Theoretical and Applied Genetics, **113**: 1103–1112.

Laurie D.A., Pratchett N., Bezant J., Snape J.W. (1995): RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter × spring barley (*Hordeum vulgare* L.) cross. Genome, **38**: 575–585.

Leonova I., Pestsova E., Salina E., Efremova T., Röder M., Börner A. (2003): Mapping of the *Vrn-B1* gene in *Triticum aestivum* using microsatellite markers. Plant Breeding, **122**: 209–212.

McCartney C.A., Somers D.J., Humphreys D.G., Lukow O., Ames N., Noll J., Cloutier S., McCallum, B.D. (2005): Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 × 'AC Domain'. Genome, **48**: 870–883.

McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers J., Morris C., Somers D.J., Apples R., Devos K.M. (2008): Catalogue of gene symbols for wheat. Available at http://www.shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp

Miura H., Nakagawa M., Worland A.J. (1999): Control of ear emergence time by chromosome 3A of wheat. Plant Breeding, **118**: 85–87.

Neumann K., Kobiljski B., Denčić S., Börner A. (2011): Genome-wide association mapping: a case study in bread wheat (*Triticum aestivum* L.). Molecular Breeding, **27**: 37–58.

Pritchard J.K., Stephens M., Donnelly P.J. (2000): Inference of population structure using multilocus genotype data. Genetics, **155**: 945–959.

Quarrie S.A., Dodig D., Pekiç S., Kirby J., Kobiljski B. (2003): Prospects for marker-assisted selection of improved drought responses in wheat. Bulgarian Journal of Plant Physiology, XXIX (Special Issue): 83–95.

Tóth B., Galiba G., Fehér E., Sutka J., Snape J.W. (2003): Mapping genes affecting flowering time and frost resistance on chromosome 5B of wheat. Theoretical and Applied Genetics, **107**: 509–514.

UN Country code (2009): http://unstats.un.org/unsd/methods/m49/m49alpha.htm

Yu J., Buckler E.S. (2006): Genetic association mapping and genome organization of maize. Current Opinion in Biotechnology, **17**: 155–160.