

Detection of segregation distortion loci in triticale (*x Triticosecale* Wittmack) based on a high-density DArT marker consensus genetic linkage map

Alheit *et al.*

RESEARCH ARTICLE

Open Access

Detection of segregation distortion loci in triticale (*x Triticosecale* Wittmack) based on a high-density DArT marker consensus genetic linkage map

Katharina V Alheit¹, Jochen C Reif¹, Hans Peter Maurer¹, Volker Hahn¹, Elmar A Weissmann², Thomas Miedaner¹ and Tobias Wüerschum^{1*}

Abstract

Background: Triticale is adapted to a wide range of abiotic stress conditions, is an important high-quality feed stock and produces similar grain yield but more biomass compared to other crops. Modern genomic approaches aimed at enhancing breeding progress in cereals require high-quality genetic linkage maps. Consensus maps are genetic maps that are created by a joint analysis of the data from several segregating populations and different approaches are available for their construction. The phenomenon that alleles at a locus deviate from the Mendelian expectation has been defined as segregation distortion. The study of segregation distortion is of particular interest in doubled haploid (DH) populations due to the selection pressure exerted on the plants during the process of their establishment.

Results: The final consensus map, constructed out of six segregating populations derived from nine parental lines, incorporated 2555 DArT markers mapped to 2602 loci (1929 unique). The map spanned 2309.9 cM with an average number of 123.9 loci per chromosome and an average marker density of one unique locus every 1.2 cM. The R genome showed the highest marker coverage followed by the B genome and the A genome. In general, locus order was well maintained between the consensus linkage map and the component maps. However, we observed several groups of loci for which the colinearity was slightly uneven. Among the 2602 loci mapped on the consensus map, 886 showed distorted segregation in at least one of the individual mapping populations. In several DH populations derived by androgenesis, we found chromosomes (2B, 3B, 1R, 2R, 4R and 7R) containing regions where markers exhibited a distorted segregation pattern. In addition, we observed evidence for segregation distortion between pairs of loci caused either by a predominance of parental or recombinant genotypes.

Conclusions: We have constructed a reliable, high-density DArT marker consensus genetic linkage map as a basis for genomic approaches in triticale research and breeding, for example for multiple-line cross QTL mapping experiments. The results of our study exemplify the tremendous impact of different DH production techniques on allele frequencies and segregation distortion covering whole chromosomes.

Background

For scientists as well as commercial breeders the hexaploid man-made wheat-rye hybrid triticale (*x Triticosecale* Wittmack; $2n = 6 \times = 42$) is considered a promising crop with a broad genetic potential. Triticale is a partially outcrossing species and stands out due to a wide adaptation to abiotic stress conditions like

salinated or acid soil, aluminium toxicity, drought, and waterlogged soils [1,2]. Furthermore, the cereal has attained importance as feed stock owing to a valuable composition of amino acids and a stable performance in less productive environments [3]. Providing raw material for the generation of bioenergy and biofuels, triticale produces more biomass for a comparable grain yield than other crops and, therefore, can increase the industrially useable biomass without increasing competition with food production on arable land [4].

* Correspondence: tobias.wuerschum@uni-hohenheim.de

¹State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany

Full list of author information is available at the end of the article

Modern genomic approaches to enhance the breeding progress such as association mapping or genomic selection [5-7] require the availability of high-quality and high-density genetic linkage maps. For triticale, only one genetic map based on 73 DH lines and 356 markers with an average map density of 6.9 cM has been published so far [8]. Markers are, however, neither distributed homogeneously among the different genomes (50.7% located on R genome) nor on the chromosomes. For example only two markers are located on each of the 1A, 4A, and 3B chromosomes. Therefore, a highly saturated genetic linkage map for triticale is urgently required to enable genomics research and knowledge-based breeding.

A prerequisite for the construction of genetic linkage maps is the availability of polymorphic molecular markers. Diversity Arrays Technology (DArT) markers [9] have been identified as a valuable tool in cereals and have been employed successfully to create linkage maps of the triticale parents wheat [10-14] and rye [15]. Badea et al. [16] have recently reported the development of a triticale-specific DArT array combining markers developed in wheat, rye and triticale.

Consensus maps are genetic maps that are created by a joint analysis of the data from several segregating populations and different approaches are available for the construction of such maps. In the joint data process, segregation data from individual populations are pooled and loci orders and genetic map distances are computed based on mean recombination frequencies and combined LOD scores (implemented in JoinMap[®] 4) [17]. Merged linkage maps not only provide a mean to assess associations among individual linkage maps but are also the basis for QTL studies in multiple segregating populations [18].

The phenomenon that alleles at a locus deviate from the Mendelian expectation has been defined as segregation distortion [19] and has been described in many species, for example maize [20], rice [21], tomato [22] and alfalfa [23]. Sample size and genotyping errors are some non-biological factors that can contribute to segregation distortion. Biologically, segregation distortion can be due to selection among gametes and/or zygotes [24]. If an allele at a locus diminishes gametic or zygotic fitness, then that locus and other loci linked to it will deviate from the expected Mendelian segregation ratio. Biological segregation distortion will, therefore, always effect a cluster of markers within the chromosomal region surrounding the segregation distortion locus [25-27]. In general, the contribution of different factors to segregation distortion may vary in different populations [21]. Several studies have identified QTL that affect DH production and these can lead to segregation distortion in regions surrounding these QTL [28]. For doubled

haploid (DH) plant development several techniques are available which differ in (i) the sex of the source material (female vs. male gametes), (ii) the genesis of plantlets (zygotic embryos vs. embryoid bodies from microspores) and (iii) growth conditions in *in vitro* culture. The study of segregation distortion is, therefore, of particular interest in DH populations of diverse origin.

The objectives of this study were to (i) develop a high-quality saturated consensus linkage map combining DArT marker data from six triticale mapping populations, (ii) evaluate the colinearity and thus the reliability of the component maps by a comparison with the consensus map, and (iii) assess the presence and the extent of segregation distortion in the six triticale mapping populations developed by different approaches.

Results

Genetic diversity analysis

This study was based on six segregating populations derived from nine parental lines of which three were used as a common parent each contributing to two populations (Table 1). The DArT markers from the integrated map were used to assess the genetic similarity between the parental lines in a principal coordinate analysis based on the modified Rogers' distances of the individuals. The first two principal coordinates explained 42.8% and 20.7% of the total variation, respectively (Figure 1). The principal coordinate analysis revealed different degrees of relatedness but none of the parental lines was clearly separated from the others. The genetic distances among the six segregating populations revealed that populations F2_LxT and DH_LxA had the highest degree of genetic similarity (0.79) whereas populations EAW74 and DH07 were most distant (0.53) (Table 2).

DArT marker linkage maps of individual populations

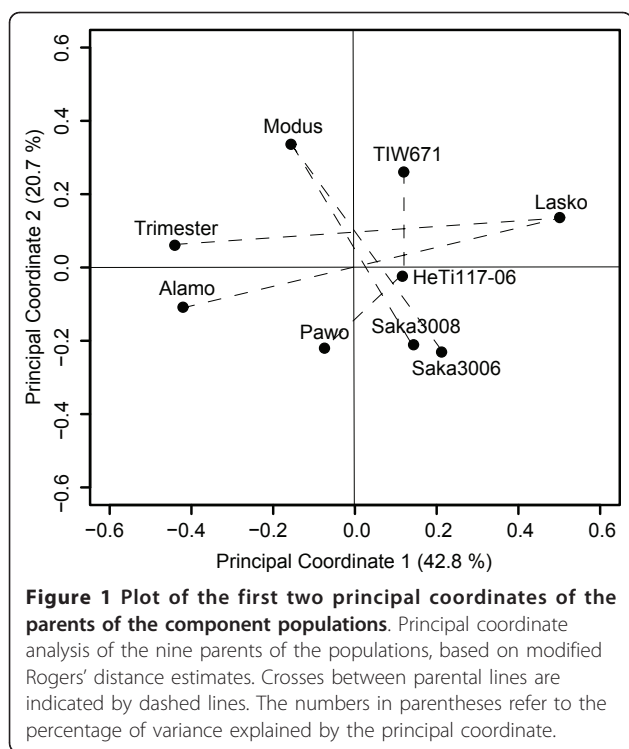
The component maps were constructed from six datasets containing between 114 (F2_LxT) and 200 (EAW74, EAW78) individuals and between 510 (F2_LxT) and 1244 (DH06) markers (Table 1). Out of these initial datasets 481 (F2_LxT) to 1209 (DH06) loci could be mapped on the 21 chromosomes (Table 3). The total

Table 1 Description of the mapping populations

Pop ^a code	Pop ^a pedigree	Pop ^a type	Pop ^a size	Markers ^b
DH06	Modus × Saka3006	DH	131	1244
DH07	Modus × Saka3008	DH	120	1064
EAW74	HeTi117-06 × Pawo	DH	200	713
EAW78	HeTi117-06 × TiW671	DH	200	673
DH_LxA	Lasko × Alamo	DH	146	979
F2_LxT	Lasko × Trimester	F ₂	114	510

^a Pop = population.

^b Number of markers polymorphic in this population after quality pre-selection.



lengths of the six resulting linkage maps varied between 1745.0 and 3270.9 cM. To estimate average marker densities, each group of co-segregating markers was considered as single marker (unique locus) to avoid artifacts leading to an overestimation of the actual values. Depending on the respective map length, single marker position counts from 435 (F2_LxT) to 755 (DH06) led to an average density of one unique marker locus every 4.0 (F2_LxT) to 5.2 cM (EAW78). With genetic map distances between unique loci varying from 0.1 to 111.6 cM, DArT markers were not evenly distributed along the chromosomes and in addition, showed cluster formation in different regions of the chromosomes. Ninety percent (2861 out of 3186) of the intervals between adjacent markers, however, were smaller than 10 cM and only 4.4% (140 out of 3186) exceeded 20 cM.

Table 2 Common markers and genetic similarity between mapping populations

	DH06	DH07	EAW74	EAW78	DH_LxA	F2_LxT
DH06		681	293	277	353	129
DH07	0.72		260	276	263	135
EAW74	0.57	0.53		254	244	104
EAW78	0.57	0.54	0.65		191	81
DH_LxA	0.65	0.64	0.66	0.64		204
F2_LxT	0.60	0.62	0.64	0.64	0.79	

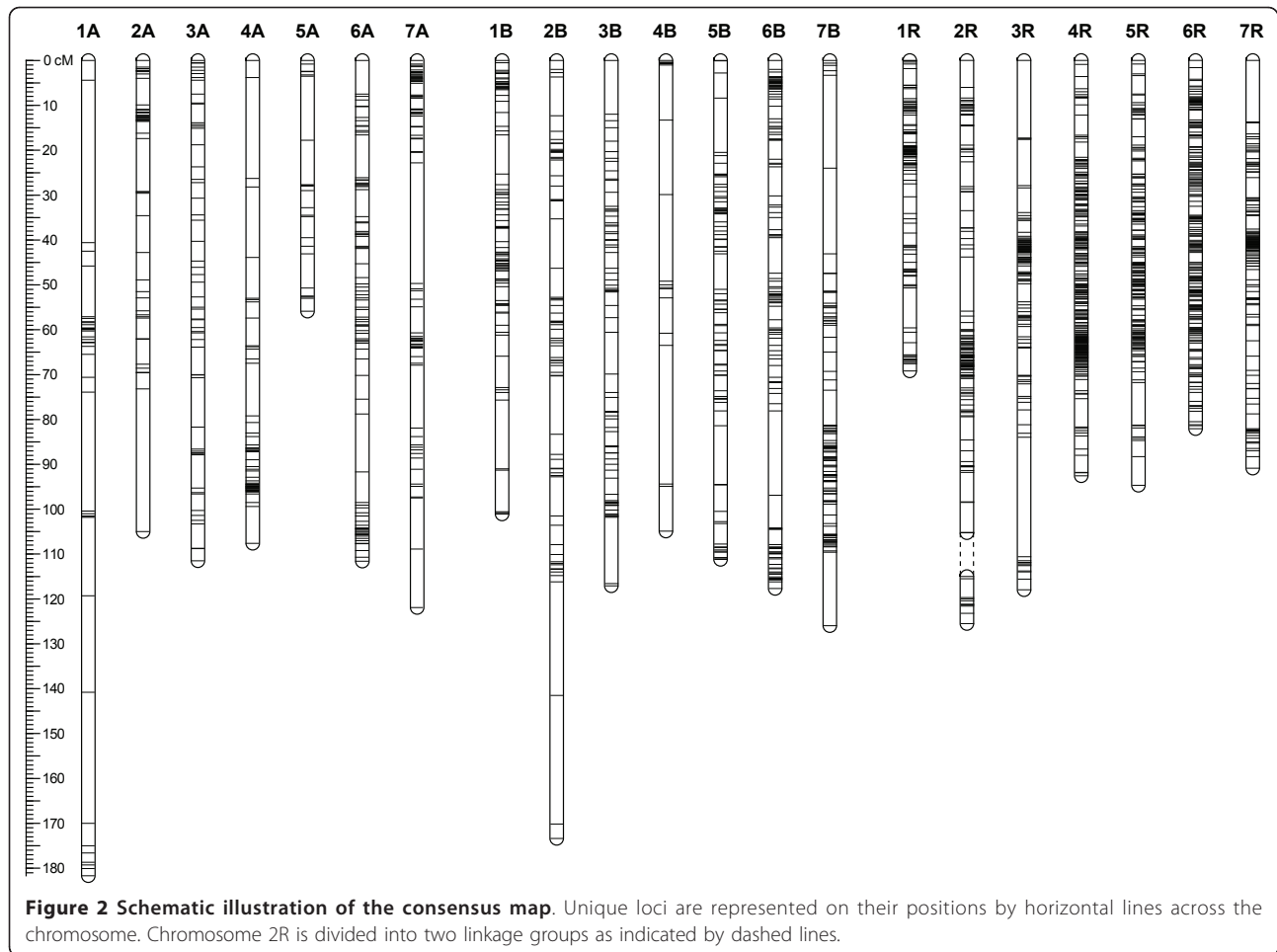
Number of markers common between component maps is shown above the diagonal and genetic similarity between the component mapping populations below the diagonal.

A comparison of the genetic maps with regard to the three triticale genomes, A, B and R, revealed that the A genome chromosomes showed the lowest marker coverage with an average number of 146.7 loci (109.2 unique) and an average marker distance of 6.5 cM. In contrast, the R genome chromosomes containing 451.7 loci (279.3 unique) and a map length of on average 879.3 cM featured the highest marker density with a mean genetic map distance of 3.1 cM between unique loci (Table 3).

A total of 81 (EAW78 and F2_LxT) to 681 markers (DH06 and DH07) were in common between pairs of component maps and we observed a high correlation ($r = 0.72$; $P = 0.02$) between the number of common markers and the genetic similarity between the populations (Table 2).

A consensus genetic linkage map from the combined datasets

To construct the triticale DArT marker consensus linkage map, the datasets containing the markers mapped on the component maps were merged for a joint analysis. The final consensus map incorporated 2555 markers mapped to a total of 2602 loci of which 1929 loci (74.1%) were unique (Table 3 Figure 2). Among the 2602 mapped loci, 1553 were computed on the basis of information derived from at least two mapping populations and 53 loci were based on five or six mapping populations. A subset of 47 out of the 2555 DArT markers (1.8%) mapped to two different loci in the consensus map. In more than 80% of the cases the genetic map positions of the different copies of these markers were found to be located either on a different chromosome within the same genome, or on a homeologous chromosome. The consensus map spanned a total length of 2309.9 cM resulting in an average marker density of one unique locus every 1.2 cM. Chromosome sizes ranged from 55.9 cM (chromosome 5A) to 181.7 cM (1A). Chromosome 2R was composed of two linkage groups 2R-1 and 2R-2 which could not be integrated applying the chosen mapping parameters. Linkage groups 2R-1 and 2R-2 covered 105.3 cM and 10.5 cM, respectively, with a total of 151 mapped loci of which 123 were unique. On the consensus map the number of loci per chromosome ranged from 21 (17 unique) on 4B to 344 (231 unique) on 4R with an average of 123.9 loci (91.9 unique) per chromosome. The average marker density based on unique loci ranged from 0.4 (4R, 6R) to 4.9 cM (1A). In the consensus map the largest gap between single marker positions spanned 36.2 cM and 99% of the intervals were smaller than 10 cM. Also in the consensus map, clustering of DArT markers became evident in different regions of the chromosomes and to a greater extent on chromosomes of the A and B



genomes (Figure 2). Furthermore, markers were not distributed equally among the A, B and R genomes. The number of mapped loci increased from the A genome with 430 (372 unique) to 664 (515 unique) on the B genome. The R genome featured the highest number of mapped loci totaling 1508 loci (1042 unique) with a total length of 663.3 cM and also the highest genetic map density among the triticale genomes, with an average marker density of 0.6 cM. We found that 16 DArT markers originating from the wheat D genome could be mapped on various chromosomes of the consensus map. Key data of the consensus map useful for further research including marker names, mapped chromosomes and positions are available as additional file (Additional file 1).

Comparison of the consensus map with the individual linkage maps

To evaluate the quality of the consensus linkage map, we assessed the consistency of locus order between the consensus map and the component maps. For this comparison, the marker positions of the consensus map were

plotted against the positions in the individual component maps separately for each chromosome (Figure 3). In general, locus order was well maintained between the consensus linkage map and the component maps. We observed few groups of loci, however, for which the colinearity between the component maps and the consensus map was uneven. On chromosome 1A of the DH07 map, for example, loci in the interval between 35-50 cM showed the same sequence but reverse orientation compared to the consensus map. Similarly, this was observed for chromosome 3B of the F2_LxT population (interval 90-125 cM) and for chromosomes 6A and 6R of population EAW78 (intervals 150-170 and 200-250 cM, respectively). For chromosomes 3R and 6A of the DH06 and DH07 specific maps the colinearity plots revealed that loci on the 0-10 cM and 0-40 cM intervals, respectively, were located at the opposite end of the chromosome compared to the consensus map. The locus sequence in these intervals, however, was well conserved. Highly dense marker regions of the consensus map, mainly found on chromosomes of the R genome, exhibited partial marginal shifts in locus order compared to the

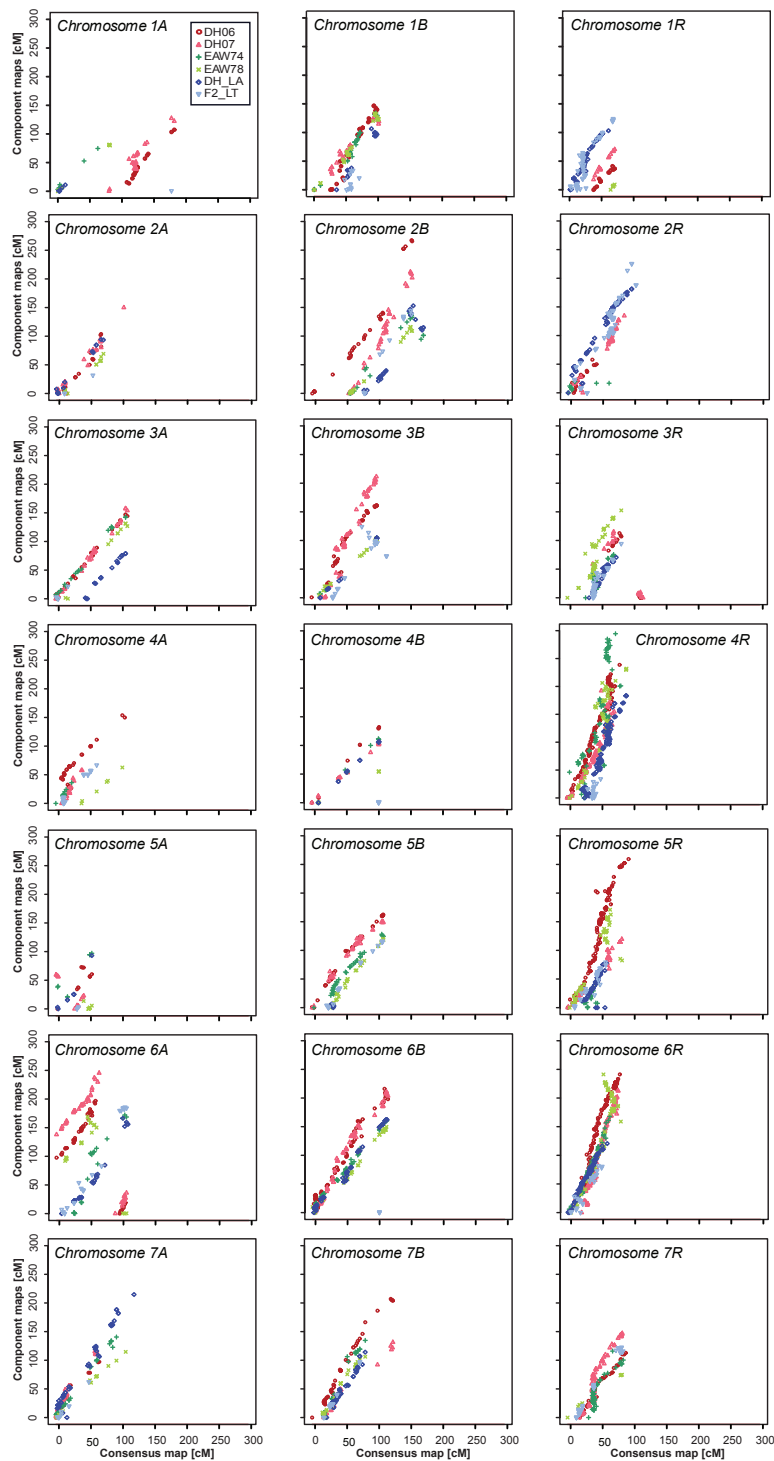


Figure 3 Comparison of loci positions in component maps and the consensus map. DH06, DH07, EAW74, EAW78; DH_LxA, and FZ_LxT are indicated by red circles, pink triangles, green crosses, light-green exes, blue diamonds and light-blue triangles, respectively.

component maps. On a more global level, the colinearity plots revealed disparities in length between the component maps and the consensus map. In the consensus map identical pairs of loci resulted in shorter genetic map distances condensing most chromosomes.

Segregation distortion

Among the 2602 loci mapped on the consensus map, 886 (34.0%) showed distorted segregation ($P < 0.01$) in at least one of the individual mapping populations, whereas only 311 of these loci (11.9%) deviated in two or more

populations. The rate of distorted markers in the individual populations averaged 19.9% and totalled 11.5% (156 out of 1352) in DH06, 8.3% (98 out of 1187) in DH07, 33.6% (329 out of 980) in EAW74, 34.3% (357 out of 1042) in EAW78, 13.2% (145 out of 1102) in DH_LxA and 27.2% (161 out of 592) in F2_LxT. On the level of single chromosomes, the proportion of markers deviating from the expected segregation ratio ranged from 0 to 100%. Markers on chromosomes 5A (DH06) and 1A, 3A, 7A, 4B and 7R (DH_LxA) followed the expected 1:1 segregation ratios, whereas all markers on chromosome 1R (EAW74) showed a significant deviation. As shown in Figure 4, there were several chromosomes with large regions in which the markers exhibited a distorted segregation pattern, namely in population DH06 on chromosome 7R, in populations EAW74 and EAW78 on chromosomes 2B, 3B, 1R, 4R and 7R, and in population DH_LxA on chromosome 2R. The distorted regions exhibited a pattern with slight segregation distortion at the flanks, which increased towards more highly distorted loci in the centre. In addition, we observed evidence for segregation distortion between pairs of loci caused either by a predominance of parental or recombinant genotypes (Figure 5).

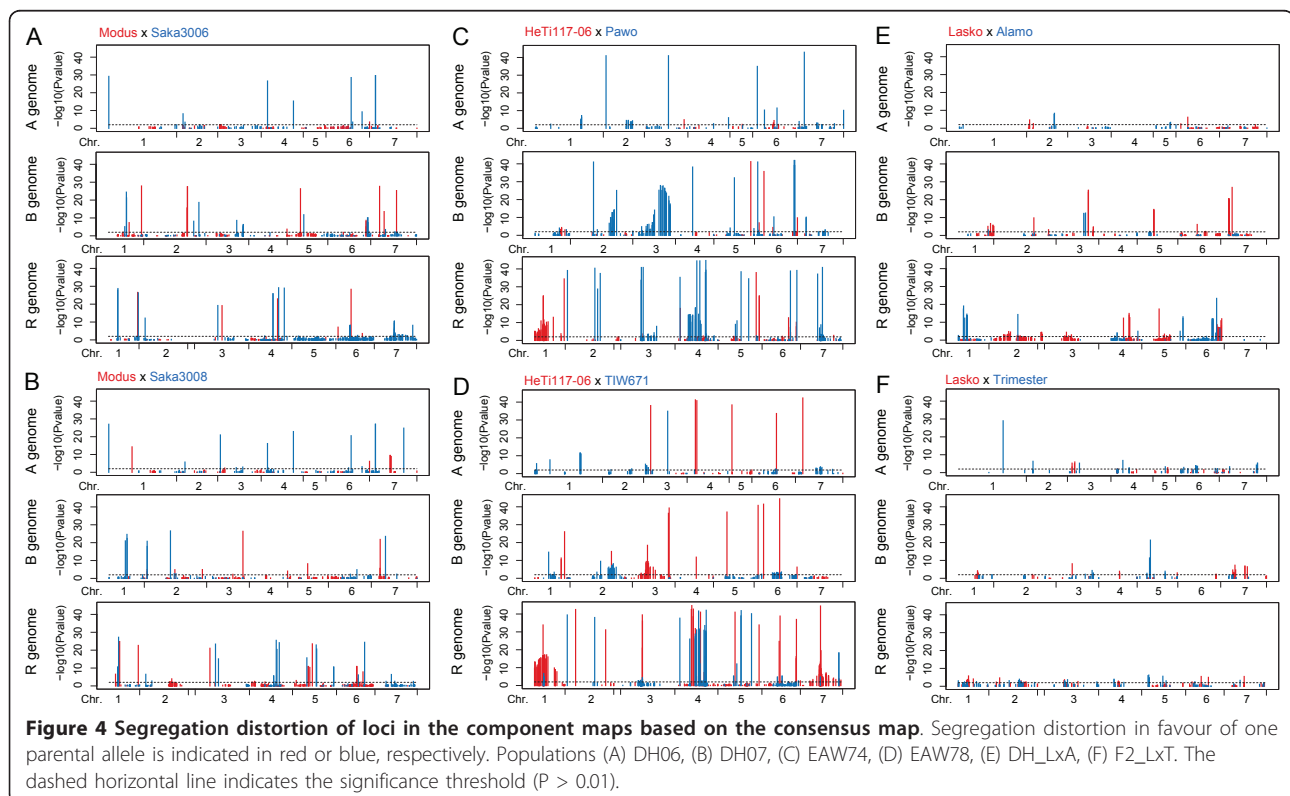
Discussion

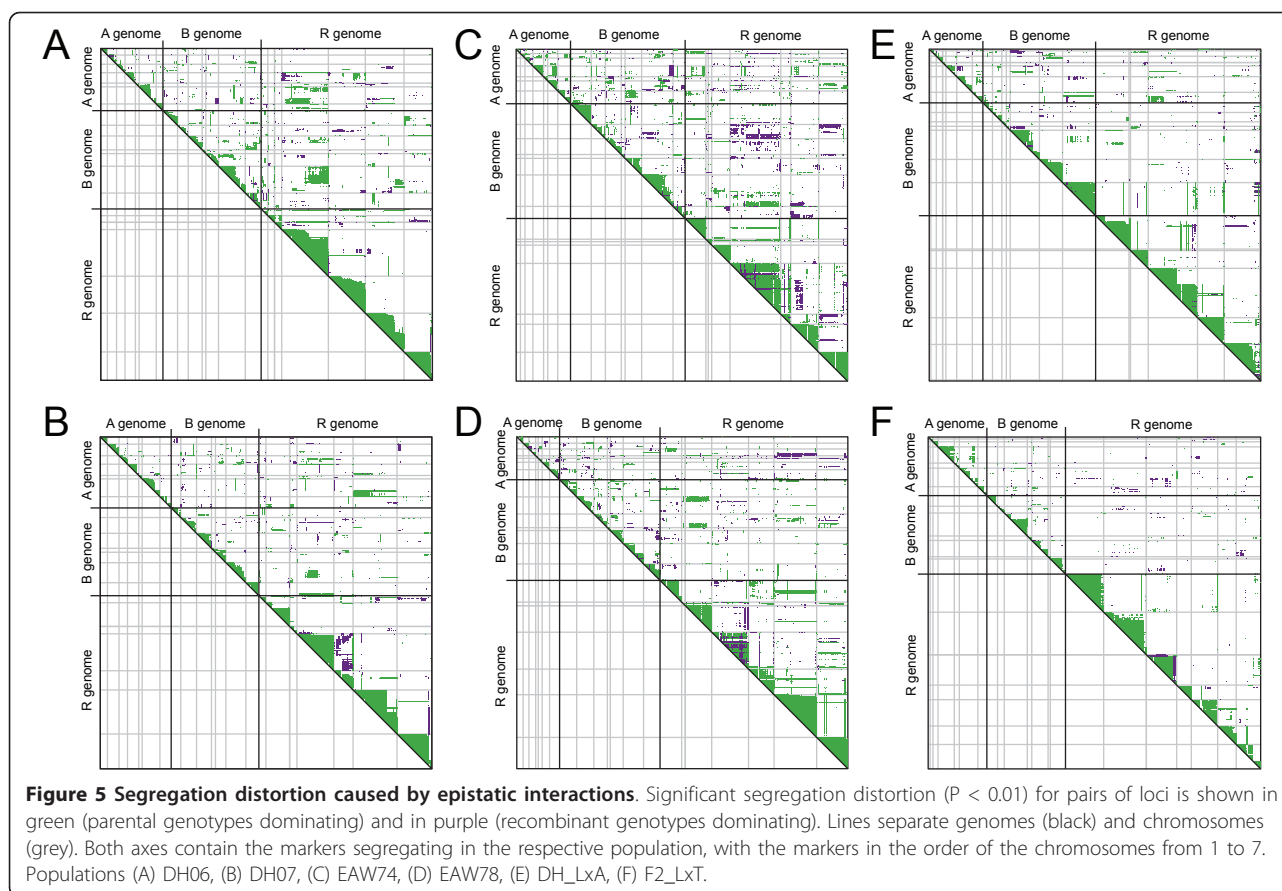
Genome-wide high-density genetic linkage maps specific to a species are both useful and essential for a number

of reasons. These include the identification of marker-trait associations via linkage analysis and association mapping, for isolation of genes by map-based cloning, for comparative mapping and for exploration of genome organisation [29,30]. In our study, we have created the first high-density consensus genetic linkage map integrating DARt marker data from six triticale mapping populations.

Evaluation of method and test conditions

In this study, the integrated genetic linkage map was calculated by using the joint data approach and regression mapping algorithm of JoinMap[®] 4 [17]. For effective linkage map pooling (and bridging) the individual mapping populations should share some common genetic bases (populations possessing a common parent for example) or common statistics (similar linkage information), but genetic map construction also relies on genome variation at loci detectable by molecular markers [31]. Our study was based on segregation data from six partially connected mapping populations derived from nine triticale elite breeding lines (Table 1). In reference to the first two principal coordinates the parental lines of each cross were derived from distinct quadrants (Figure 1), generating a maximum amount of genetic diversity. On the other hand, genetic similarity values between individual populations and the amount





of markers found to be in common between them indicate a certain connectedness between the populations (Table 2). These findings clearly show that the plant material underlying this study meets both requirements for linkage map construction and pooling. That is, sufficient variation for polymorphic segregation data as well as common genetic bases for a statistically firm integration.

Genetic linkage maps are established by heuristical algorithms and statistical approaches and thus, have an inherent statistical error. This holds true for the estimation of the recombination frequencies as well as for the integration of data from several populations. Sample size is crucial for genetic map construction as it affects the power of linkage detection and the accuracy of recombination fraction estimation [31]. Random variation and potential biological variation can cause differences in estimated pairwise distances between individual populations, particularly if populations are of small size [17]. We investigated six mapping populations incorporating 114 to 200 progenies, and a total sample size of 911 individuals (Table 1) which is similar to recent studies in rye [32], sorghum [33], red clover [34] and barley [35]. We observed differences in the estimation of

recombination frequency between identical pairs of markers in different populations [31] and found that for intervals up to 10 cM this heterogeneity increased successively and declined again for larger genetic distances (Additional file 2). The extent of heterogeneity between our individual mapping populations was expected due to both the effect of random sampling and biological variation which is occurring regardless of common genetic bases among populations.

In conclusion, the mapping populations used in our study were well suited for reliable consensus linkage map building. However, our results also highlight the fact that consensus maps always constitute a compromise which must be kept in mind.

Distribution of DArT markers

During the development and assessment of DArT markers in triticale, a larger number of polymorphic markers originated from the rye genome and more wheat markers from the B genome [16]. With 57.8% of markers mapped to the R genome, 25.5% to the B genome and only 16.5% to the A genome (Table 3) we could verify these results in an applied mapping experiment with a large number of populations and individuals.

Similar results have been found in studies of wheat for DArT [10-12] as well as other marker types [36-39]. We, therefore, conclude that the bias observed in our study is not attributable to the mapping individuals or the type of marker. Instead it may be due to the design of the triticale DArT array and the number of markers originating from the different genomes, but is also likely to reflect the different polymorphic nature of the A, B, and R genomes [16]. In our study, 1.8% of the DArT markers mapped to two different loci on the consensus map, but never on the same chromosome (Additional file 1). In terms of the ratio of markers that occur in a multicopy manner, our results agree well with those reported for hexaploid wheat (2%), barley (1.4%) and sorghum (1.8%) [10,33,35]. This may be attributable to the polyploid nature of hexaploid triticale having an impact on the accuracy of DArT markers due to alternative binding sites on homeologous chromosomes or may be ascribed to paralogous sequences. Molecular markers are known for their tendency to cluster, caused either by an unbalanced distribution of recombination events along chromosomes or an unequal representation of chromosomal regions on the genotyping array [33]. In accordance with this expectation we found that DArT markers clustered in several chromosomal regions (Figure 2). A possible explanation for regions with higher marker density on chromosomes could be that recombination occurs more frequently in gene-rich regions [40] which are present in clusters comprising physically small chromosomal regions and account for only 5-10% of the wheat genome [41,42]. The observed gaps in the consensus map may, on the other hand, be caused by identity-by-descent of the parental genotypes in these genomic regions. Taken together, clustering of tightly linked loci and gaps with low marker density in the consensus map either reflect the genetic situation in triticale or are due to specific properties of the applied DArT markers (e.g. complexity reduction step, or redundant clones). Further research including alternative high-density marker systems, e.g. SNPs, will help to address this question.

Consensus map features

Marker coverage and genetic map density are influenced by many criteria such as genome length, number of markers, distribution of markers and crossovers in the genome, mapping population size and mapping strategy [31]. As a result of the integration of datasets from six mapping populations our final triticale consensus map incorporated 2602 loci (74.1% unique) covering 2309.9 cM (Table 3). The previously published triticale genetic linkage map [8] comprised 356 markers (AFLP, RAMP, RAPD, SSR) and spanned 2465.4 cM. Thus, despite different marker types being used, these results agree well

in regard to computed map lengths for the triticale genome.

Likewise our results are in good accordance with the reported map length of 2383 cM based on 339 DArT markers for the related hexaploid bread wheat (AABBDD) [10]. In our study the linkage groups of the A and B genome together covered 1646.6 cM and those of the R genome 663.3 cM. Contrary to these results, the published genetic maps of durum wheat (AABB) [12] and rye (RR) [15] based on DArT markers spanned 2022 and 3144.6 cM, respectively. As mapping functions were similar in the studies, the disparities may be explained by a different number of markers mapped and/or the different mapping algorithms applied. Referring to unique markers, many appeared redundant in component maps and were just 0.1 cM apart after integration. This raises the question whether these loci are really distinct from each other or are just a product of statistics and therefore should not be regarded as unique in further QTL mapping studies.

The average density in our consensus map was one unique marker every 1.2 cM, and 98% of all intervals between adjacent loci were smaller than 10 cM. This is sufficient for QTL mapping and many modern genomics approaches [43]. It may, however, be worthwhile amending larger gaps by target-oriented employment of additional markers. Thus, we conclude that through integration of datasets from six mapping populations we were able to improve both the density and the quality of the component maps up to the final high-density DArT marker triticale consensus map.

Colinearity

One possible method to assess the quality of a consensus map is to compare the locus arrangement of the consensus map, which was optimised at the multi-population level, with the arrangement of loci in the component maps (each one optimised independently) [35]. A consistent order is hypothesised if the markers identify identical chromosomal locations and if there are no incorrect or missing scores [44]. In addition, this colinearity comparison can identify chromosome rearrangements in individual populations. Our tests for colinearity resulted in an overall good consistency affirming the high quality of our consensus map (Figure 3). We also found, however, regions where groups of neighbouring loci showed identical order of loci but inversion within a linkage group or even positioning of a region with conserved order at the opposite end of the chromosome. Such inconsistencies were reported for other species including red clover [34] and sorghum [33]. In triticale, chromosomal rearrangements are known to occur [45]. Thus, these inversions of marker order or positioning could reflect real genetic events such as small chromosome rearrangements or, as

they occur mostly after gaps, they could also be caused by statistical uncertainty due to many weak linkages contributing to the adjustment. Furthermore, marginal shifts in locus order were found in regions with highly dense markers. Similar results were reported before in several mapping experiments [32-35,46]. Despite a certain heterogeneity of recombination frequencies between mapping populations this must mainly be attributed to the dependency of estimated gene orders on sample size [31]. Especially for the high-density regions, extremely large mapping populations would be required to resolve the correct order of markers.

The colinearity plots revealed that respective linkage groups were generally longer in the component maps than in the consensus map and this effect was even larger in denser linkage groups. The application of different algorithms (maximum likelihood or regression approach) during component and consensus map construction has been reported to affect map lengths, despite the same mapping function [32,34,46]. Another explanation could be that the condensed map length may be the intended outcome of the addition of more markers during the integration process [32].

Segregation Distortion

Segregation distortion is known to strongly impact genetic map construction and QTL mapping [26,27] but distorted markers may also be beneficial for QTL mapping if handled properly [47]. Whereas highly deviating markers cannot be placed in the respective component maps, they can be included in the consensus map through integration of unbiased information available from other populations without segregation deviation. Our experimental design with multiple segregating populations thus offered an excellent basis for the evaluation of segregation distortion and the mapping of segregation distortion QTL.

The component maps of two populations were completely lacking certain linkage groups (2A and 1R in EAW74, and 2R in EAW78) (Table 3). As the three linkage groups were well covered in other populations and the same markers were also positively scored in the populations with the lacking chromosomes we can exclude a scarcity of markers. The majority of the markers on these chromosomes (79 and 100% of the markers mapped on chromosomes 2A and 1R in the consensus map, respectively), however, showed significant segregation distortion in population EAW74. This illustrates the possible consequences of segregation distortion which not only affects genetic map distances and ordering of loci, but can even result in complete chromosomes being absent from genetic maps.

The populations underlying our study were five DH and one F₂ population. The DH populations were

produced by different methods, either by maize pollination of the oocytes (DH06, DH07, DH_LxA) or by microspore culture (EAW74, EAW78) and were not pre-selected for any trait. We observed segregation distorted regions caused by biological factors (regions of distorted markers which had the same skew direction, distinguishable from deviating loci scattered apart along the chromosomes which are likely due to genotyping errors or non-functional markers) only in DH populations and mainly in microspore culture derived populations (Figure 4). The patterns observed in segregation distortion regions can be explained by both the distance between the markers and the segregation distortion loci linked to them and the effects of those loci [48]. Our test for segregation distortion in the microspore derived populations resulted in clusters of distorted markers on chromosomes 2B, 3B, 1R, 4R, and 7R. In studies of wheat DH populations, chromosome 2B was reported to harbour QTL responsible for green spot initiation and plant regeneration [49] and a different type of *in vitro* culture response in anther culture [28]. Furthermore, QTLs located on chromosomes 1R, 4R and 7R were previously reported to have an effect on the yield of green plantlets from anthers in culture and embryo induction (7R) in triticale [8]. Segregation distortion on chromosome 7R was, besides EAW74 and EAW78, also observed in DH06. This chromosome has been implicated in the selection of zygotes or female gametes in rye [50]. To our knowledge no QTL affecting segregation distortion have been described for chromosomes 3B (EAW74 and EAW78) and 2R (DH_LxA) yet. Due to the consistent occurrence in both microspore derived DH populations we assume that this region located on chromosome 3B harbours a novel QTL responsible for *in vitro* or androgenetic response in triticale.

Epistasis refers to interactions between two or several loci [51,52] and has recently been shown to contribute to segregation distortion [53,54]. In accordance with this we observed epistatic interactions involved in segregation distortion (Figure 5). These epistatic interactions point towards selection for specific allele combinations for *in vitro* or androgenetic response in triticale.

The DH technology has become an indispensable part of both research and breeding of triticale and many other agronomically important crops. Depending on the effect of the segregation distortion locus, it can influence the allele frequencies on an entire chromosome, including all genes involved in the expression of important agronomic traits. If the segregation distortion locus and the QTL for the agronomic trait are linked in repulsion, the agronomic QTL will be underrepresented or in the most extreme case absent from the population. The same holds true for the introgression of traits, if the QTL are by chance located on chromosomes harbouring

segregation distortion loci. The relatively high number of segregation distortion loci identified in our study highlights this problem both for research and applied breeding. Further characterisation or even identification of the nature of segregation distortion loci may facilitate solving these issues.

Conclusions

We have constructed the first DArT marker consensus genetic linkage map for triticale by integrating segregation data from six mapping populations. The colinearity of the consensus map was well maintained and it is, therefore, sufficiently reliable for use in multiple-line cross QTL mapping experiments and in addition may serve as a reference for genetic maps created from other triticale germplasm. The results of our study underpin the impact of different DH production techniques on segregation distortion and allele frequencies covering whole chromosomes. In this context we identified a previously unknown region located on chromosome 3B likely to be responsible for *in vitro* or androgenetic response in triticale. Our results imply that caution must be exerted when DH populations are utilised in research or applied breeding.

Methods

Plant material and DNA extraction

This study was based on 911 triticale lines (*Triticosecale* Wittmack L.) from six mapping populations derived from nine parental lines (Table 1). Populations DH06 (131 individuals), DH07 (120), EAW74 (200), EAW78 (200), and DH_LxA (146) were doubled haploid (DH) lines whereas the F₂_LxT population (114 individuals) was an F₂ population. Leaf tissue was harvested around five leaf stage from single plantlets and dried in silica gel. High-quality genomic DNA for genotyping was isolated from 20-25 mg of dried leaf tissue according to a modified CTAB method [55] and adjusted to a concentration of ca. 50 ng/μl for marker analysis.

Marker and molecular analysis

Marker data for genetic mapping were obtained by DArT genotyping of all samples [9]. DArT genotyping of the individuals used in this study was carried out by Triticarte Pty Ltd, Yarralumla, ACT, Australia [56] with the current triticale array. The nine parental varieties were also DArT genotyped as described above. Associations among the nine parental lines were analysed by applying a principal coordinate analysis [57] based on the modified Rogers' distances of the individuals [58]. Similarity among the populations was estimated as one minus the modified Rogers' distances among the populations. Segregation distortion was calculated based on the *P* values obtained by a chi-square test and the

parent contributing the allele with the higher frequency is indicated in Figure 4. To identify loci with epistatic interactions causing segregation distortion, we tested all possible pairs of markers using Fisher's exact test. Principal coordinate analysis and segregation distortion computations were performed with the software packages Plabsoft [59] and R [60], respectively.

Genetic linkage map construction

For the construction of the genetic linkage maps, DArT markers polymorphic in a population were transformed into genotype codes according to the score of the parents. For quality filtering a pre-selection with regard to their segregation ratio was performed. Markers which significantly ($P \leq 0.001$) deviated from the expected 1:1 (DHs) and 3:1 (F₂) ratio in a chi-square test were excluded from further analyses. The genetic linkage maps were constructed with the software JoinMap[®] 4.0 [17]. Markers were assigned to linkage groups applying the independence LOD (logarithm of the odds) parameter with LOD threshold values ranging from 2.0 to 20.0. The test for independence is not affected by segregation distortion which allowed for the liberal level of significance in terms of deviation. Certain ungrouped markers were added to groups on a by-case basis according to the indicated strongest cross link (SCL) LOD values. Chromosome names and orientation were assigned to linkage groups based on a subset of markers in each linkage group for which the positions have recently been published ([10,15]; chromosome assignment by Triticarte Pty Ltd, Yarralumla, ACT, Australia [56]). The available information on the position of some of the DArT markers also allowed us to link chromosome regions that appeared unlinked at the LOD = 2.0 level. Markers causing suspect linkages due to an estimated recombination frequency > 0.6 were excluded in the particular population. During the calculation of the individual maps from the six populations the locus order within chromosomes and estimation of recombination frequencies were established employing the provided maximum likelihood algorithm with modified calculation settings. For an adjusted map order optimisation, chain length and stopping criterion were extended to 5000, the cooling control parameter was decreased to 0.0001. The maximum likelihood algorithm was used to establish the map order of the markers within a defined linkage group and the genetic distances in centimorgan (cM) values were output converted with Kosambi's mapping function [61]. After each run post-mapping quality filtering tools provided by JoinMap[®] 4 for the maximum likelihood method such as the plausible position matrix and the fit and stress monitoring were studied and markers causing a poorer fit were excluded. For the construction of the consensus linkage

maps quality filtered data sets from individual populations related to the same chromosomes were joined together in one data set. In JoinMap[®] the calculations of consensus maps are based on mean recombination frequencies and combined LOD scores of pairwise data from multiple populations. To screen for deviant pairs the heterogeneity test using a standard G^2 statistic was used. In order to be able to exclude differences more likely to be due to the effect of random sampling or technical or statistical failure, and hence provide a basis for adequate linkage map pooling without significant differences as postulated by [31], pairs of loci highly deviating in their estimated recombination frequencies ($P \leq 0.001$) were excluded from computation of the consensus linkage map. Thus, loci were not forfeit but placed more trustable by the use of multi-locus models.

Locus order and map distances were calculated using the regression mapping approach with the following settings: recombination frequency used < 0.49 , LOD > 1.0 , goodness-of-fit jump threshold = 5.0, number of added loci after which to perform a ripple = 1, mapping function = Kosambi, and third round = yes. After each run the post-mapping quality filtering tool indicating the mean chisquare-contributions was studied and markers causing a poorer fit were excluded. The graphical representation of the map was drawn using MapChart software [62].

Additional material

Additional file 1: Locus positions on the consensus map. Excel spreadsheet containing a list of all consensus map loci. Data include (i) the locus referred to as marker name plus chromosome it was derived from (NA not available), (ii) chromosome and position the locus was mapped to and (iii) multicopy marker details.

Additional file 2: Heterogeneity of distances between unique loci among component populations. Heterogeneity was tested by comparing observed numbers of recombinants (calculated from recombination frequency and LOD score) in the individual component populations with the expected numbers based on the mean recombination frequency [17]. P values denote values of the G^2 statistic for each pair of loci summed up over all component populations contributing information for this pair. Respective map distances are based on the consensus map.

Acknowledgements

This research is funded by the Federal Ministry of Education and Research (BMBF) under the promotional reference 0315414A. This publication reflects the views only of the authors. We acknowledge the handling of the funding by the Project Management Organisation Jülich (PTJ). We thank Angela Harmsen for excellent technical assistance in the laboratory and Agnes Rölfing-Finze, Hans Häge, Jacek Till and Justus von Kittlitz for outstanding work in the greenhouse and field. We are grateful to M. Tucker for carefully reading the manuscript. We greatly appreciate the helpful comments and suggestions of two anonymous reviewers.

Author details

¹State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany. ²Saatzucht Dr. Hege GbR Domäne Hohebuch, 74638 Waldenburg, Germany.

Authors' contributions

KVA carried out the mapping analyses, performed parts of the statistical analyses and drafted the manuscript. JCR participated in the design of the study and edited the manuscript. HPM and VH developed the DH06 and DH07 populations. EAW developed the EAW74 and EAW78 populations. TM developed the DH_LxA and F2_LxT populations. TW participated in the design of the study, performed parts of the statistical analyses and helped to draft the manuscript. All authors read and approved the final manuscript.

Received: 18 April 2011 Accepted: 28 July 2011 Published: 28 July 2011

References

1. Kuleung C, Baenziger PS, Dweikat I: **Transferability of SSR markers among wheat, rye, and triticale.** *Theor Appl Genet* 2004, **108**(6):1147-1150.
2. Oettler G: **The fortune of a botanical curiosity - Triticale: Past, present and future.** *J Agric Sci* 2005, **143**(5):329-346.
3. Tams SH, Bauer E, Oettler G, Melchinger AE: **Genetic diversity in European winter triticale determined with SSR markers and coancestry coefficient.** *Theor Appl Genet* 2004, **108**(7):1385-1391.
4. Pronyk C, Mazza G: **Optimization of processing conditions for the fractionation of triticale straw using pressurized low polarity water.** *Bioresour Technol* 2011, **102**(2):2016-2025.
5. Meuwissen THE, Hayes BJ, Goddard ME: **Prediction of total genetic value using genome-wide dense marker maps.** *Genetics* 2001, **157**(4):1819-1829.
6. Veyrieras J-B, Camus-Kulandaivelu L, Gouesnard B, Manicacci D, Charcosset A: **Bridging genomics and genetic diversity: Linkage disequilibrium structure and association mapping in maize and other cereals.** *Crop Sci* 2007, **47**(SUPPL DEC):60-71.
7. Jannink J-L, Lorenz AJ, Iwata H: **Genomic selection in plant breeding: From theory to practice.** *Briefings in Functional Genomics and Proteomics* 2010, **9**(2):166-177.
8. González JM, Muñoz LM, Jouve N: **Mapping of QTLs for androgenetic response based on a molecular genetic map of × Triticosecale Wittmack.** *Genome* 2005, **48**(6):999-1009.
9. Jaccoud D, Peng K, Feinstein D, Kilian A: **Diversity arrays: a solid state technology for sequence information independent genotyping.** *Nucleic Acids Res* 2001, **29**(4):e25.
10. Akbari M, Wenzl P, Caig V, Carling J, Xia L, Yang S, Uszyński G, Mohler V, Lehmensiek A, Kuchel H, Hayden MJ, Howes N, Sharp P, Vaughan P, Rathmell B, Huttner E, Kilian A: **Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome.** *Theor Appl Genet* 2006, **113**(8):1409-1420.
11. Semagn K, Bjørnstad Å, Skjenes H, Marøy AG, Tarkegne Y, William M: **Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population.** *Genome* 2006, **49**(5):545-555.
12. Mantovani P, Maccaferri M, Sanguineti MC, Tuberosa R, Catizone I, Wenzl P, Thomson B, Carling J, Huttner E, DeAmbrogio E, Kilian A: **An integrated DArT-SSR linkage map of durum wheat.** *Mol Breed* 2008, **22**(4):629-648.
13. Peleg Z, Saranga Y, Suprunova T, Ronin Y, Röder MS, Kilian A, Korol AB, Fahima T: **High-density genetic map of durum wheat × wild emmer wheat based on SSR and DArT markers.** *Theor Appl Genet* 2008, **117**(1):103-115.
14. Jing H-C, Bayon C, Kanyuka K, Berry S, Wenzl P, Huttner E, Kilian A, Hammond-Kosack KE: **DArT markers: Diversity analyses, genomes comparison, mapping and integration with SSR markers in Triticum monococcum.** *BMC Genomics* 2009, **10**:458.
15. Bolibok-Bragoszewska H, Heller-Uszyńska K, Wenzl P, Uszyński G, Kilian A, Rakoczy-Trojanowska M: **DArT markers for the rye genome - genetic diversity and mapping.** *BMC Genomics* 2009, **10**:578.
16. Badea A, Eudes F, Salmon D, Tuvesson S, Vrolijk A, Larsson C-T, Caig V, Huttner E, Kilian A, Laroche A: **Development and assessment of DArT markers in triticale.** *Theor Appl Genet* 2011.
17. Van Ooijen JW: **JoinMap[®] 4, Software for the calculation of genetic linkage maps in experimental populations.** Edited by: Kyazma BV. Wageningen, Netherlands; 2006:.
18. Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L: **Connected populations for detecting quantitative trait loci and testing for epistasis: An application in maize.** *Theor Appl Genet* 2006, **113**(2):206-224.

19. Sandler L, Hiraizum Y, Sandler I: Meiotic drive in natural populations of *Drosophila melanogaster*. I. The cytogenetic basis of segregation-distortion. *Genetics* 1959, **44**:233-250.
20. Lu H, Romero-Severson J, Bernardo R: Chromosomal regions associated with segregation distortion in maize. *Theor Appl Genet* 2002, **105**(4):622-628.
21. Xu Y, Zhu L, Xiao J, Huang N, McCouch SR: Chromosomal regions associated with segregation distortion of molecular markers in F₂ backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Molecular and General Genetics* 1997, **253**(5):535-545.
22. Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD: Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 1988, **335**(6192):721-726.
23. Brummer EC, Bouton JH, Kochert G: Development of an RFLP map in diploid alfalfa. *Theor Appl Genet* 1993, **86**(2-3):329-332.
24. Falconer DS, Mackay TFC: In *Introduction to Quantitative Genetics*. 4 edition. Edited by: Harlow, Essex. UK: Longmans Green; 1996.
25. Luo L, Xu S: Mapping viability loci using molecular markers. *Heredity* 2003, **90**(6):459-467.
26. Xu S, Hu Z: Mapping quantitative trait loci using distorted markers. *International Journal of Plant Genomics* 2009, **2009**.
27. Vogl C, Xu S: Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics* 2000, **155**(3):1439-1447.
28. Torp AM, Hansen AL, Andersen SB: Chromosomal regions associated with green plant regeneration in wheat (*Triticum aestivum* L.) anther culture. *Euphytica* 2001, **119**(3):377-387.
29. Varshney RK, Hoisington DA, Tyagi AK: Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol* 2006, **24**(11):490-499.
30. Varshney RK, Langridge P, Graner A: Application of Genomics to Molecular Breeding of Wheat and Barley. *Adv Genet* 2007, **58**:121-155.
31. Liu B-H: Statistical genomics: linkage, mapping, and QTL analysis. Edited by: Stern RB, Didier D, Boca Raton. CRC Press LLC; 1998.
32. Gustafson JP, Ma X-F, Korzun V, Snape JW: A consensus map of rye integrating mapping data from five mapping populations. *Theor Appl Genet* 2009, **118**(4):793-800.
33. Mace ES, Rami J-F, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR: A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biology* 2009, **9**:13.
34. Isobe S, Kölliker R, Hisano H, Sasamoto S, Wada T, Klimenko I, Okumura K, Tabata S: Construction of a consensus linkage map for red clover (*Trifolium pratense* L.). *BMC Plant Biology* 2009, **9**:57.
35. Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesná J, Cakir M, Poulsen D, Wang J, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A: A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* 2006, **7**:206.
36. Liu Y-G, Tsunewaki K: Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat. *Jpn J Genet* 1991, **66**(5):617-633.
37. Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW: A microsatellite map of wheat. *Genetics* 1998, **149**(4):2007-2023.
38. Peng J, Korol AB, Fahima T, Röder MS, Ronin YI, Li YC, Nevo E: Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: Genoma-wide coverage, massive negative interference, and putative Quasi-linkage. *Genome Res* 2000, **10**(10):1509-1531.
39. Chao S, Zhang W, Akhunov E, Sherman J, Ma Y, Luo M-C, Dubcovsky J: Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L.) cultivars. *Mol Breed* 2009, **23**(1):23-33.
40. Sandhu D, Champoux JA, Bondareva SN, Gill KS: Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. *Genetics* 2001, **157**(4):1735-1747.
41. Gill KS, Gill BS, Endo TR, Taylor T: Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics* 1996, **144**(4):1883-1891.
42. Gill KS, Gill BS, Endo TR, Boyko EV: Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics* 1996, **143**(2):1001-1012.
43. Piepho H-P: Optimal marker density for interval mapping in a backcross population. *Heredity* 2000, **84**(4):437-440.
44. Beavis WD, Grant D: A linkage map based on information from four F₂ populations of maize (*Zea mays* L.). *Theor Appl Genet* 1991, **82**(5):636-644.
45. Bento M, Pereira HS, Rocheta M, Gustafson P, Viegas W, Silva M: Polyploidization as a retraction force in plant genome evolution: Sequence rearrangements in triticale. *PLoS ONE* 2008, **3**(1).
46. Alsop BP, Farre A, Wenzl P, Wang JM, Zhou MX, Romagosa I, Kilian A, Steffenson BJ: Development of wild barley-derived DArT markers and their integration into a barley consensus map. *Mol Breed* 2011, **27**(1):77-92.
47. Xu S: Quantitative trait locus mapping can benefit from segregation distortion. *Genetics* 2008, **180**(4):2201-2208.
48. Chen J, Walsh B: Method for the mapping of a female partial-sterile locus on a molecular marker linkage map. *Theor Appl Genet* 2009, **119**(6):1085-1091.
49. Ben Amer IM, Korzun V, Worland AJ, Börner A: Genetic mapping of QTL controlling tissue-culture response on chromosome 2B of wheat (*Triticum aestivum* L.) in relation to major genes and RFLP markers. *Theor Appl Genet* 1997, **94**(8):1047-1052.
50. Hackauf B, Wehling P: Development of microsatellite markers in rye: map construction. *Plant Breeding and Seed Science* 2003, **48**(2):143-151.
51. Carlborg Ö, Haley CS: Epistasis: Too often neglected in complex trait studies? *Nat Rev Genet* 2004, **5**(8):618-625.
52. Carlborg Ö, Jacobsson L, Åhlgren P, Siegel P, Andersson L: Epistasis and the release of genetic variation during long-term selection. *Nat Genet* 2006, **38**(4):418-420.
53. Bombliès K, Lempe J, Eppe P, Warthmann N, Lanz C, Dangl JL, Weigel D: Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biology* 2007, **5**(9):1962-1972.
54. McDaniel SF, Willis JH, Shaw AJ: A linkage map reveals a complex basis for segregation distortion in an interpopulation cross in the moss *Ceratodon purpureus*. *Genetics* 2007, **176**(4):2489-2500.
55. Saghai-Marouf MA, Soliman KM, Jorgensen RA, Allard RW: Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 1984, **81**:8014-8018.
56. Triticarte Website. [http://www.triticarte.com.au/].
57. Gower JC: Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 1966, **53**:325-338.
58. Wright S: *Evolution and genetics of populations*. Chicago, The University of Chicago Press; 1978IV.
59. Maurer HP, Melchinger AE, Frisch M: Population genetic simulation and data analysis with Plabsoft. *Euphytica* 2008, **161**(1-2):133-139.
60. R Software. [http://www.R-project.org/].
61. Kosambi DD: The estimation of map distances from recombination values. *Ann Eugenetic* 1944, **12**(3):172-175.
62. Voorrips RE: Mapchart: Software for the graphical presentation of linkage maps and QTLs. *J Hered* 2002, **93**(1):77-78.

doi:10.1186/1471-2164-12-380

Cite this article as: Alheit et al.: Detection of segregation distortion loci in triticale (x *Triticosecale* Wittmack) based on a high-density DArT marker consensus genetic linkage map. *BMC Genomics* 2011 **12**:380.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

