

RESEARCH ARTICLE

Open Access

# Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and *Brachypodium* genomes

Pasqualina Colasuonno<sup>1</sup>, Mastrangelo Anna Maria<sup>2</sup>, Antonio Blanco<sup>1</sup> and Agata Gadaleta<sup>1\*</sup>

## Abstract

**Background:** The importance of wheat to the world economy, together with progresses in high-throughput next-generation DNA sequencing, have accelerated initiatives of genetic research for wheat improvement. The availability of high density linkage maps is crucial to identify genotype-phenotype associations, but also for anchoring BAC contigs to genetic maps, a strategy followed for sequencing the wheat genome.

**Results:** Here we report a genetic linkage map in a durum wheat segregating population and the study of mapped DArT markers. The linkage map consists of 126 gSSR, 31 EST-SSR and 351 DArT markers distributed in 24 linkage groups for a total length of 1,272 cM. Through bioinformatic approaches we have analysed 327 DArT clones to reveal their redundancy, syntenic and functional aspects. The DNA sequences of 174 DArT markers were assembled into a non-redundant set of 60 marker clusters. This explained the generation of clusters in very small chromosome regions across genomes. Of these DArT markers, 61 showed highly significant (Expectation < E-10) BLAST similarity to gene sequences in public databases of model species such as *Brachypodium* and rice. Based on sequence alignments, the analysis revealed a mosaic gene conservation, with 54 and 72 genes present in rice and *Brachypodium* species, respectively.

**Conclusions:** In the present manuscript we provide a detailed DArT markers characterization and the basis for future efforts in durum wheat map comparing.

**Keywords:** Wheat, DArT marker, Genetic map, Synteny

## Background

Wheat (*Triticum* spp.) is among the most widely grown crops in the world. It is a polyploid species with a nuclear genome characterized by seven homoeologous chromosome groups (A, B and D genomes).

The availability of high density linkage maps is crucial to identify genotype-phenotype associations, but also for anchoring BAC contigs to genetic maps, a strategy followed for sequencing the wheat genome, promoted by the International Wheat Genome Sequencing Consortium (IWGSC, <http://www.wheatgenome.org/>) [1]. In

wheat, large genome size, high percentage of repetitive regions and low level of polymorphism [2] complicated the obtainment of high-resolution genetic maps by molecular markers. The availability of SSR and EST-SSR markers and the development of high-throughput systems, such as diversity array technology (DArT) [3], have overcome these difficulties.

SSR markers, together with expressed sequence tag (EST)-derived SSR markers, have become the markers of choice for cereal genetic mapping because of abundance in plant genomes, high information content, codominant inheritance, and easily detection and reproducibility. In particular EST-SSR markers have been extensively studied in barley [4] and maize [5] facilitating genomic information transfer from related species to wheat.

\* Correspondence: [agata.gadaleta@uniba.it](mailto:agata.gadaleta@uniba.it)

<sup>1</sup>Department of Soil, Plant and Food Sciences, University of Bari "Aldo Moro", Via Amendola 165/A, Bari 70126, Italy

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

The DArT system was originally validated for the simple genomes of rice [3] and *Arabidopsis thaliana* [6] and subsequently applied to other crops such as barley [7] and wheat [8]. Deriving from hybridization-based strategy, DArT markers generate whole-genome fingerprints by scoring the presence versus absence of DNA fragments in genomic DNA. Several molecular linkage maps have been developed in wheat, and many of them use DArT markers as anonymous markers [9-13]. The release of the sequence of 2000 wheat DArTs revealed their functional features in relation to their genetic position. The availability of the DArT marker sequences and their abundance in grass genomes changed the role of these markers, proposing them as an optimal tool for comparative genetic mapping or for identification of functional markers.

Comparative mapping is successfully applied in the assembly of sequences or chromosome regions. In this case molecular markers display an important role for their high degree of genetic collinearity across grass genomes because of their common ancestor [14]. Early syntenic analyses were based on RFLP markers, facilitating the evolutionary study of divergence between rice, *Brachypodium* and wheat [15]. The linear order of genes and their strong conservation in large chromosomal regions is described as macro-collinearity. Because of their small and compact genomes, rice and *Brachypodium* were among the first grass genomes to be completely sequenced and were proposed to study the larger and more complex wheat genome as "bridge genomes" [16]. Macro-collinearity relation across rice, *Brachypodium* and wheat genomes was demonstrated at the *Lr34* locus in wheat [17]. However, different studies also have shown that the sequence comparison at the level of individual genes breaks down the macro-collinearity and this was reported as micro-collinearity. At this level it is observed a mosaic gene conservation, with some genes present in both grass species, and some in only one species [16,18,19]. These frequent exceptions in collinearity could depend by translocation, duplication and deletion events [20].

In the present study, we report a complete genetic linkage map developed by crossing two durum wheat varieties characterized by different carotenoid pigment content. The objectives within this topic were: (1) to describe general results of the development of a durum wheat map by gSSR and DArT markers; (2) to extend the sequence analysis of DArT markers in terms of redundancy, comparative and functional analysis; (3) to characterize genetic regions involved in the grain carotenoid content by use of sequenced DArT markers.

## Results

### Genetic linkage map description

To show the distribution of the molecular markers, in particular of DArT markers in relation of their available

sequences in durum wheat genome, a genetic linkage map is reported in Figure 1.

The total number of polymorphic markers used to generate the durum linkage map was 546 (141 gSSR, 40 EST-SSR and 365 DArT markers) (Table 1). Moreover five loci corresponding to five genes (two for Glutamine synthetase 2, *Psy-1*, *Psy-2* and *Psy-3*) were included in the genetic map since they resulted polymorphic between the two parents [21-23]. After the first screening between the two parental lines, the 546 markers were scored in  $F_2:F_3$  families.

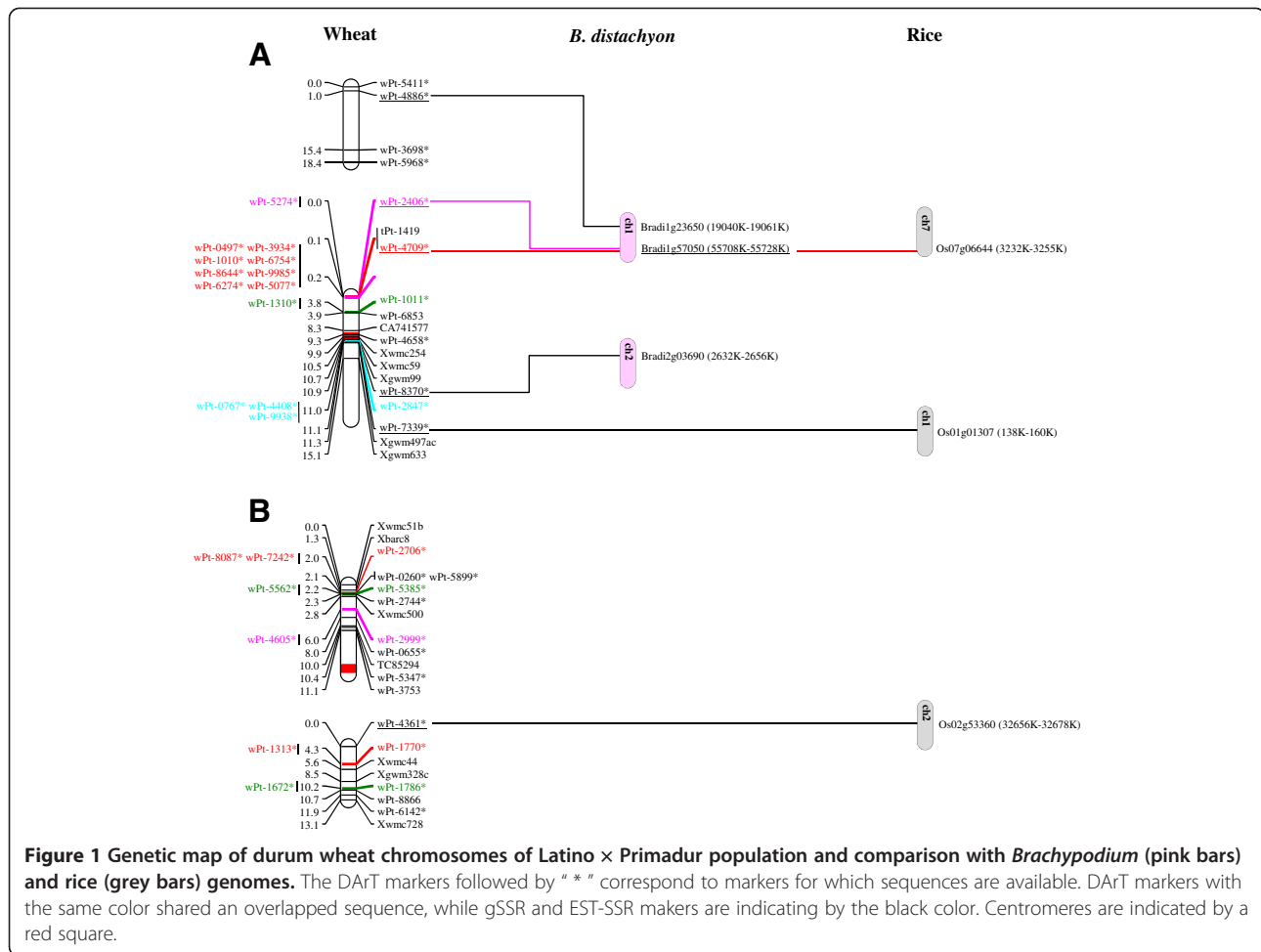
The complete linkage map consisted of a total of 505 molecular markers (122 gSSR, 32 EST-SSR and 351 DArT markers) distributed in 26 linkage groups (Table 2). The total length of the map accounted 1,272 cM, with an average chromosome length of 90 cM. The total marker number was highest in homoeologous group 7 (total 97 loci), as well as the total map length (259 cM). Homoeologous group 1 had the lowest marker number (total 58 loci) and shortest map length (58 cM).

Forty-one markers (19 SSR, 8 EST-SSR and 14 DArT markers) remained unlinked, probably due to the insufficient coverage of the genome. In fact in some cases we were not able to have a unique linkage group representing one chromosome and no markers were mapped on several bins.

Linkage groups were assigned to chromosomes by comparing markers on the generated map to previously published durum maps. For some chromosome regions (1A, 2A and 6A) the absence of marker anchor loci was replaced by the DArT markers location provided by Triticarte Pty. Ltd.

The 26 linkage groups were assigned to A genome (14) and B genome (12) (Figure 1). The chromosomes 2B, 4A, 5A and 6B were assembled in a single linkage group. The chromosomes 2A and 6A were constructed on 3 linkage groups, while 1A, 1B, 3A, 3B, 4B, 5B, 7A and 7B were represented by two linkage groups. Differences were found between the A and B genomes. B genome had 283 (56%) mapped markers in 688 cM and A genome had 222 (44%) mapped loci in 584 cM (Table 2). The whole 505 mapped loci covered the durum wheat map with a number per chromosome ranging from 16 (3A and 4B) to 61 (3B) (Table 2).

The primer pairs that amplified two or more loci mapped to homoeologous as well as to non-homoeologous sites. The highest number of loci was produced by *Xwmc51*, with two loci mapping on homoeologous group 2 and one locus on chromosome 1B, and by *Xwmc500* with three loci mapping on non-homoeologous chromosomes 4A, 1B and 7B. In five cases SSR markers mapped on the two homoeologous chromosomes: *Xwmc51* and *Xwmc382* primer pairs mapped on homoeologous group 2, *Xgwm132* marker mapped on homoeologous group 6



and *Xcfd6* and *Xcfa2040* loci mapped on homoeologous group 7.

Co-linearity between multi-allelic markers located on homoeologous chromosomes was observed for the group 2 and 7, indeed the two markers of each linkage group maintained the same genetic order. Instead for the *Xgwm132* loci on homoeologous group 6 there was not an evident co-linearity because of the different chromosome location across the arm (*Xgwm132ef-6AL4-0.55-1.00* and *Xgwm132b-6BS5-0.76-1.05*).

**Table 1 Number of markers screened among the parental lines Latino and Primadur and analyzed in the LP population**

Marker class	Marker analysed (n.)	Marker polymorphic (n.)	Marker analysed in the population (n.)	Polymorphism (%)
gSSR	400	141	141	35.3
EST-SSR	450	175	40	38.9
DArT	380	365	365	96.1
Total	1230	681	546	55.4

The analysis of the marker distributions revealed that there were 56 clusters of DArT markers across the durum wheat genome. Most of them were made of 2 or 3 DArT markers, but others consisted of clusters ranging from 4 to 11 markers at the same locus. In particular, 9 DArT markers clustered in a small region of 0.01 cM on chromosome 6B and 12 DArT markers were grouped in 0.13 cM on chromosome 3B. Other clusters were identified on chromosome 1A (9 DArT markers in 0.04 cm) and 7B (7 DArT markers in 0.15 cM).

The majority of loci in the LP map showed the expected 1:2:1 ratio for co-dominant markers (110 mapped/unmapped gSSR markers) or 3:1 for dominant markers ratio (31 mapped/unmapped gSSR and 365 DArT markers). However, 1 gSSR and 10 DArT markers, out of the 546 markers, presented a significant segregant distortion. Exactly the *Xwmc489*, *wPt-1830*, *wPt-8153*, *wPt-6017*, *wPt-7745-6B*, *wPt-3376-6B* and *wPt-5006* presented distortion in favor of the Primadur allele whereas the other DArT markers (*tPt-7209*, *wPt-8776-1B/2B*, *wPt-0694-2B* and *wPt-9261*) showed distorted segregation in favor of the Latino allele. The loci characterized

**Table 2 Chromosome assignment, marker distribution, length of linkage groups and marker density in the genetic map constructed with the F<sub>2</sub>:F<sub>3</sub> population Latino Primadur**

Linkage group	SSR	DArT	Total markers	Length (cM)	Density (cM/marker)
1A	6	25	31	33.5	1.1
1B	7	20	27	24.2	0.9
2A	14	29	43	115.3	2.7
2B	12	23	35	142.1	4.1
3A	2	14	16	43.4	2.7
3B	12	49	61	151.8	2.5
4A	7	27	34	75.0	2.2
4B	9	7	16	61.3	3.8
5A	20	6	26	115.6	4.4
5B	21	21	42	132.0	3.1
6A	6	20	26	65.4	2.5
6B	5	46	51	54.1	1.1
7A	16	30	46	136.1	3.0
7B	17	34	51	123.0	2.4
group 1	13	45	58	57.7	1.0
group 2	26	52	78	257.4	3.3
group 3	14	63	77	195.2	2.5
group 4	16	34	50	136.3	2.7
group 5	41	27	68	247.6	3.6
group 6	11	66	77	119.5	1.6
group 7	33	64	97	259.1	2.7
A genome	71	151	222	584.3	2.6
B genome	83	200	283	688.5	2.4
<b>Total</b>	154	351	505	1,272.8	2.5
<b>Average</b>	11.0	25.1	36.1	90.9	2.6

by segregation distortion were localized on chromosomes 1B, 2B and 6B. These loci were not used in the mapping population.

#### Analysis of DArT markers sequences

After scoring 365 DArT markers in the LP population, 351 DArT markers were randomly mapped along chromosomes. Only for 327 mapped DArT markers, sequences were available and they were analysed to search redundant sequences. Of these DArT markers, 201 sequences were not available in the public dataset made of 2000 sequences, already analysed by Marone et al. [10].

Based on high-quality DNA sequence, a total of 174 DArT markers were grouped into two contigs (characterized by overlapping and extended regions) and 58 marker clusters (identical sequences and longer than 400 bp). The remaining 153 DArT clones represented non-redundant sequences since they did not show any overlapping or identical regions with each other. In particular, assembling procedure detected 2 contigs on

linkage groups 3B and 6B and 58 marker grouped in clusters. The contig on 3B group has the highest DArT marker number (12 DArTs for a final length of 866 bp), whereas the *wPt-4386-6B* marker grouped into its sequence 9 DArT markers (final length of 637 bp). The majority of sequence DArT clusters shared high identity sequences and were characterized by < 3% mismatches.

No DArT clones sharing high identity sequences appeared to be positioned at different position on the genetic map. This underlines the reproducibility of DArT scores and genetic analysis since markers were in the same allelic phase.

Since grass genomes display a high degree of genetic collinearity because of their common ancestor [14], we used sequences from DArT markers to deduce syntenic relationships between wheat and two model genome species as *Brachypodium* and rice.

The non-redundant set of 153 sequences of mapped DArT markers was searched against the complete non-redundant protein, DNA and EST database (<http://geno>

mevolution.org/CoGe/CoGeBlast.pl) of *Brachypodium* and rice genomes to identify segments of collinearity. A compilation of the best BLAST hits from each searched sequence is shown in Additional file 1: Table S1.

The analysis revealed a mosaic gene conservation, with 46 genes (59%) present in both grass species, and 26 and 8 genes only in *Brachypodium* and rice species, respectively.

The identity percentages ( $E$  value less than  $e^{-10}$ ) for the 72 DArT sequences ranged from 75% (for the multi-loci *wPt-2587-6B* marker), to 100% (for the multi-loci *wPt-5432-3B* marker). Overall, 31 DArT sequences had significant matches to annotated *Brachypodium* proteins (Additional file 1: Table S1).

For the 54 DArT sequences aligned with rice (*Oryza sativa*) genome database, we observed an identity percentages ranging from 74% (*wPt-9277-2A*) to 100% (*wPt-9989-3B*) at  $E$  value less than  $e^{-10}$ . Out of the 54 matched sequences, 22 DArT sequences were recognized as annotated rice proteins.

In both analyses, the remaining DArT sequences were classified as uncharacterized proteins or unknown sequences. The comparisons also revealed that among all sequences (annotated, uncharacterized and not matched proteins), 14 *Brachypodium* proteins were also present in the rice EST database, sharing the same functions.

To further identify and validate collinear regions among the grass species, we considered also the EST-SSR or gene markers used for the development of the LP linkage groups as the sequences are available in NCBI. In particular out of the 32 mapped EST markers, 18 ESTs showed high sequence identity with *Brachypodium* genes, whereas 19 markers were aligned with rice genome. Among these 19 markers, 14 EST markers were detected in both genomes with similar annotated functions, identity scores higher than 76% and aligned regions bigger than 150 bp. Only for *TC77302-3A*, *CA694714-4B*, *TC8448-6A* and *BJ236800-6B* markers, we observed a small region of alignment (less than 100 bp) and an unknown function protein in *Brachypodium* and rice genomes. These markers mapped in different loci at  $E$  value less than  $e^{-10}$  with high identity percentages.

Overall syntenic regions were identified on durum wheat chromosomes 3B, 4A, 5B, 6B and 7A, where more than three sequenced DArT were present. Probably due to the gap showed on wheat chromosome 3B, 21 markers (3 SSR and 18 DArT markers) in 22.9 cM matched with two different regions in *Brachypodium* (chromosome 2 and 1) and one rice region (chromosome 1).

Two syntenic regions were also found for wheat chromosome 4A. A region of 54.1 cM (with 3 SSR, 1 EST-SSR and 10 DArT markers) and an overlapped one of 19.2 cM (with 1 SSR, 1 EST-SSR and 19 DArT markers) showed the best hit in the *Brachypodium* sequences of chromosome

5 and 1, while in rice genome it was not possible to identify a well-defined collinearity segment, since it was found on one DArT marker per each chromosome.

An interesting correspondence was found between *Brachypodium* chromosome 1 and a region of 22 cM of durum wheat on chromosome 5B, including a gap on its long arm. A total of 7 markers (2 EST-SSR and 5 DArT markers) had a better alignment with *Brachypodium* genome than rice, in which only two markers were mapped on the same chromosome 3.

A collinear segment was detected on chromosome 3 of *Brachypodium* with 37 cM of wheat segment on chromosome 6B (with 1 SSR, 2 EST-SSR and 42 DArT markers). Most of these DArT sequences had an uncharacterized protein role but showed a better match with *Brachypodium* genome than that of rice.

Remarkable results were detected on wheat chromosome 7A: 33 markers (1 SSR, 9 EST-SSR and 23 DArT markers) were considered in a region of 56.7 cM. Two collinear regions were found in both grass model genomes. The best hits mapped on *Brachypodium* chromosome 1 and rice chromosome 6.

In general, the characterized wheat genomic regions showed more correspondence with the *Brachypodium* genome than that of rice. The high frequency of gene similarity of DArT sequences among both model species provided a good indication that many DArT markers were within rich gene regions.

#### Functional analysis of DArT sequences

The availability of DArT sequences can provide opportunities to identify gene loci involved in the carotenoid biosynthesis considering the QTL regions detected (data in [24]). The analysis of DArT markers mapped in all the detected QTL regions allowed to find candidate genes for this trait only on chromosomes 3B (*Xbarc84-Xgwm299* and *Xbarc84-Xgwm340* intervals).

Since DArT sequences were short in many cases, matches were searched in Cerealsdb database (<http://www.cerealsdb.uk.net/>) to expand genomic region. The derived longer genomic clones were used as a query for a new BLASTX search against *Brachypodium* and rice protein databases.

The bioinformatic analysis showed that in the QTL region on chromosome 3B, the *wPt-9786*, *wPt-8959*, *wPt-3342* and *wPt-6929* markers matched with genes involved in abscisic acid response, while the *wPt-9989* marker presented homology with gene related to geranylgeranyl diphosphate accumulation. In details, the *wPt-9786* marker highlighted the best hit with both model species for the zinc finger CCCH domain-containing protein, involved in ABA-regulated dormancy [25]. Even the *wPt-8959* marker showed high sequence similarity with *Brachypodium* and rice genome and corresponded

to the abscisic stress-ripening protein (*Asr*) gene. This one seems to be involved in some phases of plant development (such as senescence and fruit development) and in responses to abiotic stresses (such as salt and limited light stresses). There are not precise information concerning the biological functions of *Asr* proteins, but it has been demonstrated a web of interactions around abscisic acid [26].

As confirmed in *Brachypodium* and rice genome, the *wPt-3342* marker belonged to a large transcription factor superfamily related to ABA and auxin responses: the B3 domain-containing proteins [27]. The *wPt-6929* marker had a best hit with the *Brachypodium* genome for G-type lectin S-receptor-like serine/threonine protein kinase, a positive regulator of plant tolerance to salt stress induced by ABA [28].

Instead, the *wPt-9989* marker showed the best hit with the *Brachypodium* and rice protein geranylgeranyl diphosphate reductase. This protein catalyzes the bidirectional reduction of geranylgeranyl diphosphate to phytol diphosphate, providing phytol for both tocopherol and chlorophyll synthesis [29].

To investigate in details the roles of these selected DArT markers, we performed a comparative matching analysis of DArTs to wheat 61 k microarray platform using the Plant Expression Database (PLEXdb) Blast [30]. PLEXdb identifies the corresponding probe sets on wheat GeneChip with E value < -10 through similarity searches. A total of 6 probe sets were identified as differentially transcribed from the *wPt-0668*, *wPt-0343*, *wPt-8959*, *wPt-3342*, *wPt-6929* to *wPt-9989* marker.

Although for the *wPt-0668* and *wPt-0343* markers we observed proteins of unknown function, the analysis in PLEXdb revealed that together with *wPt-9989* marker, they showed the highest level of expression in drought stress conditions, starting from root seeds germination phase. The *wPt-8959* marker seemed to be expressed higher in root tissue, while the *wPt-6929* and *wPt-3342* markers showed an invariable expression level in all microarray experiments.

## Discussion

DArT markers are among the most widely used tools in the generation of dense genetic maps. In the present study, we report a 1,272 cM intervarietal genetic map of durum wheat using the LP F<sub>2</sub>:F<sub>3</sub> segregating population in which SSR and DArT markers were integrated. Latino is a cultivar with high grain yield and adaptability but poor pasta quality parameters, while Primadur is a French variety with a high pigment color. Thus, we analysed the distribution of SSR and DArT markers into LP linkage map and used the DArT sequences to investigate them in collinear studies in *Brachypodium* and rice species.

The primary goal was to provide a molecular and functional map with 122 gSSR, 32 EST-SSR and 351 DArT markers. The LP map confirms other published intervarietal maps according to the positions and orders of commonly mapped markers [9,10,31] and was also used to develop a dense consensus map together with other 5 durum wheat bi-parental maps (Marone et al, 2012b). Efforts were made to minimize the gaps on chromosomes 1B, 2A, 3A, 3B, 4B, 5B, 7A and 7B, but in some cases we were unable to substantially close them for the lack of polymorphisms or for the low marker density, as described in several wheat maps [32,33]. The LP map also shows inter-marker distances higher than 20 cM and chromosome regions 4AS, 4BL and 7AS not covered at all by markers. Lack of genome coverage of homoeologous group 4 was observed in other published wheat maps [9,31,32,34].

The construction of the genetic linkage map indicated a high tendency of DArT markers to have a cluster distribution, determined through marker co-location. We presume that the genomic representations obtained with *PstI* reflect the presence of methylation sites, producing markers in the hypomethylated gene-rich regions [35]. The DArT clones appeared to be located mostly in the telomeric chromosome regions, probably because of the high G + C content of *PstI* sites and of the high recombination rate of these gene-rich regions [9]. The higher density of clusters in one specific region could depend by repeated clones, as confirmed by sequence analysis of clusters sharing highly similar sequences [9].

Based on DArT sequences, we selected contigs or representative markers (from DArT marker clusters) to get a distribution of these markers on *Brachypodium* and rice genomes and to provide information about their functional nature. Our analysis revealed that about one-third of the DArT markers had similar sequences in both model species. As extensively confirmed also by previous studies [17,33], *Brachypodium* resulted the closest wheat relative, allowing the identification of 72 genes in contrast of the 54 genes for rice genome. About 43% and 20% of predicted genes in *Brachypodium* and rice, respectively, were annotated as hypothetical genes. The sequences of *Brachypodium* could be helpful to verify these hypothetical genes in the rice genome.

In light of these results, we identified microsyntenic regions on chromosomes 3B, 4A, 5B, 6B and 7A with bigger size in *Brachypodium* than rice genome. Shatalinal et al. [36], Gu et al. [37], and La Rota and Sorrells [38] revealed the same conserved gene order for these regions starting from large insert clones and not considering microcollinearity through locations with a high number of molecular markers. Many DArT markers (about 80%) were absent in *Brachypodium* genes and about 14% of the rice genes failed in the collinear positions in *Brachypodium*.

It is difficult to determine if DArT sequences were lost in the model genomes, if there are alternative isoforms or if they consist of intronic regions, less conserved across grass species. The identity percentages ( $E$  value less than  $e^{-10}$ ) and the small region size of DArT markers are important characteristics that reduce the gene search into *Brachypodium* and rice species.

A detailed analysis was carried out expanding the DArT sequences on chromosome region 3BL with Chinese Spring genomic clones, when they were available. We focused on this region because it explains a QTL region for yellow pigment [24] with 2 SSR and 14 mapped DArT markers. A high proportion of the DArT clones correspond to sequences involved in the plant response mediated by abscisic acid (ABA) and geranylgeranyl diphosphate metabolism. The ABA corresponds to apocarotenoid molecules that coordinate plant response to water deficit and gene expression, while geranylgeranyl diphosphate is the precursor of all carotenoid components. It is not clear how the ABA-regulated genes (represented by *wPt-9786*, *wPt-8959*, *wPt-3342* and *wPt-6929* markers) and the genes involved in carotenoid biosynthesis (*wPt-6929* marker) could be related.

The analysis in wheat expression database reveals the highest expression level of *wPt-9989* marker in drought conditions, sign of the important function of the carotenoid molecules in stress response. Indeed carotenoids provide precursors for ABA (main apocarotenoid) that plants can accumulate under drought conditions, having effect on plant growth and development. The higher expression level in root tissue of the *wPt-8959* marker (abscisic stress-ripening protein gene) shows its involvement in ABA biosynthesis but not directly in stress conditions. This needs to be better clarified since in barley expression databases the best sequence hit for the *wPt-8959* marker is expressed mostly in roots and after pathogen infection.

## Conclusion

In conclusion, the current analysis provides relevant information on the use of DArT markers as a tool for collinearity studies in grass species and for the identification of genes controlling traits under investigation. Indeed the identification of genes involved in carotenoid biosynthesis makes these markers an important start point for future wheat breeding programs.

## Methods

### Plant material and genotyping

The genetic map obtained from Latino  $\times$  Primadur (LP) mapping population is reported here in all its parts and integrated with additional molecular markers. Only 5 chromosomes (2A, 3B, 5A, 7A and 7B) were previously published by Blanco et al. [24] for the subsequent QTL

analysis in carotenoid pigment content. This is achieved mainly with SSR and EST-SSR markers, and integrated with DArT markers provided by Triticarte Pty. Ltd (Canberra, Australia; <http://www.triticarte.com.au>) as whole-genome profiling service laboratory.

SSR sequences and PCR conditions were provided and termed from GWM (*Gatersleben Wheat Microsatellite*, [39]), WMC (*Wheat Microsatellite consortium*; [40]; Prasad et al, 2002; [41]), BARC (*Bangladesh Agricultural Research Council*; [42]) and CFA (molecular probes from RFLP). The development of EST-SSR primer pairs was reported by La Rota et al, [43] and available in the public database <http://wheat.pw.usda.gov>. Primer pairs were chosen in order to represent each microsatellite class and to produce PCR products ranging from 100 to 350 bp in length.

The primer pairs for candidate GS genes mapped on chromosome 2A and 4A were reported by Gadaleta et al. [22]; the primers for candidate Psy1 gene located on chromosome 7A were reported by He et al. [44] and those for candidate Psy3 gene located on chromosome 5A derived by Dibari et al. [21]. The primer pair for candidate Psy2 gene (*forward* TTGGAAATCGAGGT ATATGACCT, *reverse* ACTGGACGAACTGGCACAG) located on chromosome 5B derived from a contemporary study not yet published. All the 365 provided DArT markers were indicated with an ID number starting with "wPt" code, followed by a clone number.

### Segregation analysis and map construction

For all of the loci the segregation ratio of 1:2:1 or 3:1 was tested using Chi-squared analysis. The linkage analysis was performed using the Kosambi mapping function within the JoinMap v. 4.0 software [45]. Linkage groups were assembled using a LOD score  $\geq 3$ , after preliminary analysis using LOD score ranging from 2 to 10.

Somers et al. [46], Sourdille et al. [47,48], and Gadaleta et al. [9] previously mapped gSSR markers used as anchor loci and for assigning linkage groups to a particular chromosome. When necessary markers were removed and the order recalculated until a stable and consistent order with the physical position of the markers on chromosome bins was reached.

### Sequence analysis of the DArT clones

The DArT sequences mapped in LP map were kindly obtained by Triticarte Pty. Ltd. The sequences analysis of 327 DArT markers was carried out using the CLC Genomics Workbench software v. 5.5.1 (CLC bio, Muehlthal, Germany; limited free version) with the following parameters: nucleotide mismatch cost = 2; insertion = deletion costs = 2; length fraction = 0.4; similarity = 0.9; conflict among individual bases resolved by voting for the base with the highest frequency. CLC Genomics Workbench is

a program used for assembling similar DARt fragments into longer consensus sequences (known as “contig”). BLASTn homology searches and functional annotation were carried out using COGE terms (<http://genomevolution.org/CoGe/CoGeBlast.pl>) [49] to compare the non-redundant DARt sequences to rice (*Oryza sativa*) and *Brachypodium* genomes. An expectation value (E) of  $e^{-10}$  was used as the significant threshold. The DARt markers were assigned to individual grass chromosomes based on their best match to the model genomes. The DARt sequences were also analyzed by BLASTX against NCBI non-redundant protein database (<http://www.ncbi.nlm.nih.gov/BLAST>). The sequence from DARt markers were also screened in BLASTN search against the Triticeae Repeat Sequence (TREP) database (<http://wheat.pw.usda.gov/ITMI/Repeats>).

For each DARt analysis EST-SSR or gene sequences were used as control for confirmation of collinear chromosome region.

## Additional file

**Additional file 1: Table S1.** Genetic and functional features of wheat DARt and EST-SSR markers. All sequences detected in *Brachypodium* and rice species were selected basing on identity percentages (>70%) and E value (less than  $e^{-10}$ ).

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

PC performed linkage map development and bioinformatic analysis. AG, AMM and AB contributed to data interpretation and assisted in drafting the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

This research was supported by a grant from the Italian Ministry of Agriculture, project “Mappa 5A”, by a grant from MIUR project PON01\_01145-ISCOCSEM, project PAQ-Industria 2015 and projects PRIN-2010/2011.

## Author details

<sup>1</sup>Department of Soil, Plant and Food Sciences, University of Bari “Aldo Moro”, Via Amendola 165/A, Bari 70126, Italy. <sup>2</sup>CRA, Cereal Research Centre, SS16 km 675, Foggia 71122, Italy.

Received: 1 August 2013 Accepted: 25 November 2013

Published: 5 December 2013

## References

- Moolhuijzen P, Dunn DS, Bellgard M, Carter M, Jia J, Kong X, Gill BS, Feuillet C, Breen J, Appels R: **Wheat genome structure and function: genome sequence data and the international wheat genome sequencing consortium.** *Aust J Agric Res* 2007, **58**(6):470–475.
- Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD: **RFLP-based genetic maps of wheat homoeologous group 7 chromosomes.** *Theor Appl Genet* 1989, **78**:495–504.
- Jaccoud D, Peng K, Feinstein D, Kilian A: **Diversity arrays: a solid state technology for sequence information independent genotyping.** *Nucleic Acids Res* 2001, **29**:25.
- Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, Kota R, Varshney RK, Perovic D, Grosse I, Graner A: **A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics.** *Theor Appl Genet* 2007, **114**:823–839.
- Davis GL, McMullen MD, Baysdorfer C, Musket T, Grant D, Staebell M, Xu G, Polacco M, Koster L, Melia-Hancock S, Houchins K, Chao S, Coe EH Jr: **A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map.** *Genetics* 1999, **152**:1137–1172.
- Wittenberg AHJ, Van der Lee T, Cayla C, Kilian A, Visser RGF, Schouten HJ: **Validation of the high throughput marker technology DARt using the model plant *Arabidopsis thaliana*.** *Mol Genet Genom* 2005, **274**:30–39.
- Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Klein-hofs A, Kilian A: **Diversity arrays technology (DARt) for whole-genome profiling of barley.** *Proc Natl Acad Sci U S A* 2004, **101**:9915–9920.
- Akbari M, Wenzl P, Caig V, Carling J, Xia L, Yang S, Uszynski 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**:1409–1420.
- Gadaleta A, Giancaspro A, Giove SL, Zacheo S, Mangini G, Simeone R, Signorile A, Blanco A: **Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat.** *Theor Appl Genet* 2009, **118**:1015–1025.
- Marone D, Panio G, Ficco DBM, Russo MA, De Vita P, Papa R, Rubiales D, Cattivelli L, Mastrangelo AM: **Characterization of wheat DARt markers: genetic and functional features.** *Mol Genet Genom* 2012, **287**:741–753.
- Peleg Z, Saranga Y, Suprunova T, Ronin Y, Röder MS, Kilian A, Korol AB, Fahima T: **High-density genetic map of durum wheat and wild emmer wheat based on SSR and DARt markers.** *Theor Appl Genet* 2008, **117**:103–115.
- Zhang W, Chao S, Manthey F, Chicaiza O, Brevis JC, Echenique V, Dubcovsky J: **QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat.** *Theor Appl Genet* 2008, **117**(8):1361–1377.
- Mantovani P, Maccaferri M, Sanguineti MC, Tuberosa R, Catione I, Wenzl P, Thomson B, Carling J, Huttner E, De Ambrogio E, Kilian A: **An integrated DARt-SSR linkage map of durum wheat.** *Mol Breed* 2008, **22**:629–648.
- Gale MD, Devos KM: **Plant comparative genetics after 10 years.** *Science* 1998, **282**:656–659.
- Keller B, Feuillet C: **Colinearity and gene density in grass genomes.** *Trends Plant Sci* 2000, **5**:246–251.
- Huo N, Vogel JP, Lazo GR, You FM, Ma Y, McMahon S, Dvorak J, Anderson OD, Luo MC, Gu YQ: **Structural characterization of *Brachypodium* genome and its syntenic relationship with rice and wheat.** *Plant Mol Biol* 2009, **70**:47–61.
- Bossolini E, Wicker T, Knobel PA, Keller B: **Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation.** *Plant J* 2007, **49**:704–717.
- Bossolini E, Krattinger SG, Keller B: **Development of SSR markers specific for the Lr34 resistance region of wheat using sequence information from rice and *Aegilops tauschii*.** *Theor Appl Genet* 2006, **113**:1049–1062.
- Swigonova Z, Bennetzen JL, Messing J: **Structure and evolution of the r/b chromosomal regions in rice, maize and Sorghum.** *Genetics* 2005, **169**:891–906.
- Lysák MA, Schubert I: **Mechanisms of chromosome rearrangements.** In *Plant genome diversity, Physical structure, behaviour and evolution of plant genomes, Volume 2*. Edited by Leitch IJ, Greilhuber J, Doležel J, Wendel JF. Wien: Springer; 2013:137–147.
- Dibarí B, Murat F, Chosson A, Gautier V, Poncet C, Lecomte P, Mercier I, Bergès H, Pont C, Blanco A, Salse J: **Deciphering the genomic structure, function and evolution of carotenogenesis related phytoene synthases in grasses.** *BMC Genomics* 2012, **13**:221.
- Gadaleta A, Nigro D, Giancaspro A, Blanco A: **The glutamine synthetase (GS2) genes in relation to grain protein content of durum wheat.** *Funct Integr Genom* 2011, **11**(4):665–670.
- Pozniak CJ, Knox RE, Clarke FR, Clarke JM: **Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat.** *Theor Appl Genet* 2007, **114**:525–537.
- Blanco A, Colasuonno P, Gadaleta A, Mangini G, Schiavulli A, Simeone R, Digesù AMD, Vita P, Mastrangelo AM, Cattivelli L: **Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat.** *J Cereal Sci* 2011, **54**:255–264.
- He Y, Gan S: **A novel zinc-finger protein with a proline-rich domain mediates ABA-regulated seed dormancy in *Arabidopsis*.** *Plant Molec Biol* 2004, **54**:1–9.



26. Hu W, Huang C, Deng X, Zhou S, Hou S, Chen L, Li Y, Wang C, Ma Y, Yuan Q, Wang Y, Cai R, Liang X, Yang G, He G: **TaASR1, a transcription factor gene in wheat, confers drought stress tolerance in transgenic tobacco.** *Plant Cell Environ* 2013, **36**(8):1449–1464.
27. Wang Y, Deng D, Zhang R, Wang S, Bian Y, Yin Z: **Systematic analysis of plant-specific B3 domain-containing proteins based on the genome resources of 11 sequenced species.** *Mol Biol Rep* 2012, **39**(5):6267–6282.
28. Sun XL, Yu QY, Tang LL, Ji W, Bai X, Cai H, Liu XF, Ding XD, Zhu YM: **GsSRK, a G-type lectin S-receptor-like serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress.** *J Plant Physiol* 2012, **170**(5):505–515.
29. Della Penna D, Pogson BJ: **Vitamin synthesis in plants: tocopherols and carotenoids.** *Ann Rev* 2006. doi:10.1146/annurev.arplant.56.032604.144301.
30. Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA: **PLEXdb: gene expression resources for plants and plant pathogens D1194–D1201.** *Nucl Ac Res* 2012, **40**:1194–1201.
31. Marone D, Laidò G, Gadaleta A, Colasuonno P, Ficco DBM, Giancaspro A, Giove S, Panio G, Russo MA, De Vita P, Cattivelli L, Papa R, Blanco A, Mastrangelo AM: **A high-density consensus map of A and B wheat genomes.** *Theor Appl Genet* 2012. doi:10.1007/s00122-012-1939-y.
32. Torada A, Koike M, Mochida K, Ogihara Y: **SSR-based linkage map with new markers using an intraspecific population of common wheat.** *Theor Appl Genet* 2006, **112**:1042–1051.
33. Quraishi UM, Abrouk M, Murat F, Pont C, Foucrier S, Desmazieres G, Charmet G, Paux E, Murigneux A, Guerreiro L, Lafarge S, Le Gouis J, Feuillet C, Salse J: **Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution.** *Plant J* 2011, **65**:745–756.
34. Elouafi I, Nachit MM: **A genetic linkage map of the Durum x Triticum dicoccoides backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits.** *Theor Appl Genet* 2004, **108**:401–413.
35. Van Os H, Stam P, Visser RGF, Van Eck HJ: **RECORD: a novel method for ordering loci on a genetic linkage map.** *Theor Appl Genet* 2006, **112**:389.
36. Shatalinal M, Wicker T, Buchmann JP, Oberhaensli S, Šimková H, Doležal J, Keller B: **Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno.** *Plant Biotech J* 2013, **11**:23–32.
37. Gu YQ, Ma Y, Huo N, Vogel JP, You FM, Lazo GR, Nelson WM, Soderlund C, Dvorak J, Anderson OD, Luo MC: **A BAC-based physical map of *Brachypodium distachyon* and its comparative analysis with rice and wheat.** *BMC Genomics* 2009, **10**:496.
38. La Rota M, Sorrells ME: **Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat.** *Funct Integr Genom* 2004, **4**:34–46.
39. Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier NH, Laroy P, Ganal MV: **A microsatellite map of wheat.** *Genetics* 1998, **149**:2007–2023.
40. Varshney RK, Prasad M, Roy JK, Kumar N, Harjit-Singh Dhaliwal HS, Balyan HS, Gupta PK: **Identification of eight chromosomes and a microsatellite marker on 1AS associated with QTL for grain weight in bread wheat.** *Theor Appl Genet* 2000, **8**:1290–1294.
41. Gupta P, Balyan H, Edwards K, Isaac P, Korzun V, Röder M, Gautier MF, Joudrier P, Schlatter A, Dubcovsky J, De la Pena R, Khairallah M, Penner G, Hayden M, Sharp P, Keller B, Wang R, Hardouin J, Jack P, Leroy P: **Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat.** *Theor Appl Genet* 2002, **105**:413–422.
42. Song QJ, Fickus EW, Cregan PB: **Characterization trinucleotide SSR motifs in wheat.** *Theor Appl Genet* 2002, **104**:286–293.
43. La Rota M, Kantety RV, Yu JK, Sorrells ME: **Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley.** *BMC Genomics* 2005, **6**:23–30.
44. He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX, Xia XC: **Characterization of phytoene synthase 1 gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker.** *Theor Appl Genet* 2008, **116**(2):213–221.
45. Van Ooijen JW, Voorrips RE: *JoinMap 3, software for the calculation of genetic linkage maps.* Plant Research International: Wageningen; 2001.
46. Somers DJ, Isaac P, Edwards K: **A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.).** *Theor Appl Genet* 2004, **109**:1105–1114.
47. Sourdille P, Singh S, Cadalen T, Brown-Guedira GL, Gay G, Qi Bikram S, Gill L, Dufour P, Murigneux A, Bernard M: **Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.).** *Funct Integr Genom* 2004, **4**(1):15–25.
48. Sourdille P, Cadalen T, Guyomarch H, Snape JW, Perretant MR, Charmet G, Boeuf C, Bernard S, Bernard M: **An update of the courtot x Chinese spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat.** *Theor Appl Genet* 2003, **106**:530–538.
49. Lyons E, Pedersen B, Kane J, Alam M, Ming R, Tang H, Wang X, Bowers J, Paterson A, Lisch D, Freeling M: **Finding and comparing syntenic regions among Arabidopsis and the out-groups papaya, poplar, and grape: CoGe with rosids.** *Plant Physiol* 2008, **148**:1772–1781.

doi:10.1186/1471-2156-14-114

**Cite this article as:** Colasuonno et al.: Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and *Brachypodium* genomes. *BMC Genetics* 2013 **14**:114.

**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

