Exploitation of SNP markers located on wheat 5A chromosome for the study of syntenic relationship with model species

Giancaspro A., Nigro D., Giove S.L., Zacheo S.A., Simeone R., Piarulli L., Colasuonno P., Blanco A.

in

Porceddu E. (ed.), Damania A.B. (ed.), Qualset C.O. (ed.).
Proceedings of the International Symposium on Genetics and breeding of durum wheat

Bari : CIHEAM
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 110

2014
pages 229-233

To cite this article / Pour citer cet article

Giancaspro A., Nigro D., Giove S.L., Zacheo S.A., Simeone R., Piarulli L., Colasuonno P., Blanco A. **Exploitation of SNP markers located on wheat 5A chromosome for the study of syntenic relationship with model species.** In : Porceddu E. (ed.), Damania A.B. (ed.), Qualset C.O. (ed.). *Proceedings of the International Symposium on Genetics and breeding of durum wheat.* Bari : CIHEAM, 2014. p. 229-233 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 110)

Exploitation of SNP markers located on wheat 5A chromosome for the study of syntenic relationship with model species

Angelica Giancaspro, Domenica Nigro, Stefania Lucia Giove, Silvana Addolorata Zacheo, Rosanna Simeone, Luciana Piarulli, Pasqualina Colasuonno, Antonio Blanco

Department of Soil, Plant and Food Sciences, Section of Genetic and Plant Breeding,
University of Bari "Aldo Moro", Bari, Italy

Abstract. Wheat is one of the most important world's crop species, however genomic studies are made difficult by the high complexity and width of its genome, which is very large and polyploid. For these reasons, sequence information from sequenced genomes of model species together with the study of the syntenic relationships among them represent a very important tool to elucidate the structure and the function of the complex wheat genome. In the present work, we analyzed the syntenic relationship existing between wheat, rice and *Brachypodium distachyon*, focusing in particular on SNP markers previously mapped on wheat 5A chromosome. The comparative genomic study described in the present work confirmed the great importance of sequence information associated with SNP markers combined with comparative genomics studies. In particular, the work evidenced a co-linearity and a good conservation of loci order between wheat and the two model reference species; such studies will be very useful for understanding the evolution of grass genomes, map based cloning of important genes, and for future wheat genome sequencing and the acceleration of genomic based improvement of these important crops.

Keywords. Wheat – SNP – Synteny – *B. distachyon* – Rice.

Exploitation des marqueurs SNP situés sur le chromosome 5A du blé pour l'étude des relations synténiques entre des espèces modèles

Résumé. *Le blé est une des cultures les plus importantes au monde et cependant, les études génomiques s'avèrent être difficiles à cause de la grande complexité et de la taille de son génome et de sa polyploïdie. Pour ces raisons, les informations de séquence obtenues à partir des génomes séquencés des espèces modèles et l'étude de leurs relations synténiques représentent un outil très important pour expliquer la structure et la fonction du génome du blé si complexe. Dans le présent travail, nous avons analysé la relation synténique entre le blé, le riz et Brachypodium distachyon, en nous concentrant, en particulier, sur les marqueurs SNP précédemment cartographiés sur le chromosome 5A du blé. L'étude génomique comparative illustrée dans ce travail a confirmé la grande importance des informations de séquence associées à des marqueurs SNP, combinées avec des études de génomique comparative. En particulier, les travaux ont mis en évidence une co-linéarité et une bonne conservation de l'ordre des loci entre le blé et les deux espèces modèles de référence ; ces études seront très utiles pour la compréhension de l'évolution des génomes des graminées, pour le clonage de gènes importants à l'aide de la cartographie, et pour le séquençage futur du génome du blé et l'accélération de l'amélioration basée sur la génomique de ces cultures importantes.*

Mots-clés. *Blé – SNP – Synténie – B. distachyon – Riz.*

I – Introduction

Despite the economic importance of wheat, genomic studies on this crop are very difficult due to the high complexity and size of its genome (16,937 Mb), its polyploid nature ($2n = 4x = 28$ for durum wheat or $2n = 6x = 42$ for bread wheat), and the large extension (about 80%) of repetitive elements (Vitulo *et al*., 2011). Moreover, the extension of the coding portion of wheat genome, which is expected to account for about 90,000 genes, is only a very small percentage of the total, and it is quite similar to other grass species such as rice (*Oryza sativa,* 430 Mb) or *Brachypodium distachyon* (300 Mb) with have much smaller genomes (Arumuganathan and Earle, 1991). For all these reasons, sequence information from sequenced genomes of model species, and the study of the syntenic relationships among them represent a very important tool to elucidate the structure of the complex wheat genome, and to infer information about gene order and function on the different chromosomes.

The syntenic relationships existing among *Poaceae* have been well described since many years (Moore *et al*., 1995). In particular, rice and *B. distachyon* have been recently recognized as model species for both comparative and functional genomics in grass species, due to the limited size and to the availability of complete sequences of their genomes. Several authors showed a good co-linearity among large chromosome segments of these two genomes with that of wheat (Sorrels *et a*l., 2003). In particular, regions of synteny to wheat chromosome 5A have been found in many grass species (Sorrels *et al*., 2003); moreover, studies by Moore *et al.* (2005) and Sorrels *et al*. (2003) reported that wheat homoeologous group 5 shows one of the most complex syntenic relationships with rice.

Wheat 5A chromosome accounts for a total size of 827 Mb, representing about 4.9% of wheat entire genome (Safar *et a*l., 2010), with predicted chromosome arms lengths of 295 Mb and 532 Mb for 5AS and 5AL, respectively (Safar *et al*., 2010). The importance of this chromosome is due to the fact that it carries lots of genes controlling important traits, such as *domestication* (*Q*) (Simons *et al*., 2006), resistance to abiotic stresses, frost tolerance (*Fr1*, *Fr2*), vernalization requirement (*Vrn)*, regulation of homoeologous chromosome pairing (*Ph1*), loci (QTLs) for yield and productivity, and several genes for pathogens resistance (e.g. Fusarium head blight).

Several studies have described a co-linearity between wheat 5A short arm and rice chromosomes R12, R11, R5 and R1 (Sarma *et al*., 2000; Qi end Gill, 2001). In particular, Linkiewicz *et al.* (2004) confirmed that the co-linearity between wheat chromosome 5A and rice chromosome 12 spans from 5A short arm to the proximal region of the long arm, suggesting the presence of similar centromere location. Wu *et al.* (1998) hypothesized that similarity between wheat 5AS and rice R11 could have been originated from the duplication of a larger segment of chromosome R12 in R11. On the other hand, 5A long arm shows syntenic regions with the rice chromosomes 9 and 3 (Vitulo *et al*., 2011). The 5A short arm is co-linear to *Brachypodioum* chromosome 4, whereas the long arm shows a close synteny with chromosomes 4 and 1 (Vitulo *et al*., 2011).

In the present study the SNP markers genetically and physically mapped on wheat 5A chromosome by Gadaleta *et al*. (2014, in press) were exploited to study the syntenic and functional relationships among wheat and the model genomes of rice and *Brachypodium distachyon*..

II – Material and Methods

The 90K iSelect array developed by Illumina CSPro® and described by Wang *et al*. (2014) was used to survey a set of 81,587 SNP markers across a recombinant inbred line (RIL) population developed by Gadaleta *et al*. (2012) by crossing the wheat cv. Chinese Spring (CS) and the line Chinese Spring-5A dicoccoides, which is a disomic substitution line carrying the 5A chromosome of CS replaced by the 5A chromosome from *Triticum turgidum* ssp. *dicoccoides*. The same SNP markers were also assayed on a set of aneuploid lines derived from Chinese Spring, including nulli-tetrasomic (NT), di-telosomic (DT) and 12 deletion bin lines, of which four dividing the short arm into 5 bins and 8 dividing the long arm into 9 bins. SNP markers data were integrated into a previous genetic linkage map of wheat 5A chromosome developed by Gadaleta *et al*. (2012), in order to produce a high-density and more saturated map. A sub-set of 50 SNP markers genetically and physically mapped on wheat 5A chromosome were employed to carry out a comparative study among wheat and the sequenced genomes of two model species, represented by *Brachypodium* *distachyon* and *Oryza sativa*, with the objective to detect orthologous and syntenic relationships among the three genomes. The similarity search was carried out by launching each wheat SNP sequence against *B. distachyon* and *O. sativa* genomes by means of the BLAST tool at the site http://www.phytozome.it. The orthologous SNP sequences were also subjected to a BLASTX analysis against non-redundant protein database for assigning putative functions. For both analysis, an E value cut-off of 10^{-7} was used.

III – Results and Discussion

In the present work we analysed the orthology relationships among wheat, *Brachypodium distachyon* and *Oryza sativa* by focusing on wheat chromosome 5A, and in particular exploiting 50 of the SNP markers genetically and physically mapped on this chromosome by Gadaleta *et al*. (2014, in press). We choose this chromosome because different studies (Moore *et al*., 2005; Sorrels *et al*., 2003) reported that wheat homoeologous group 5 has one of the most complex syntenic relationships with rice and other grass species among all the chromosome groups. Out of 50 SNP sequences analysed, 45 of them identified a putative orthologous with an E value $> 10^{-5}$; in particular, 35 wheat SNP sequences showed similarity with both model species, 10 SNP sequences only with *B. distachyion* and five only with *O. sativa*. The majority of the orthologous wheat SNPs were found to be related to *Brachypodium* sequences located on chromosomes 4 and 1, while the remaining 2 showed similarity with sequences on chromosomes 2 and 3, respectively. Rice sequences with the highest level of homology to wheat SNPs were identified on chromosomes 3, 12, 9. In particular, 30 wheat.

SNPs matched with genomic sequences located on rice chromosomes 12 and/or 3, while seven wheat SNPs matched with rice chromosome 9, and three with rice chromosome 7. These results are in agreement with what reported by previous studies (Vitulo *et al*., 2011), as the wheat SNPs genetically and physically mapped on 5A chromosome short arm showed a close relationship to the rice chromosome 12, while the majority of SNP sequences located on 5A long arm showed a closer synteny with rice sequences on chromosomes 9 and 3.

By comparing the map position of the analyzed 50 wheat SNP markers with the genomic position of the syntenic sequences found on *Brachipodium distachyon* and rice chromosomes, we confirmed the good co-linearity existing among these three genomes. In fact, a good conservation of sequence order was observed among the wheat chromosome 5A, and the chromosomes 12, 9, 3 of rice and 1 and 4 of *Brachipodium*, as already reported by other several authors (Sidhu *et al*., 2008; Sorrels *et al*., 2009; Vitulo *et al*., 2011). Similar homologous relationships were also reported by Sorrels *et a*l. (2009), in fact a small number of SNPs found homologous sequences to the *Brachipodium* chromosomes 2 and 3 and to rice chromosomes 2 and 7. Comparative genomics studies and the elucidation of the syntenic relationships existing among species in the grass family, is of particular importance because the family comprises a number of economically important crops, and the genomic analysis using model species could offer a potentially useful strategy for the development of highly saturated genetic linkage maps and for gene discovery in cultivated wheat (Gadaleta *et al*., 2012).

In order to identify some of the functional relationships underlying the synteny among wheat, rice and *Brachypodium* genomes, we tried to attribute a putative function to all the 50 analysed SNPs (Fig.1).

Figure 1. Graphic representation of putative function of 50 SNP sequences mapped on chromosome 5A by Gadaleta *et al*. (2014, in press).

Interestingly, we found that most of the wheat SNP sequences fell into coding regions, while only 20% matched with sequences belonging to non-coding regions of the two reference genomes. A very high percentage of the orthologous sequences identifying coding regions (25%) matched with predicted uncharacterized proteins, i.e., proteins of unknown function belonging to rice, whereas the other identified genes were found to be involved in several metabolic pathways. In particular, 30% of wheat SNP sequences matching with genes identified calmodulin-like proteins, while 9% were found to match with polypeptides of known function represented by a protein kinase superfamily. The remaining 16% represented other classes of genes. In conclusion, we can say that this work confirmed the great importance of sequence information associated with SNP markers combined with comparative genomics studies. Because of the syntenic relationships existing among grass species, knowledge from a model species could greatly facilitate the study of other important cereal crops. In particular in wheat, which is characterized by a very wide and complex genome, not yet completely sequenced, SNPs represent a powerful tool which could facilitate comparison with model species with the aim to infer information about the evolutionary mechanisms, order of genes and the putative function of genomic sequences.

Acknowledgments

The research project was supported by grants from Italy’s Ministero dell’Istruzione, dell’Università e della Ricerca, projects “PON-ISCOCEM” and “PRIN-2010-11”.

References

Kumar S., Mohan A., Balyan H., Gupta P.K., 2009. Orthology between genomes of *Brachypodium*, wheat and rice. *BMC Research Notes,* 2, pp. 93.

Arumuganathan K., Earle E.D., 1991. Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.,* 9, pp. 208-218.

Gadaleta A., Giancaspro A., Giove S.L., Zacheo S., Incerti O., Simeone R., Colasuonno P., Nigro D., Valè G., Cattivelli L., Stanca M., Blanco A., 2012. Development of a deletion and genetic linkage map for the 5A and 5B chromosome of wheat (*Triticum aestivum*). *Genome,* 55, pp. 417-427.

Gadaleta A., Giancaspro A., Nigro D., Giove S.L., Zacheo S.A., Incerti O., Simeone R., Piarulli L., Colasuonno P., Blanco A., 2014. High-Density genetic and physical Map of Wheat Chromosome 5A for the detection of candidate genes in QTL analysis. *Molecular Breeding.* DOI10.1007/s11032-014-0185-1.

Qi L.L., Gill B.S., 2001. High-density physical maps reveal that the dominant male-sterile gene Ms3 is located in a genomic region of low recombination in wheat and is not amenable to map-based cloning. In: *Theor. Appl. Genet*., 103, pp. 998-1006.

Safar J., Simkova H., Kubalakova M., Cihalikova J., Suchankova P., Bartos J., Dolezel J., 2010. Development of Chromosome-Specific BAC Resources for Genomics of Bread Wheat. *Cytogenet, Genome Res.,* 129, pp. 211–223.

Sarma R.N., Fish L., Gill B.S., Snape J.W., 2000. Physical characterization of the homoeologous group 5 chromosomes of wheat in terms of rice linkage blocks, and physical mapping of some important genes. *Genome,* 43, pp. 191-198.

Somers D.J., Fedak G., Savard M., 2003. Molecular mapping of novel genes controlling fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. *Genome,* 46, pp. 555–564.

Sorrells M.E., La Rota M., Bermudez-Kandianis C.E., Greene R.A., Kantety R., Munkvold J.D., Miftahudin, Mahmoud A., Ma X., Gustafson P.J., Qi L.L., Echalier B., Gill B.S., Matthews D.E., Lazo G.R., Chao S., Anderson O.D., Edwards H., Linkiewicz A.M., Dubcovsky J., Akhunov E.D., Dvorak J., Zhang D.,. Nguyen H.T., Peng J., Lapitan N.L.V., Gonzalez-Hernandez J.L., Anderson J.A., Hossain K., Kalavacharla V., Kianian S.F., Choi D.W., Close T.J., Dilbirligi M., Gill K.S., Steber C., Walker-Simmons M.K., McGuire P.E., Qualset C.O., 2003. Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.,* 13, pp. 1818-1827.

Vitulo N., Albiero A., Forcato C., Campagna D., Dal Pero F., Bagnaresi P., Colaiacovo M., Faccioli P., Lamontanare A., Simkova H., Kubalakova M., Perrotta G., Facella P., Lopez L., Pietrella M., Gianese G., Dolezel J., Giuliano G., Cattivelli L., Valle G., Stanca A.M., (2011). First survey of the Wheat Chromosome 5° composition through a Next Generation Sequencing Approach. *Plos One,* 6 (10).

Wang S., Wong D., Forrest K., Allen A., Chao S., Huang E., Maccaferri M., Salvi S., Milner S.G., Cattivelli L., Mastrangelo A.M., Stephen S., Luo M.C., Dvorak J., Mather D., Appels R., Dulferos R., Brown-Guedira G., Akhunova A., Feuillet C., Salse J., Morgante M., Pozniak C., Wieseke R., Plieske J., Morell M., Dubcovsky J., Ganal M., Tuberosa R., Lawley C., Mikoulitch I., Cavanagh C., Edwards K.J., Hayden M., Akhunov E., 2014. Characterization of polyploid wheat genomic diversity using the high-density 90,000 SNP array. *Plant Biotechnology Journal, In press.*

Wu J. Z., Kurata N., Tanoue H., Shimokawa T., Umehara Y., Yano M., Sasaki T., 1998. Physical mapping of duplicated genomic regions of two chromosome ends in rice. *Genetics,* 150, pp. 1595-1603.