Journal of Cereal Science 75 (2017) 25–34

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Genome-wide association mapping of phenolic acids in tetraploid wheats

Domenica Nigro [a], Barbara Laddomada [b], Giovanni Mita [b], Emanuela Blanco [c], Pasqualina Colasuonno [d], Rosanna Simeone [a], Agata Gadaleta [d], Antonella Pasqualone [e], Antonio Blanco [a, *]

[a] Department of Soil, Plant & Food Sciences, Plant Breeding Section, University of Bari, Via Amendola 165/A, 70126, Bari, Italy
[b] Institute of Sciences of Food Production, ISPA, CNR, Via Prov.le Monteroni, 73100, Lecce, Italy
[c] Institute of Biosciences and Bioresources, IBBR, CNR, Via Amendola 165/A, 70126, Bari, Italy
[d] Department of Agricultural & Environmental Science, Research Unit of "Genetics and Plant Biotechnology", University of Bari, Via Amendola 165/A, 70126, Bari, Italy
[e] Department of Soil, Plant & Food Sciences, Food Science Section, University of Bari, Via Amendola 165/A, 70126, Bari, Italy

ARTICLE INFO

Article history:
Received 28 November 2016
Received in revised form
23 January 2017
Accepted 24 January 2017
Available online 20 March 2017

Keywords:
GWAS
Phenylpropanoid pathway
SNPs
Phenolic acids genes
Wheat

ABSTRACT

Phenolic acids are major components of cell walls in wheat and have important implications on human health as antioxidants with anti-tumor activity. Our objectives were to identify phenolic acid genes in wheat by single nucleotide polymorphisms (SNPs) detected within the coding sequences of candidate genes, and to identify chromosomal regions associated with single phenolic acids and total soluble phenolic compounds. A set of candidate genes involved in the biosynthesis of hydroxycinnamic acid derivatives were identified by comparative genomics. SNPs found in the coding sequences of six genes (*PAL1*, *PAL2*, *C4H*, *C3H*, *COMT1* and *COMT2*) were used to determine their chromosomal location and accurate map position on two reference consensus linkage maps. The genome-wide association study (GWAS), based on genotyping a tetraploid wheat collection with 81,587 gene-associated SNPs, detected 22 quantitative trait loci (QTL) distributed on almost all durum wheat chromosomes. Two QTL for *p*-coumaric acid were coincident with the phenylalanine ammonia-lyase (*PAL2*) and *p*-coumarate 3-hydroxylase (*C3H*) genes on chromosome arms 2AL and 1AL, respectively. The availability of candidate gene-based markers can allow elucidating the mechanism of phenolic acids accumulation in wheat kernels and exploiting the genetic variability of phenolic acids content for the nutritional improvement of wheat end-products.

1. Introduction

Common wheat (*Triticum aestivum* L. subsp. *aestivum*, genome AABBDD, 2n = 6x = 42) and durum wheat (*Triticum turgidum* L. subsp. *durum* (Desf.) Husnot, genome AABB, 2n = 4x = 28) are polyploid species grown globally for the production of bread, pasta, couscous and other local food products. The high content of carbohydrate and protein make wheat-end products important in the human diet. Wheat species contain a variety of bioactive components, such as tocols, sterols, alkylresorcinols, folates, phenolic acids and fiber components (Shewry et al., 2010). In particular, phenolic acids have received great attention due to their health-promoting and disease-preventing value (review by Verpoorte et al., 2002). These compounds are associated in plant species with numerous biological functions, including photosynthesis, nutrient uptake, protein synthesis, structural components, seed dormancy, biotic and abiotic stress responses (Bravo, 1998). In cereals, phenolic acids are particularly abundant and include hydroxycinnamic acid derivatives (e.g. *p*-coumaric, caffeic, ferulic, and sinapic acids) and hydroxybenzoic acid derivatives (e.g. *p*-hydroxybenzoic, vanillic, syringic, and gallic acids) (Li et al., 2008). Phenolic acids in the wheat grain are typically located in the bran and germ fraction as soluble or esterified to saccharides and other low molecular mass components (e.g. organic acids), and primarily as insoluble bound forms linked to cell wall polymers. The general phenylpropanoid pathway is involved in the biosynthesis of

* Corresponding author.
E-mail address: antonio.blanco@uniba.it (A. Blanco).

http://dx.doi.org/10.1016/j.jcs.2017.01.022
0733-5210/

hydroxycinnamic-acid derivatives (reviewed by Vogt, 2010). Briefly, the hydroxylation of phenylalanine by phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H) and *p*-coumaric acid hydroxylase (C3H) leads to the formation of hydroxycinnamates and their activated forms. Methylation of caffeic acid, catalyzed by caffeic acid 3-O-methyltransferase (COMT), yields ferulic and sinapic acids (Supplementary Fig. 1). The hydroxycinnamic-acid derivative group of phenolics includes ferulic acid (4-hydroxy-3-methoxycinnamic acid), which is the most abundant phenolic acid of wheat at all stages of development, representing 90% of the total. The concentration of ferulic acid increases steadily during grain development, prior to a 50% decrease during grain ripening. This phenolic acid arises from the metabolism of phenylalanine and tyrosine and it is ubiquitously present in plant cell walls (Bravo, 1998).

The antioxidant properties of phenolic compounds are well known and extensively studied in a variety of plants (Rice-Evans et al., 1997). The presence of the CH=CH–COOH group in its structure is considered to be the key for significantly higher antioxidative efficiency compared to that of hydroxybenzoic acids. Significant correlations also are known between the level of phenolic compounds and the antioxidant activity of whole-meal semolina in large sets of durum wheat samples (Pasqualone et al., 2014). Phenolic compounds display antioxidant activity as terminators of free radicals by donating a hydrogen atom (Bravo, 1998). Moreover, phenoxy radical intermediates are resonance stabilized; therefore, a new chain reaction is not easily initiated. Phenolic compounds are subject to the activity of polyphenol oxidases (PPO) (E.C. 1.14.18.1), a class of enzymes that catalyse the oxidation of phenolics to quinones in presence of oxygen. In bread wheat, PPO causes undesired discoloration of oriental noodles and dough browning in durum wheats (Taranto et al., 2012).

Phenotypic variation in phenolic acid content is extensively studied in wheat germplasm and cultivars (Gawlik-Dziki et al., 2012; Laddomada et al., 2016; Li et al., 2008; Narwal et al., 2014; Pasqualone et al., 2014; Ragaee et al., 2012; Shewry et al., 2010; Yilmaz et al., 2015; Verma et al., 2008) and indicates that total and individual phenolic acid content are complex traits influenced by both genotype and environmental factors. The heritability estimates for total phenolic acids in winter and spring bread wheat genotypes was shown to be low due to the strong influence of environmental factors on the trait (Shewry et al., 2010). Actually, a recent study on tetraploid genotypes showed a higher ratio of genotypic variance to total variance both for individual and total phenolic acids, suggesting that it might be realistic to improve the trait in durum cultivars through appropriate breeding programs (Laddomada et al., 2016). The dissection of quantitative traits, such as phenolic acids, typically uses DNA-based molecular markers and biparental mapping populations. This approach requires developing specific segregating populations, and QTL detection is limited to loci segregating between crosses. Moreover, the detected QTL cover many cM and additional steps are required to narrow the QTL region or clone the genes. Linkage disequilibrium-based association mapping (AM) is a recent, alternative approach that uses a set of genotypes (germplasm accessions, breeding lines, cultivars) representing the products of hundreds of recombination cycles, thus providing higher resolution QTL mapping (Rafalski, 2010). The limitation of AM studies (genome-wide association study and candidate genes approaches) is the high frequency of false-positive and false-negative associations, which depend on population structure, relative kinship among individuals, and on multiple testing of thousands of markers.

The genetic control of phenolic acids in cereals has been investigated in rice, barley and sorghum, with studies limited to determining the total phenolic acid content and not considering individual phenolic compounds (Cai et al., 2015; Jin et al., 2009; Mohammadi et al., 2014; Rhodes et al., 2014). As far as we know, no study on QTL and genes coding for individual phenolic acids has been carried out in wheat. We recently explored the genetic variability of phenolic compounds in a core collection composed of 112 tetraploid wheat (*T. turgidum* L.) genotypes, about half of which were represented by durum cultivars and the remainder by landraces and wild types (Laddomada et al., 2016; Pasqualone et al., 2014). This core collection was derived from a larger set of 237 genotypes that were screened previously for genetic variability by Simple Sequence Repeat (SSR) and Diversity Arrays Technology (DArT) markers. The molecular data were submitted to cluster analysis to group the genotypes (Laidò et al., 2013), and a core collection was generated by picking genotypes from each cluster in order to maintain the genetic variability characterizing the full set.

The present study was designed to identify candidate genes for the hydroxycinnamic acid derivatives in wheat and to investigate association between regions of the durum wheat genome and the accumulation of individual phenolic compounds as well as total soluble phenolic components. With this aim, we analyzed the 112 core collection of tetraploid wheats using a molecular marker array including 81,587 gene-associated SNPs (Wang et al., 2014). The importance of identifying genes and QTL for phenolic acid composition and content in wheat grain is based on the lack of information on the genetic basis of phenolic acid metabolism in wheat. The characterization of key genes involved in the biosynthetic pathway of phenolic acids could enable the improvement of wheat cultivars by traditional and molecular breeding, and by further advanced biotechnology, such as metabolic engineering. Durum wheat cultivars with higher phenolic acid content would lead to end-products with enhanced health-promoting properties.

2. Material and methods

2.1. Plant materials

The set of 112 tetraploid wheat genotypes included 65 old and modern cultivars of durum wheat and various *T. turgidum* subspecies, namely subsp. *turgidum* (12 accessions), subsp. *turanicum* (8 accessions), subsp. *polonicum* (8 accessions), subsp. *carthlicum* (3 accessions), subsp. *dicoccum* (9 accessions) and subsp. *dicoccoides* (7 accessions). Plant material was grown in the experimental field of the University of Bari at Valenzano (Bari, Italy) in the 2011-12 and 2012-13 growing seasons in a randomized complete block design with three replicates and plots consisting of 1-m rows, 30 cm apart, with 50 germinating seeds per plot. During the growing season, 100 kg/ha of N was applied and standard cultivation practices were adopted. Plots were hand-harvested at maturity and grain samples from each plot were separately ground on a laboratory mill equipped with 1-mm sieve (Cyclotec Sample Mill, Tecator Foss, Hillerød, Denmark) to obtain wholemeal semolina.

2.2. Quantitative analysis of total phenolic compounds (soluble fraction)

The total soluble phenolic compounds (TSPC, composed of free phenolic acids and phenolics bound to low-molecular-mass molecules) were extracted and determined as in Pasqualone et al. (2014). Briefly, 1 mL of methanol was added to 0.1 g wholemeal semolina, purged with a stream of nitrogen, kept on orbital shaker at 200 rev/min for 2 h in the dark, and centrifuged at 7000 × g for 5 min. The recovered supernatant was subjected to the Folin-Ciocalteu reaction and subsequently measured at 765 nm by a Cary 60 UV–Vis spectrophotometer (Agilent Technologies Inc., Santa Clara, CA, USA). A calibration curve was built by methanol solutions of ferulic