Plant Science xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Plant Science

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

Yield QTLome distribution correlates with gene density in maize

Ana Karine Martinez[a,1], Jose Miguel Soriano[a,b,1], Roberto Tuberosa[a], Rachil Koumproglou[c], Torben Jahrmann[c], Silvio Salvi[a,*]

[a] Department of Agricultural Sciences, University of Bologna, Viale Fanin 44, 40127 Bologna, Italy
[b] Field Crops Programme, IRTA (Institute for Food and Agricultural Research and Technology), 25198 Lleida, Spain
[c] Semillas Fitó, Selva de Mar 111, 08019 Barcelona, Spain

ARTICLE INFO

Article history:
Received 26 June 2015
Received in revised form
22 September 2015
Accepted 23 September 2015
Available online xxx

Keywords:
Heterosis
Meta-analysis
QTL
QTLome
Yield
Zea mays

ABSTRACT

The genetic control of yield and related traits in maize has been addressed by many quantitative trait locus (QTL) studies, which have produced a wealth of QTL information, also known as QTLome. In this study, we assembled a yield QTLome database and carried out QTL meta-analysis based on 44 published studies, representing 32 independent mapping populations and 49 parental lines. A total of 808 unique QTLs were condensed to 84 meta-QTLs and were projected on the 10 maize chromosomes. Seventy-four percent of QTLs showed a proportion of phenotypic variance explained (PVE) smaller than 10% confirming the high genetic complexity of grain yield. Yield QTLome projection on the genetic map suggested pericentromeric enrichment of QTLs. Conversely, pericentromeric depletion of QTLs was observed when the physical map was considered, suggesting gene density as the main driver of yield QTL distribution on chromosomes. Dominant and overdominant yield QTLs did not distribute differently from additive effect QTLs.

1. Introduction

Progress in molecular biology and genomics platform allows us to identify the quantitative trait loci (QTL) that govern the expression of yield and other important agronomic traits, hence providing unprecedented opportunities to enhance the effectiveness of selection targeting the key loci via genomics-assisted breeding [1–4]. The number of mapped QTLs for a given trait and species, collectively known as QTLome [5] is growing at an impressive pace, prompting increased efforts in the synthesis and interpretation of QTL information.

In maize, grain yield is the most important and genetically complex trait and is generally modelled as controlled by a large number of small effect QTLs, in the contexts of both phenotypic and marker-assisted selection [6–8]. While breeding has certainly succeeded in improving grain yield in maize during the last century, maintaining the same rate of improvement will be increasingly more difficult due to more extreme climatic conditions and the need to adopt new cropping systems requiring reduced inputs [9]. For these reasons, a better understanding of the grain yield QTLome (in terms of number of QTLs, their map position, size and type of genetic effects and interactions) remains a priority in maize breeding.

One way to synthesize QTL information is by QTL meta-analysis [10,11] which allows one to identify the regions of the genome that more frequently govern trait variation and to narrow down the confidence intervals of QTLs by leveraging existing information. In maize, MetaQTLs (MQTLs) were described for flowering time [12,13], leaf architecture [14], grain moisture [15] and ear rot resistance [16]. Hao et al. [17] used meta-analysis to identify candidate genes potentially involved in drought tolerance networks. Meta-analysis was also applied to yield and yield components QTLs [18–20]. By collecting QTL [21,22] and transcriptomic [23] data from multiple studies, a predominant centromeric location of heterotic QTLs was observed.

This study reports the results of the largest MQTL analysis so far conducted for yield and yield components in maize. QTLs were projected on the genetic and physical maps and utilized to compute MQTLs. Our main objectives were to produce an informed repository of yield QTL information in maize and to search for patterns of QTL distribution on chromosomes, with particular attention to QTLs involved in yield heterosis.

---

[*] Corresponding author.
E-mail address: silvio.salvi@unibo.it (S. Salvi).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.plantsci.2015.09.022
0168-9452/

Table 1
List of trait categories utilized in this study.

Trait	Acronym	Traits included [a]
Grain yield [b]	GY	Grain yield
Kernel weight	KW	Hundred kernel weight Kernel weight Kernel weight per ear Thousand kernel weight
Ear row number	ERN	Ear row number
Kernel number	KN	Kernel number Kernel number per ear Kernel number per plant Kernel number per row
Ear length	EL	Ear length
Ear diameter	ED	Ear diameter
Ear number	EN	Ear number Ear number per plant
Ear weight	EW	Ear weight

[a] Traits as defined from original papers which were included in the same category in our study.

[b] Grain yield trait category strictly included QTL data for overall grain production per unit of area, with the exception of QTL information from studies [18] and [37], where ear weight per plant and grain weight per plant were provided and utilized, respectively.

2. Materials and methods

2.1. Bibliographic collection and construction of a maize yield QTLome database

Literature was retrieved from Web of Science (Thomson Reuters^{TM}) (http://apps.webofknowledge.com) using the keywords 'maize yield QTL' which eventually identified 44 manuscripts published from 1992 to 2014. While papers presented QTL data for multiple traits, we only considered yield and yield components traits which were classified in eight main categories (Table 1). A summary of the QTL studies is reported in Table 2.

The yield QTLome database included information on: parents of the cross, type of cross, number of progenies, name of QTLs, trait, LOD score, proportion of phenotypic variance explained (PVE) by each QTL, QTL position on the authors' linkage map in terms of LOD peak and QTL supporting or confidence interval (CI) (Supplementary material 1). In studies reporting multi-environment results, only QTLs from the overall analysis were considered; alternatively, if two QTLs for the same trait were reported in two experiments (e.g., two different water regimes) in the same study (i.e., in the same publication), and their map positions corresponded, they were considered as the same QTL and were included in our QTL database as one.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.plantsci.2015.09.022.

For each QTL, flanking molecular markers along with their genetic position in the original map and in the reference maize map 'Genetic'(see Section 2.2), were searched and recorded. In order to avoid heterogeneity in definition of *CI* across studies, 95% *CI* values were estimated using the approach described by [66] and extended by [67] according to the population type: $CI = 163/(N \times R^2)$ for recombinant inbred line (RIL) progenies, and $CI = 530/(N \times R^2)$ for the rest of non-RIL progenies (i.e., F_2, BC, TC), where N is the size of the population and $R2$ is the proportion of variance explained by the QTLs. Estimated *CI* values were utilized for QTL projection (Supplementary material 1).

When available, the degree of dominance (often indicated as gene action) for each QTL was recorded as provided by the authors. More specifically, QTLs were declared as heterotic QTLs (HQ) when gene action was reported as dominant (D, with dominant effect/additive effect >0.8) or overdominant (OD, with $d/a > 1$). QTL were declared as additive (AQ) when gene action was reported as additive or partially dominant ($0 < d/a < 0.8$). When not available, the degree of dominance was attributed on the basis of the type of experimental population, as follows: QTLs identified in RIL populations were considered as AQ; QTLs identified in mapping populations evaluated as testcrosses were classified as potentially heterotic QTLs (PHQ).

2.2. QTL projection and meta-analysis

QTLs were projected onto the 'Genetic' maize reference map available at http://www.maizegdb.org/complete_map?id=1203637 (see also [68]), using the software Biomercator v4.1 (http://moulon.inra.fr) [11,69]. The frequency (probability) of identification of QTL for every cM position in the 'Genetic' map was estimated following the approach described as 'QTL-overview index' [12]. Meta-analysis was carried out using BioMercator v. 4.1. The input file for BioMercator is provided as Supplementary material 2.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.plantsci.2015.09.022.

For *n* individual QTLs, BioMercator tests the most likely assumption between 1–4 and *n* underlying QTLs. Decision rules are based on an Akaike-type criterion (AIC) and the one with the lowest AIC value was considered the best fit. Consensus QTL from the optimum model is regarded as meta-QTL (MQTL). In order to study the QTL distribution based on physical distances QTLs were projected onto the B73 RefGen_v2 reference genome [70] (http://www.maizegdb.org/). Additionally, variation of QTL density across the maize genome was studied by counting QTLs on each 20-cM bin (roughly equivalent to 1/100 of the genetic map) and on each 25-Mb bin (again roughly equivalent to 1/100 of maize physical map), starting from the centromeric region of each chromosome (ie. the middle point of a centromere interval was considered as position 0). A QTL was assigned to a bin based on its LOD peak as reported in original studies. Centromeres positions were obtained from the http://www.maizegdb.org/ and from [71,72]. Gene density distribution was determined from the maize genome sequence B73RefGen_v2 available at http://www.maizegdb.org/. QTLs, MQTLs and overview index were visually represented using CIRCOS [73].

3. Results

3.1. Main features of QTL mapping experiments in maize

Our survey covered 44 studies which reported QTLs for yield and yield components in maize, based on biparental populations and published from 1992 to 2014 (Table 2). The studies covered 32 different experimental crosses, which utilized 47 inbred lines as parents, the most popular being B73 and Mo17 (six and five crosses, respectively). The most common cross types were $F_{2:3}$ or $F_{2:4}$ evaluated *per se*, RIL *per se* and $F_{2:3}$ or $F_{2:4}$ evaluated as testcross/backcross (14, eight and seven populations, respectively; Table 2).

Although in many cases the original papers reported on QTLs for multiple morpho-physiological traits, we focused our survey on QTLs for grain yield (GY) and seven additional GY components of common interest in maize breeding (Table 1). The QTL database eventually included 808 unique QTLs. The number of QTLs per trait ranged from 23 to 253 (ear number and GY, respectively; Supplementary material 3). The average number of QTL per study was 7.2 for GY, while it ranged between 4.6 and 7.5 for component traits (for Ear Number and Kernel Weight, respectively; Supplementary material 3).

Please cite this article in press as: A.K. Martinez, et al., Yield QTLome distribution correlates with gene density in maize, Plant Sci. (2015), http://dx.doi.org/10.1016/j.plantsci.2015.09.022