Plant Science 242 (2016) 310–329

Contents lists available at ScienceDirect

Plant Science

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

Breeding maize for silage and biofuel production, an illustration of a step forward with the genome sequence

Yves Barrière[a,*], Audrey Courtial[b,c], Anne-Laure Chateigner-Boutin[d], Dominique Denoue[a], Jacqueline Grima-Pettenati[b]

[a] INRA, UR889, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France
[b] LRSV, Laboratoire de Recherche en Sciences Végétales, UMR5546, Université Paul Sabatier Toulouse III / CNRS, Auzeville, BP 42617, 31326 Castanet-Tolosan, France
[c] INRA, US1258, Centre National de Ressources Génomiques Végétales, CS 52627, 31326 Castanet-Tolosan, France
[d] INRA, UR1268 Biopolymères, Interactions, Assemblages, 44316 Nantes, France

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form 4 August 2015
Accepted 13 August 2015
Available online 18 August 2015

Keywords:
Maize
QTL
Candidate gene
Cell wall
Digestibility
Bioethanol
Bioenergy
Silage

ABSTRACT

The knowledge of the gene families mostly impacting cell wall digestibility variations would significantly increase the efficiency of marker-assisted selection when breeding maize and grass varieties with improved silage feeding value and/or with better straw fermentability into alcohol or methane. The maize genome sequence of the B73 inbred line was released at the end of 2009, opening up new avenues to identify the genetic determinants of quantitative traits. Colocalizations between a large set of candidate genes putatively involved in secondary cell wall assembly and QTLs for cell wall digestibility (IVNDFD) were then investigated, considering physical positions of both genes and QTLs. Based on available data from six RIL progenies, 59 QTLs corresponding to 38 non-overlapping positions were matched up with a list of 442 genes distributed all over the genome. Altogether, 176 genes colocalized with IVNDFD QTLs and most often, several candidate genes colocalized at each QTL position. Frequent QTL colocalizations were found firstly with genes encoding ZmMYB and ZmNAC transcription factors, and secondly with genes encoding zinc finger, bHLH, and xylogen regulation factors. In contrast, close colocalizations were less frequent with genes involved in monolignol biosynthesis, and found only with the *C4H2*, *CCoAOMT5*, and *CCR1* genes. Close colocalizations were also infrequent with genes involved in cell wall feruloylation and cross-linkages. Altogether, investigated colocalizations between candidate genes and cell wall digestibility QTLs suggested a prevalent role of regulation factors over constitutive cell wall genes on digestibility variations.

1. Introduction

The biological conversion of plant cell wall carbohydrates into fermentable sugars is hindered by the embedding and/or cross-linkages of carbohydrates with lignins or *p*-hydroxycinnamic acids, inducing detrimental effects on both animal feeding and biofuel production. However, large genetic variations for cell wall digestibility have been shown and *in vivo* values in early and medium early maize hybrids indeed range between 36 and 60% [1]. Correlatively, a similar range from simple to double was shown for *in vitro* cell wall enzymatic digestibility which varies between lines or hybrids from 25 to 50% [2,3]. Lignins and cross- linkages between cell wall components are nevertheless essential to grasses as they contribute to mechanical properties of plant tissues, impart hydrophobicity to vascular elements allowing water and nutriment transportation, and also contribute to disease and pest resistance [3,4]. Breeding programs which allow to achieve a balance between agronomic and energy value traits are therefore necessary for the release of maize and grass varieties suitable for cattle feeding or bioenergy production.

Marker assisted selection (MAS) has proven to be extremely efficient to improve many traits in numerous species, and in particular for (grain) maize breeding. However, in most cases, MAS is a "blind" strategy as the underlying determinants associated with markers remain unknown. In addition, QTLs have been shown for numerous traits, but with similarly unknown underlying determinants. The availability of the sequence of the maize B73 line [5] allows breeders to identify genes surrounding marker or QTL physical positions. To improve cell wall digestibility, the knowledge of genes governing such a multi-factorial trait will especially enable to set up in a

* Corresponding author. Fax: +33 5 49 55 60 44.
E-mail address: yves.barriere@lusignan.inra.fr (Y. Barrière).

http://dx.doi.org/10.1016/j.plantsci.2015.08.007
0168-9452/

Nomenclature

Abbreviations

4CL	4-Coumarate- CoA ligase
C3H	*p*- Coumaroyl-shikimate/quinate 3-hydroxylase
C4H	cinnamate 4- hydroxylase
CAD	cinnamyl alcohol dehydrogenase
COMT	caffeic acid O-methyltransferase
CCoAOMT	caffeoyl- CoA O-methyltransferase
CCR	cinnamoyl-CoA reductase
CPK	calcium dependent protein kinase
CSE	caffeoyl shikimate esterase
F5H	ferulate 5- hydroxylase
FGPS	folylpolyglutamate synthase
HCT	hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
HMT	homocysteine S-methyltransferase
MeSe	methionine synthase
MMT	S-AdoMet Methionine S-methyltransferase
MTHFR	methylene- tetrahydrofolate reductase
PAL	phenylalanine ammonia lyase
pCAT	*p*- coumaroyl CoA:hydroxycinnamyl alcohol transferase
PMT	*p*- coumarate monolignol transferase
SAHH	S-adenosyl- homocysteine hydrolase
SAMS	S-adenosyl- methionine synthetase
CPK	calcium dependent protein kinase
H unit	*p*-Hydroxyphenyl lignin unit
G unit	guaiacyl lignin unit
S unit	syringyl lignin unit
pCA	*p*- coumaric acid
FA	ferulic acid
esterFA	esterified ferulic acid
etherFA	etherified ferulic acid
NDF	neutral detergent fiber
IVNDFD	*in vitro* NDF digestibility
cM	centiMorgan
Lod	logarithm (base 10) of odds
Mbp	megabase pair
MAS	marker assisted selection
QTL	quantitative trait locus
SNP	single nucleotide polymorphism
ZFN	zinc finger nuclease
TALEN	transcription activator-like effector nuclease
CRISPR/Cas	clustered regularly interspaces short palindromic repeat/CRISPR associated system

reasoned way the balance between quality traits and the possible detrimental effects on yield, standability, or disease and pest tolerances. The objective of the current research is to illustrate how the availability of the sequence of the maize genome coupled with a list of candidate genes can be used to highlight possible or probable genes underlying cell wall digestibility QTLs.

2. Material and methods

2.1. RIL progenies and IVNDFD cell wall digestibility

Investigations were focused on cell wall digestibility because it was previously shown [1] that the energy value of whole maize plants used for cattle feeding was first explained by this trait (nearly 60%) and secondly by starch content (nearly 30%). Six RIL progenies (MBS847 x F2, F288 x F271, Fl16 x F2, F838 x F286, RIo x WM13, F7025 x F4) were thus considered for cell wall digestibility QTL and candidate gene colocalizations [6–14] , and INRA - Génoplante unpublished data. According to progenies, RIL number ranged between 100 and 242 (average number 168), and marker number ranged from 94 to 189 (average number 136). Parental lines of the RIL progenies included early flint germplasm (F2, F4, F286, and Fl16) related to Pyrenean and Northern flint landraces, early and medium-early dent germplasm (F271, F288, F7025, MBS847, RIo) related to Canadian dent, Wisconsin dent and Iodent groups, the old Wisconsin WM13 which is one of the outstanding line bred from the Minnesota13 yellow dent open- pollinated variety before the Second World War, and the medium-late flint-dent F838 line that was bred for several traits including corn borer tolerance. Phenotypic data and samples were obtained from multi-location and multi-year trials focusing on cell wall digestibility and related traits. Plants were harvested at silage maturity stage in line *per se* experiments, either as whole plant (WP) or as plants deprived of ears at harvest (PWE), and also in topcross (TC) experiments. Cell wall content was estimated from dried and ground plant samples as Neutral Detergent Fibre (NDF) according to Goering and van Soest [15]. Enzymatic solubility (EnzSol) of the same plant samples was estimated according to Aufrère and Michalet-Doreau [16] and cell wall digestibility was estimated as IVNDFD (*in vitro* NDF digestibility) according to Struik [17] with IVNDFD = 100 x [(EnzSol - (100 - NDF)) / NDF]. In addition, because cell wall digestibility variations are significantly related to variations in lignin content, lignin unit acylation by *p*-coumaric acid, and intensity of ferulate cross-linkages, available QTL data for the three latter traits were simultaneously considered. Lignin content was estimated as Acid Detergent Lignin (ADL, according to Goering and van Soest [15]) and expressed as ADL/NDF. *p*-Coumaric acid (pCA), diferulic acid and etherified ferulic acid (diFA and etherFA, both corresponding to ferulate cross-linkages) contents were estimated after mild and severe alkaline hydrolyses of the cell wall, respectively [18]. In order to avoid confusing *p*-hydroxycinnamates from grain and stover, pCA, diFA, and etherFA contents were only considered in plants deprived of ears at harvest.

2.2. QTLs for IVNDFD cell wall digestibility in six RIL progenies

Based on gathered data from the six RIL progenies, 59 QTLs for the cell wall IVNDFD trait were considered with Lod values equal to or higher than 3.5 (and simultaneously with R2 values equal to or higher than 10% in most cases), including 32 QTLs from *per se* value experiments and 27 QTLs from topcross experiments. QTL physical positions were estimated based on the physical positions of the two nearest upstream and downstream flanking markers (marker physical positions according to the MaizeGDB database RefGen_v3), assuming a constant relationship between recombination and physical distances within this interval. The physical lengths of QTL support intervals were similarly estimated based on the physical positions of markers that were the closest to upstream and downstream limits of support intervals. However, due to greatly variable physical/genetic distance ratios along chromosomes (0.3–6.0 Mbp/cM), especially in areas close to or overlapping centromers, lengths of support intervals were difficult to estimate in several instances and should be considered with caution. Positions of QTLs for ADL/NDF, pCA, etherFA, and diFA considered for colocalizations with IVNDFD QTL were similarly estimated. Furthermore, the current QTL list, based on only six RIL progenies of early and medium-early lines, is likely not exhaustive.

2.3. Candidate gene list

The list of putative candidate genes, which could underlie cell wall digestibility QTLs, was established from a large synthesis of published investigations devoted to cell wall metabolism