

RESEARCH ARTICLE

Open Access



# Genome-wide association analysis of forage quality in maize mature stalk

Hongwu Wang<sup>†</sup>, Kun Li<sup>†</sup>, Xiaojiao Hu, Zhifang Liu, Yujin Wu and Changling Huang<sup>\*</sup>

## Abstract

**Background:** Plant digestibility of silage maize (*Zea mays* L.) has a large influence on nutrition intake for animal feeding. Improving forage quality will enhance the utilization efficiency and feeding value of forage maize. Dissecting the genetic basis of forage quality will improve our understanding of the complex nature of cell wall biosynthesis and degradation, which is also helpful for breeding good quality silage maize.

**Results:** Acid detergent fiber (ADF), neutral detergent fiber (NDF) and in vitro dry matter digestibility (IVDMD) of stalk were evaluated in a diverse maize population, which is comprised of 368 inbred lines and planted across seven environments. Using a mixed model accounting for population structure and polygenic background effects, a genome-wide association study was conducted to identify single nucleotide polymorphisms (SNPs) significantly associated with forage quality. Scanning 559,285 SNPs across the whole genome, 73, 41 and 82 SNPs were found to be associated with ADF, NDF, and IVDMD, respectively. Each significant SNP explained 4.2 %–6.2 % of the phenotypic variation. Underlying these associated loci, 56 genes were proposed as candidate genes for forage quality.

**Conclusions:** Of all the candidate genes proposed by GWAS, we only found a *C3H* gene (*ZmC3H2*) that is directly involved in cell wall component biosynthesis. The candidate genes found in this study are mainly involved in signal transduction, stress resistance, and transcriptional regulation of cell wall biosynthetic gene expression. Adding high digestibility maize into the association panel would be helpful for increasing genetic variability and identifying more genes associated with forage quality traits. Cloning and functional validation of these genes would be helpful for understanding the molecular mechanism of the fiber content and digestibility. These findings provide us new insights into cell wall formation and deposition.

**Keywords:** Genome-wide association, Maize, Forage quality, Digestibility

## Background

Maize (*Zea mays* L.) is not only an important staple crop feeding billions of people but also a main forage resource for animal rearing development. Forage maize is a type of high-energy silage that supplies dry matter, organic matter and cell walls of whole plants for ruminants. The forage feeding value is usually denoted as the utilization ratio of the transformation from forage constituents to energy used by the animal. Improving the feeding value is a major objective in forage maize breeding by increasing forage digestibility, especially cell wall digestibility.

Cell walls protect plants from pests and microbial infection [1, 2] and are involved in stress sensing and signal transduction [3]. In higher plants, cell walls are mainly composed of cellulose, hemicelluloses, pectins, proteins and lignin [4]. The amount and composition of these cell wall components differ among various plant cell types [5]. Variations in cell wall structure and composition have an essential effect on plant digestibility. Among the cell wall components, lignins are important for structure integrity of stalk tissues, and contribute to the mechanical support of plants [6]. However, association and cross linkages between lignins with other cell wall components greatly increase the resistance to degradation or stover fermentation. A balance must be maintained between a robust cell wall architecture and increasing forage digestibility. Thus, quantification of

\* Correspondence: [huangchangling@caas.cn](mailto:huangchangling@caas.cn)

<sup>†</sup>Equal contributors

Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China



cell wall related traits is necessary for forage quality breeding.

To estimate lignin content of forage plants, the acid detergent lignin/neutral detergent fiber (ADL/NDF) was first proposed by Goering and Van Soest (1970) [7]. NDF mainly consists of cellulose, hemicelluloses and lignins [8], which corresponds to the cell wall content [8, 9]. After hemicelluloses are solubilized with acid detergent treatment, cellulose and lignins are left as residual parts of cell wall, the main part of ADF [10]. Consequently, hemicellulose content can be determined by NDF minus ADF, and cellulose content is assumed to be the difference between ADF and ADL [10, 11]. For digestibility, IVDMD is a routine method to evaluate whole plant enzymatic solubility. And different IVDMD methods were developed by several studies [12]. However, the chemical methods for measuring fiber content and digestibility are complex to perform and costly in breeding programs. Prediction equations based on near infra reflectance spectroscopy (NIRS) was developed and widely used in estimating cell wall content and plant digestibility [10].

Due to the importance of silage maize in livestock farming, breeders began to select maize based on digestibility levels (preferring high digestibility), which resulted in improved cultivars and germplasm [13]. A set of brown midrib mutants with a reddish-brown midrib pigmentation was found to be responsible for reduced lignin concentration, composition [14] together with increased cell wall digestibility [15]. Among these mutants, *bm1*, *bm3* and *bm5* are known to increase enzymatic degradability of maize cell walls [15]. Until now, *bm1*, *bm2*, *bm3*, and *bm4* have been confirmed to encode cinnamyl alcohol dehydrogenase (CAD) [16, 17], methylenetetrahydrofolate reductase (MTHFR) [18], caffeic acid O-methyltransferase (COMT) [19] and folylpolyglutamate synthase (FPGS) [20], respectively. However, despite the contributions of lignification level to cell wall indigestibility, the correlation between lignin content and forage digestibility varies in different genetic backgrounds [12]. Penning et al. [21], identified several QTL related to lignin abundance, none of which were associated with enzymatic hydrolysis yield. In several recent studies, lignin levels was demonstrated to not correlate cell wall digestibility by enzyme in several species [22–24]. Thus, plant digestibility can not be improved by merely decreasing the lignin content. Many researchers began to study the genetic basis of cell wall-related traits and digestibility directly. Using linkage mapping, a large number of QTL for forage quality and cell wall digestibility were identified with multiple populations in previous studies [6, 8, 25–39]. Subsequently, a meta-analysis of QTL for plant digestibility and cell wall composition in maize was performed [10]. Twenty-six

meta-QTL for digestibility traits were detected using a consensus map of 11 experiments. Approximately 42 % of meta-QTL overlapped with QTL for cell wall, which coincided with trait correlations. Furthermore, 356 potential candidate genes for cell wall biosynthesis were mapped onto the consensus map, and 39 % of the candidate genes were located within meta-QTL confidence intervals. These studies proposed numerous potential associated loci and genes for silage quality, which need further investigation and validation.

Recently, genome-wide association studies (GWAS) have played an important role in dissecting complex quantitative traits in plants due to faster analyses, numerous high resolution markers, and abundant genomic and phenotypic variation [40]. Maize has extreme genetic and phenotypic diversity, with more rapid linkage disequilibrium (LD) decay than other species [41]. The rapid development of various genotyping technologies has aided the improved resolution of GWAS with tremendous numbers of markers. In maize, GWAS has become a powerful approach that can be successfully used in dissecting the genetic architecture for many traits, but it has not been performed to dissect traits related to forage quality traits in maize. In this study, with an association panel of 368 diverse inbred lines from around the world, we performed a GWAS analysis to dissect the genetic architecture of forage quality and to identify candidate genes for fiber content and plant digestibility.

## Methods

### Germplasm and field experiments

The association panel used in the present study contains 368 diverse inbred lines (AM368), including resources from the International Maize and Wheat Improvement Center (CIMMYT), China and the USA. Most of the lines from CIMMYT belong to tropical or sub-tropical germplasm sources. Detailed information about AM368 was provided in a previous study [42]. These inbred lines were planted in Hainan and Yunnan in 2010; in Hainan, Henan, and Yunnan in 2011; and in Hainan and Yunnan in 2012. A randomized block design was conducted at all locations, without replication. Each line was planted in a single row (2.5 m in length) of 11 plants at a density of 60,000 plants/ha. Adjacent rows were spaced 0.67 m apart.

### Phenotyping methods

In the present study, when each inbred line reached physiological maturity (i.e. a black layer appeared in kernels), the ears were harvested by hand. After harvest, the second to fifth internodes above the ground of six plants from each inbred line were collected with garden scissors. All samples were immediately enzyme-deactivated at 105°C for 30 mins in a forced air oven and air-dried

for 10–14 days. Dried stalk samples were ground with a mill and screened through a mesh size of 0.1 mm. Acid detergent fiber (ADF), neutral detergent fiber (NDF) and *in vitro* dry matter digestibility (IVDMD) were detected by NIRS [43, 44]. Before measurement, stalk samples were dried at 45°C for 48 h to exclude the influence of moisture. Samples were scanned through a near-infrared reflectance spectrophotometer (VECTOR22/N; BURKER Optik, Ettlingen, Germany). ADF, NDF, and IVDMD were determined using NIRS prediction equations developed for maize plants. A modified partial least squares approach implemented in OPUS 6.0 Bruker software was used for fitting the calibration equations [45]. The coefficients of determination for cross-validation ( $R_{CV}^2$ ) and external validation ( $R_{Val}^2$ ) were 93.6 % and 94.6 % for ADF, 95.3 % and 96.5 % for NDF, and 90.2 % and 91.2 % for IVDMD, respectively.

### Genotyping

The genotyping panel consisted of two sets: the MaizeSNP50 BeadChip, containing 56,110 SNPs, and 1.03 million high quality SNPs detected using RNA sequencing [42]. Included in the genetic analysis were 525,105 high-quality SNPs with minor allele frequency (MAF) above 0.05 [46]. After combining these two sets of genotypes and removing the duplicate SNPs, 559,285 SNPs (MAF  $\geq$  0.05) were used in the GWAS analysis. The information about genotype data was described in Fu et al. (2013) [46]. The genotype data can be downloaded at <http://www.maizego.org/Resources.html>.

### Statistical analyses

The GLM procedure was performed in SAS to dissect the variance of the phenotypes in different environments. The model used for the analysis of variance was  $y_i = \mu + e_i + f_i + \varepsilon_{li}$ , where  $y_i$  is the phenotypic value of the “i”th line,  $\mu$  is the grand mean of the target trait,  $e_i$  is the environmental effect of the “i”th environment,  $f_i$  represents the genetic effect of the “i”th line, and  $\varepsilon_{li}$  is denoted as the residual error. All the effects were considered as random effects. The variance components were used to calculate the broad sense heritability as  $H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2/e)$ , where  $\sigma_g^2$  represents the genetic variance,  $\sigma_e^2$  is the residual error variance item, and  $e$  is the number of environments. The 95 % confidence intervals of the  $H^2$  were calculated following the method of Knapp et al. [47]. The Pearson correlation coefficients between traits were computed using the PROC CORR procedure in SAS.

To eliminate the environmental effect within multiple environments, we fitted a mixed linear model to calculate the best linear unbiased prediction (BLUP) value for each line:  $y_i = \mu + g_i + e_i + \varepsilon_i$ . In this equation,  $y_i$  represents the phenotype of the “i”th line,  $\mu$  is the grand

mean value of the target trait in all environments,  $g_i$  represents the genetic effect,  $e_i$  is the environmental effect, and  $\varepsilon_i$  is the random error. The grand mean was fitted as fixed effect, and genotype together with environment were considered as random effects. BLUP estimation was obtained by using the MIXED procedure (PROC MIXED) in SAS9.3 (SAS Institute), which should be denoted as the sum of the grand mean and genetic effect of each line. The BLUP values of each line were used as the phenotypic values for association mapping.

### Genome-wide association analysis

A genome-wide association analysis was performed by accounting for the population structure and kinship matrix [48], which were estimated in a previous study [42]. Using ‘no compression’ and ‘population parameters previously determined’ (P3D) algorithms, the MLM program in TASSEL 4.1.26 [49] was utilized to detect the association between the phenotype and genotype. All the candidate genes were annotated according to the information available in MaizeSequence ([http://ensembl.gma.mnsc.umn.edu/Zea\\_mays/Info/Index](http://ensembl.gma.mnsc.umn.edu/Zea_mays/Info/Index)) and in the MaizeGDB database (<http://www.maizegdb.org/gbrowse>).

### Threshold for GWAS

Background LD between the high-density SNPs used in the present study can be problematic for multiple testing [50]. In addition, the cutoff determined by the Bonferroni tests seemed to be too strict for the present study. A less stringent cutoff of  $1 \times 10^{-4}$  was applied for the detection of significant associations. To identify the most robust and stable associations, another round of GWAS was performed with a subsampling procedure [51]. Subsampling was performed 100 times with R software [52], and 80 % of inbred lines were randomly sampled each time. Finally, only the SNPs detected as significant at  $P < 1 \times 10^{-4}$  and identified at least 10 times in 100 subsamples (resample model inclusion probability (RMIP) threshold of 0.1) were denoted as significant associations.

### Linkage disequilibrium and candidate gene selection

The linkage disequilibrium measure ( $r^2$ ) with each SNP significantly associated with forage quality traits in a region of 1 Mb were estimated in TASSEL 4.1.26 [49]. We also examined the LD between significant SNPs for each trait. If the LD between adjacent peak SNPs  $> 0.2$ , they will be grouped into one unique locus. The genes containing or adjacent to significant SNPs were identified using the maizeB73 reference genome assembly v2 available on the MaizeGDB Genome database ([http://gbrowse.maizegdb.org/gb2/gbrowse/maize\\_v2/](http://gbrowse.maizegdb.org/gb2/gbrowse/maize_v2/)). The closest gene of each peak SNP was proposed to be the most likely candidate gene.

## Results

### Phenotypic variability, correlation and heritability

The phenotypic variation of ADF, NDF, and IVDMD in the diverse association panel were assessed using 368 inbred lines in seven environments. The BLUP value of three forage quality traits showed abundant variation, which ranged from 27.9 % to 43.4 % (ADF), 54.7 % to 74.3 % (NDF), and 25.1 % to 56.3 % (IVDMD) with averages of  $35.2 \pm 2.8$  %,  $64.1 \pm 3.6$  %, and  $43.2 \pm 7.0$  %, respectively (Table 1 and Fig. 1). According to a previous study [53], the association panel can be divided into four sub-groups, including Stiff stalk (SS), Tropical–subtropical (TST), Non-stiff stalk (NSS), and Mixed groups. The average ADF and NDF content in the SS group is relatively higher than that of the other three groups. In addition, the dry matter digestibility of the SS group is significantly lower than that of the NSS and MIXED group (Table 1 and Fig. 2). The population structure contributes 3.13 %, 3.09 % and 6.78 % of the variance for ADF, NDF, and IVDMD, respectively.

The results of an analysis of variance showed that both genotype and environmental effects significantly affected the forage quality traits ( $P < 0.01$ ). Using the phenotypic value across all environments, the broad sense heritability of each trait was calculated. The heritability estimates of ADF, NDF, and IVDMD were 0.81, 0.79, and 0.85, respectively (Table 2). Significantly positive correlation between ADF and NDF was observed using a correlation analysis ( $r = 0.90$ ,  $P < 0.01$ ), while IVDMD showed a strong negative correlation with ADF and NDF ( $r = -0.81$ ,  $r = -0.74$ ,  $P < 0.01$ ).

### GWAS of forage quality in mature maize stalk

Integrated with the population structure and familial relatedness, a mixed linear model was fitted to scan 559,285 SNPs across the whole genome. Quantile-quantile plots (QQ plots) implied that the population structure and familial relatedness were well controlled in the GWAS of each trait (Additional file 1: Figure S1). At a significance level of  $P < 1 \times 10^{-4}$  together with a threshold of RMIP  $\geq 0.1$ , 73, 41 and 82 SNPs were identified as being associated with ADF, NDF, and IVDMD, respectively (Fig. 3, Additional file 1: Figure S1, Table 1). For all three traits, the phenotypic variance explained by each

allele ( $R^2$ ) ranged from 4.2–6.2 %. SNPs that were significantly associated with ADF (correspond either to cellulose or lignin), which covered 24 unique significant loci, were distributed on all chromosomes except 3 (Fig. 3a and Additional file 2). In addition, 14 loci were found to be significantly associated with NDF (correspond to hemicellulose, cellulose, and lignin) (Fig. 3b), which were located on chromosomes 1, 4, 5, 6, 7, 8 and 9. The maximum number of significant loci for ADF and NDF were both observed on chromosome 4. As shown in Manhattan plots, a set of consecutive SNPs on chromosome 4 were significant for both ADF and NDF (Fig. 3a, Fig. 3b). The SNPs significantly associated with IVDMD were located on all chromosomes except 10. Of the significant SNPs associated with IVDMD, 60 % were located at the distal end of chromosome 6, generating a “hot spot” associated with forage digestibility.

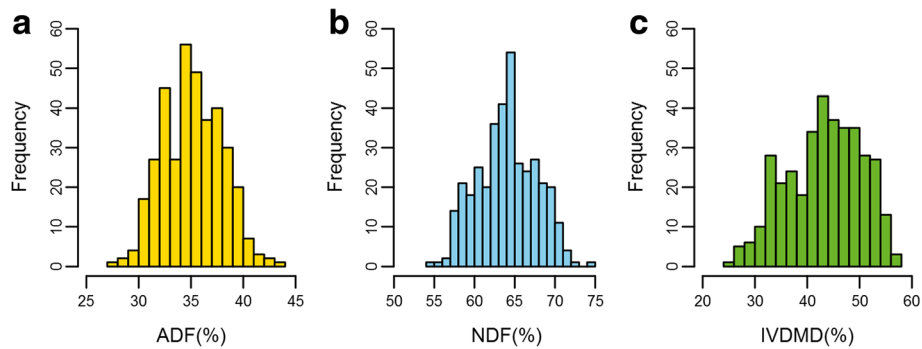
### Candidate gene selection and co-localization between different traits

The MaizeGDB genome database ([http://gbrowse.maizegdb.org/gb2/gbrowse/maize\\_v2/](http://gbrowse.maizegdb.org/gb2/gbrowse/maize_v2/)) of the maize B73 reference genome assembly v2 was used to annotate the genes that were near the significant SNPs [54]. After excluding repeated loci caused by co-localization between traits, a total of 56 genes were proposed as candidate genes for the three traits. Among these candidate genes, 10 encode uncharacterized proteins, and the proteins encoded by the remaining genes include transcription factors (GRF, LIM, ARR-B, NAC, bHLH, SBP, and hb3-type transcription factor families), enzymes involved in stress resistance, protein metabolism, signal transduction, cell wall biosynthesis and other biological processes.

In the GWAS analysis, twelve loci were identified as co-localized associations, which might be associated with two or three traits simultaneously. A strongly associated SNP (chr6.S\_155653406) is contained in the gene *GRMZM2G140817(ZmC3H2)*, which encodes a coumarate-3-hydroxylase that catalyzes hydroxylation reactions of the aromatic ring in the monolignol biosynthesis pathway [55] and significantly associated with the three traits studied herein. This SNP was significantly associated with the three traits (Fig. 3, Additional file 2) and located in the first intron of *ZmC3H2* but

**Table 1** Phenotypic variation of forage quality traits in the association panel and subpopulations

	ADF(%)		NDF(%)		IVDMD(%)	
	Range	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range	Mean $\pm$ SD
BLUP	27.9–43.4	$35.2 \pm 2.8$	54.7–74.3	$64.1 \pm 3.6$	25.1–56.3	$43.2 \pm 7.0$
MIXED	30.7–43.4	$35.2 \pm 3.0$	57.2–74.3	$64.1 \pm 4.0$	25.1–53.9	$44.4 \pm 7.1$
NSS	28.0–42.0	$34.7 \pm 2.8$	55.8–71.4	$63.8 \pm 3.4$	27.0–56.2	$45.5 \pm 6.6$
SS	30.7–41.0	$36.3 \pm 2.4$	57.7–70.5	$65.8 \pm 3.2$	28.9–49.8	$40.4 \pm 6.2$
TST	27.9–42.1	$35.2 \pm 2.8$	54.7–72.3	$63.8 \pm 3.7$	26.9–56.3	$42.2 \pm 7.1$



**Fig. 1** Distribution of forage quality traits in association panel. The frequency distribution of ADF, NDF, and IVDMD in association panel are shown in **a**, **b** and **c**, respectively

did not affect alternative splicing. Relatively moderate LD ( $R^2 < 0.6$ ) was observed between adjacent SNPs with the leading SNP (Fig. 4a). On chromosome 4, a significant SNP (PZE-104075114) was found to be associated with both ADF and NDF (Fig. 4b). The corresponding candidate gene containing this SNP encodes a LIM transcription factor (*GRMZM2G134752*) that regulates the expression level of the genes involved in the lignin biosynthesis pathway [43]. In addition to the two genes mentioned above, a candidate gene for the significant SNP on chromosome 7 at position 19,347,352 (*GRMZM2G042627*) encodes a kinase associated protein phosphatase (KAPP) that is related to the reaction response to pathogen attack and other stresses [56]. Additionally, another leading SNP which is located 15 bp downstream of chr7.S\_19347352 was found to be associated with IVDMD (Fig. 4c). These two leading SNPs were both located in the last exon of gene *GRMZM2G042627*, and the level of LD between these two SNPs was high ( $r^2 = 0.93$ , Fig. 4c).

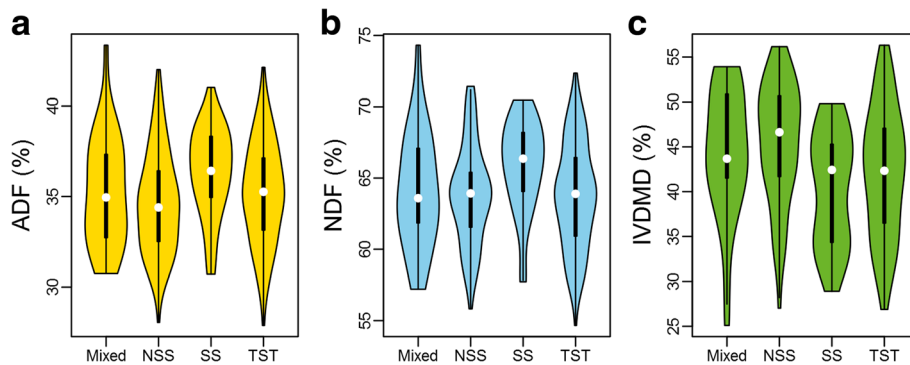
In addition to the three co-localized candidate genes mentioned above, nine more genes were also found to be common candidate genes (Additional file 6). *GRMZM2G331833*, which encodes CLP protease regulatory subunit X, was identified as candidate

genes for both NDF and IVDMD. In addition to the LIM transcription factor previously described, the co-localizations between ADF and NDF encode a small ribosomal subunit protein, an aspartic proteinase nepenthesin, an ARR-B-transcription factor, an F-box domain and LRR containing protein and an uncharacterized protein. Furthermore, three more genes were detected as candidate genes for both ADF and IVDMD; the encoded proteins included a GRF transcription factor, a mitogen-activated protein (MAP) kinase, and an uncharacterized protein.

## Discussion

### Phenotypic variation, heritability, and correlation

In the present study, a panel consisting of 368 inbred lines collected from all over the world was used to dissect the genetic architecture of forage quality traits in mature maize stalk. Approximately 1.6-, 1.3- and 2.2-fold variations were observed for ADF, NDF, and IVDMD, respectively. The huge variations in association mapping population suggests that natural germplasm with a broad genetic base could be a potential resource for improving forage quality. The sample size and phenotypic variation in the current study are larger than in the previous studies of associations between



**Fig. 2** Phenotypic variation of forage quality traits in sub-groups of association panel. The phenotype distribution of ADF, NDF, and IVDMD in each sub-group are shown in **a**, **b** and **c**, respectively. Mixed: Mixed group; SS: Stiff stalk; NSS: Non-stiff stalk; TST: Tropical and sub-tropical

**Table 2** Analysis of variance, heritability and correlation

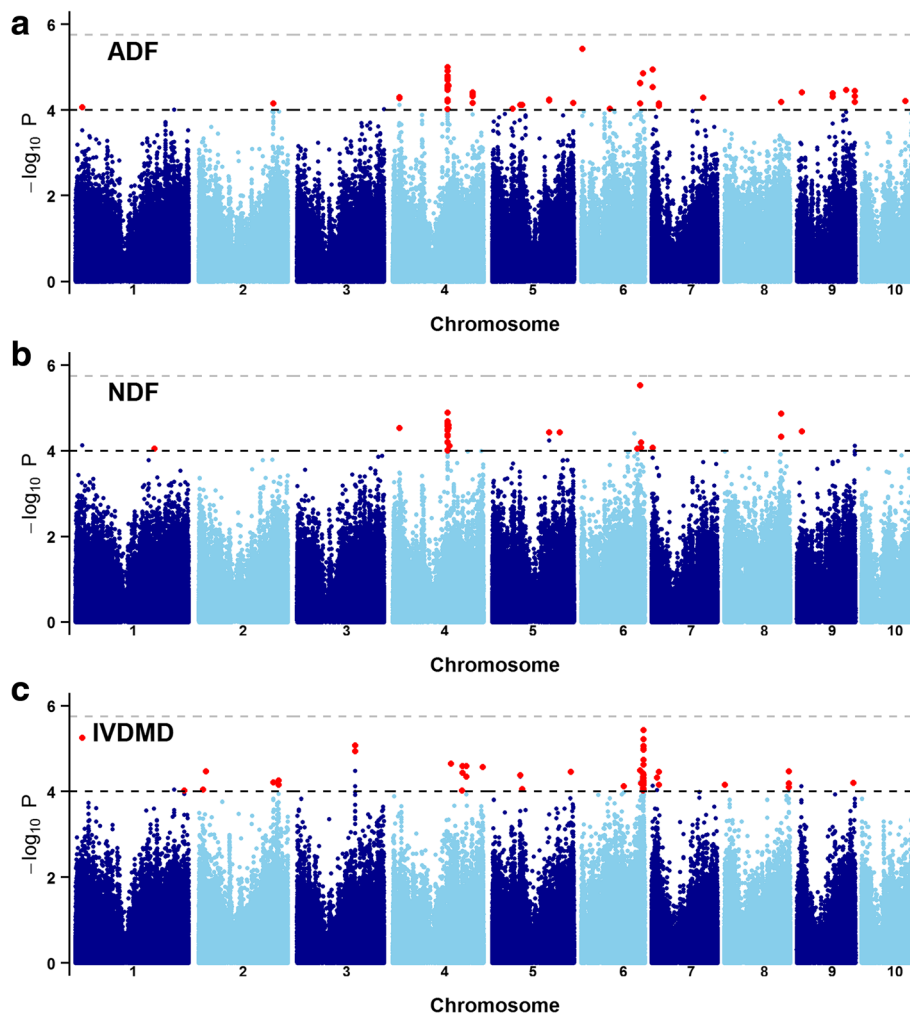
Traits	Mean Squares		$(H^2)^b$	CI <sup>c</sup>	Correlation		
	E <sup>a</sup>	G <sup>a</sup>			ADF	NDF	IVDMD
ADF	4404.92**	68.63**	0.81	0.78–0.83	1		
NDF	12449.61**	116.44**	0.79	0.76–0.82	0.90**	1	
IVDMD	16051.09**	393.83**	0.85	0.83–0.87	-0.81**	-0.74**	1

\*\*Significant at  $P < 0.01$ <sup>a</sup>Mean square values for environmental and genotypic factors<sup>b</sup>Broad-sense heritability<sup>c</sup>95 % confidence interval of broad-sense heritability

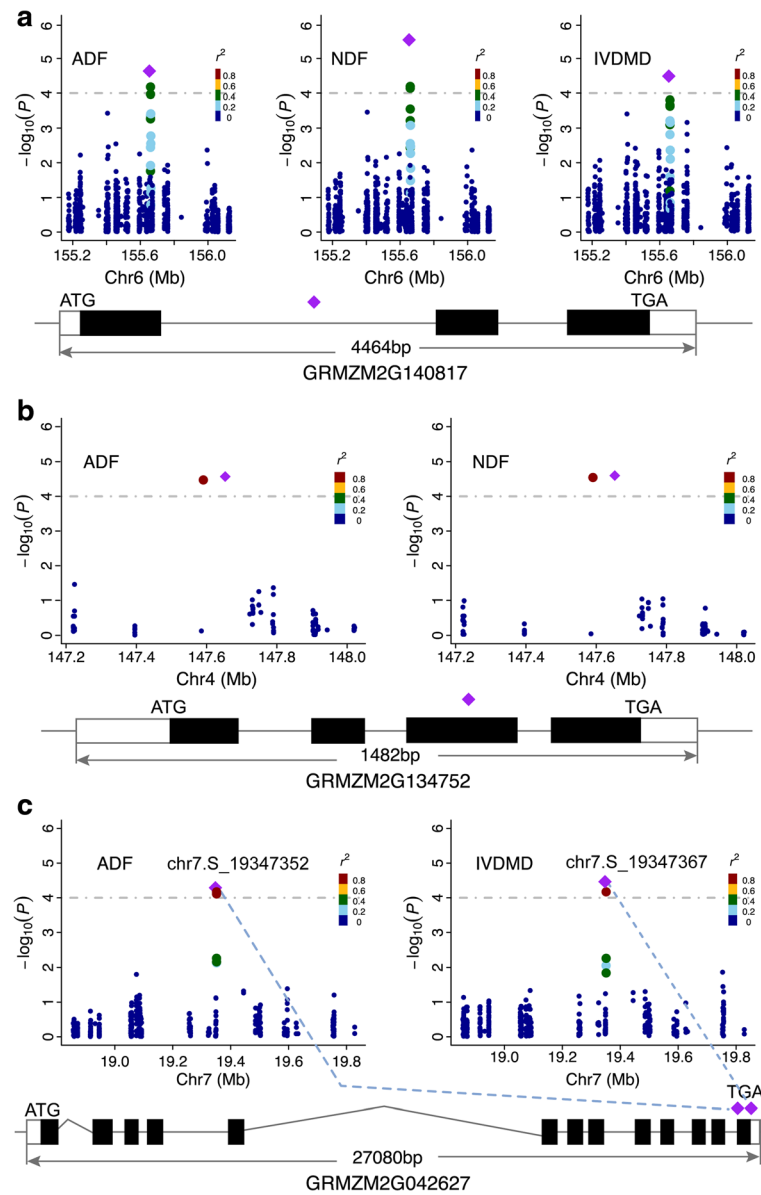
lignin pathway genes and forage quality traits [57–60]. According to the population stratification, the association panel was divided into four sub-groups [53]. We found relatively higher fiber content and lower digestibility in the SS group, which contains 45 lines with high stalk strength (data not shown) and is represented by B73, a classic line

from the Iowa Stiff Stalk Synthetic (BSSS) maize populations. These results suggest that population structure has an effect on forage quality and confirm the correlation between stalk strength and cell wall-related traits.

The phenotypic variation of the three traits across seven environments were dissected by an analysis of variance. Both genotype and environmental effects were significant for the three traits. In the previous studies about forage quality traits, heritability estimates ranged from 0.51 to 0.92 [30, 31, 36, 38, 39, 57]. In the present study, the heritability estimates for ADF, NDF, and IVDMD in the current study were relative high, 0.79 to 0.85, consistent with previous studies. These results reveal the strong genetic bases of these traits. A significant negative correlation between IVDMD and fiber content ( $r = -0.81$  for ADF and  $-0.74$  for NDF) was found, which suggests that high fiber content genotypes will be difficult to digest.



**Fig. 3** GWAS results for forage quality traits in a maize association panel. Manhattan plots for ADF, NDF, and IVDMD are shown in **a**, **b** and **c**, respectively. Grey and black dashed lines correspond to the thresholds of Bonferroni correction ( $P < 1.8 \times 10^{-6}$ ) and  $P < 1 \times 10^{-4}$ . Red dots indicate significant SNPs ( $P < 1.0 \times 10^{-4}$  and RIMP > 0.1) associated with each cell wall-related trait



**Fig. 4** Association and genomic location of known and new loci associated with forage quality traits. **(a-c)** Three identified genes were associated with forage quality traits. *GRMZM2G140817* (*ZmC3H2*) **(a)** was associated with ADF, NDF, and IVDMD. *GRMZM2G134752* **(b)** was associated with ADF and NDF. *GRMZM2G042627* **(c)** was associated with ADF and IVDMD, with different leading SNPs. (Top) A 0.5-Mb region on each side of the leading SNP, which is denoted by a purple diamond. The color of the remaining SNPs reflects the  $r^2$  values with the most significantly associated SNP. Dashed horizontal lines depict the significance threshold ( $1 \times 10^{-4}$ ). (Bottom) Gene structure according to the information from the B73 genome sequence in the GRAMENE database ([http://ensembl.gramene.org/Zea\\_mays/Info/Index](http://ensembl.gramene.org/Zea_mays/Info/Index))

### Genetic architecture of forage quality in mature maize stalk

Population structure can create unexpected LD between loci on separate chromosomes and lead to spurious associations for GWAS [41]. Although population structure contributes only a small proportion of the phenotypic variation in the present study, we still included population structure and familial relatedness matrices into the mixed model of the GWAS. As shown in Quantile-Quantile plots (Additional file 1: Figure S1),

false positive associations were well controlled for the GWAS of ADF, NDF, and IVDMD.

GWAS have been a powerful tool for studying the genetic basis of complex traits in maize. Because of the rapid decay of linkage disequilibrium in the association panel, the genome of maize was broken up into small LD blocks. The resolution for identifying an association with ultra-high density SNPs is at the gene level. In the present study, we scanned 559,285 SNPs to identify

SNPs associated with ADF, NDF, and IVDMD. With this high-density genotype data, background LD between SNPs may be a problem for multiple testing. The threshold determined by the Bonferroni correction seemed to be too strict for detecting significant associations. To resolve this problem, a suggestive threshold that is lower than the Bonferroni correction with a subsampling procedure were employed in previous studies [61, 62]. In the current study, using  $P < 1 \times 10^{-4}$  and RMIP  $> 0.1$  as the significance level, we found 56 unique loci that were significantly associated with forage quality traits. The phenotypic variation explained by each locus ranged from 4.2 %–6.2 %, which revealed that the genetic basis of the three traits in our association panel are mainly controlled by a number of minor effect quantitative trait genes (QTG).

Up to now, hundreds of QTL related to cell wall related traits have been identified across the whole genome of maize [8, 25, 28, 30–32, 34–39, 63, 64]. The previous studies suggest that the genetic architecture of these cell wall-related traits were controlled by a few major QTL and a large number of minor effect QTL, and these cell wall and digestibility QTL cover 77 % and 58 % of the maize genome, respectively [10]. Therefore, it is not surprising that most of the associations detected in the current study were co-localized with previously identified QTL regions. On the other hand, no silage maize was contained in the association panel of this study. The frequency of the favorable allele in silage maize may be reduced in our GWAS analysis. These rare favorable alleles generally explain a large proportion of the phenotypic variance [65]. However, association mapping has limited statistical power for detecting the effects of rare alleles [66], which may lead to a lack of major effect associations for the traits studied herein. In conclusion, combining the results of GWAS with linkage mapping or adding silage germplasm into the association population would be helpful to better understand the genetic architecture of cell wall-related traits.

#### Potential candidate genes and underlying pathways

Identifying the genes associated with forage quality is crucial for understanding the molecular mechanism of cell wall biosynthesis. Among the candidate genes found in the present study, *GRMZM2G140817* (*ZmC3H2*), which encodes a *p*-coumarate 3-hydroxylase, was found to be related to ADF, NDF, and IVDMD. This enzyme belongs to the CYP98 cytochrome P450 family and catalyzes the hydroxylation reaction of the aromatic rings of *p*-coumaric acid, then converts the substrate to caffeic acid (Additional file 3: Figure S2). *C3H* provides a watershed between the non-methoxylated *p*-hydroxyphenyl (H) branch and guaiacyl (G)/syringyl (S) branch in the lignin pathway. In *Arabidopsis*, a *C3H*-deficient mutant

was reported to be displaying a *reduced epidermal fluorescence (ref)* phenotype [67]. This mutant causes a decrease in the lignin content and produces lower amounts of G and S monolignols than the wild type. The downregulation of *C3H* in poplar and alfalfa caused an increase of H lignin units and a reduction of the total lignin content [68–70]. Other than the dicotyledonous species, *C3H* down-regulation was performed in maize in a recent study [71]. *ZmC3H1* knock-down maize plants were generated and led to a moderate increase of *p*-hydroxyphenyl (H) lignin subunits. The authors concluded that *ZmC3H2* might compensate for the reduced level of *ZmC3H1* in these *C3H1* repressed plants. In view of all these results, the two *C3H* genes may together play an important role in the lignin biosynthesis of maize. The functional validation of *C3H* genes should be performed to study the mechanism of how these two genes affect lignin accumulation in maize.

Other than the key enzyme genes participating in the cell wall biosynthesis pathway, the modulation of the transcription level also significantly affects the rate of flux through the secondary cell wall biosynthesis pathway [72]. A number of transcription factors have been found to be involved in the transcriptional network and responsible for regulating the lignin biosynthetic genes [72–75]. The AC-rich element is a common *cis* element contained in the promoter region of most lignin biosynthetic genes. This type of transcription activator is recognized by the DNA-binding domain of R2R3-MYB transcription factors [76–78]. In addition to the MYB transcription factor, a member of the LIM family transcription factor (*NtLIM1*) has been reported to bind with AC elements and to regulate the expression of lignin biosynthetic genes in tobacco [79, 80]. The suppression of *NtLIM1* led to a reduction in the expression level of *PAL*, *4CL*, and *CAD*, which caused a 27 % decrease of lignin content in antisense lines [80]. Furthermore, a group of NAC transcription factors were proven to regulate secondary wall biosynthesis and deposition in various lignification tissues of *Arabidopsis* [81–85]. In the present study, other than one LIM transcription factor (*GRMZM2G134752*) and two NAC transcription factors (*GRMZM2G031200*, *GRMZM5G857701*), five more transcription factor genes (*GRF*, *ARR-B*, *bHLH*, *SBP*, and *hb3*-type transcription factor families) were found to be related to fiber content and digestibility (Additional file 2). These results demonstrate the importance of transcriptional regulation in cell wall biosynthesis. Due to the complexity of the process for cell wall biosynthesis and decomposition, the function of five newly found transcription factor genes should be fully validated with a molecular biology analysis.

Plant cells provide not only mechanical strength for the plant stem but also barriers for attacks by insects



and diseases [1]. A strong phenotypic correlation between stress resistance and cell wall-related traits was observed during the selection process of breeding programs [86–88]. QTL co-localization also revealed a significant genotypic correlation between cell wall composition and resistance to pests and diseases [1]. Recently, two key enzymes in lignin biosynthesis (HCT and CCoAOMT) were reported to cooperate with the NLR Rp1 protein to regulate the hypersensitive defense response in maize [89]. In the current study, several candidate genes for cell wall-related traits were found to be involved in stress resistance and signal transduction, which includes protein responses to pest, low sulfur, and drought (Additional file 6). A kinase-associated protein phosphatase gene (KAPP) (*GRMZM2G042627*), which was significantly associated with ADF and IVDMD, was also found to be associated with resistance to herbivore attack [62]. These results provide evidence for a connection between cell wall formation and various stress responses.

In previous candidate gene-based association studies, ten lignin biosynthetic genes were investigated to detect associations between polymorphisms and forage quality traits [57–60]. The results showed that polymorphisms within the *PAL*, *ZmC3H1* and *F5H* were associated with forage quality traits. However, these associations were affected by population structure and multiple tests. Only individual polymorphisms in the *4CL1*, *CCoAOMT2* and *ZmPox3* were stably associated with digestibility [57, 59, 60]. In the present study, we did not find any other gene involved in lignin biosynthesis in association with ADF, NDF or IVDMD with the exception of *ZmC3H2*. In addition, the candidate genes identified in the current study mainly included genes encoding transcription factors, protein kinases, and proteins involved in response to stress and other biological processes. Therefore, we conclude that in diverse association panels, only genetic variants of several key enzyme genes in the lignin pathway may play roles in controlling the genetic basis of forage quality. More phenotypic variation of the forage quality may be attributed to the transcriptional regulation of lignin biosynthesis and genes involved in other cell wall biosynthesis pathways.

## Conclusions

In the present study, a genome-wide association study for forage quality traits was conducted with nearly 560 thousand SNPs. There were 24, 14 and 31 loci found to be associated with ADF, NDF, and IVDMD, respectively. Each of these loci contributes a small proportion of the phenotypic variance. Twelve loci were found to be co-localized for the three traits. Excitingly, underlying a significant associated SNP on chromosome 6, a *C3H* gene, which catalyzes a key step in the lignin pathway, was proposed as a candidate gene for all the three forage

quality traits. Most of the candidate genes underlying the associated loci were involved in transcriptional regulation of cell wall biosynthesis and the response to stress, suggesting that the transcriptional regulation of cell wall biosynthesis may play a vital role in cell wall formation and deposition. In conclusion, absence of silage maize in the association panel in the present study may be responsible for the lack of cell wall component biosynthesis genes found using GWAS. This study improves our understanding of the genetic basis of forage quality traits and provides insight into potential areas to improve forage digestibility. Future research will focus on cloning and validating the function of the potential candidate genes.

## Additional files

**Additional file 1:** Quantile-Quantile plots for the GWAS results for ADF, NDF, and IVDMD. The QQ plots for ADF, NDF, and IVDMD are shown in a, b and c, respectively. The horizontal grey solid line and black dashed line correspond to the thresholds of the Bonferroni correction and  $P = 1 \times 10^{-4}$ . (TIF 361 kb)

**Additional file 2:** Summary of leading SNPs and candidate genes significantly associated with forage quality traits. (XLSX 16 kb)

**Additional file 3:** The phenylpropanoid pathway, adapted from previous studies [55, 89]. 4CL, 4-hydroxycinnamoyl-CoA ligase; C3H, *p*-coumarate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; CAD, cinnamyl-alcohol dehydrogenase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; COMT, caffeic/5-hydroxyferulic acid O-methyltransferase; F5H, ferulate 5-hydroxylase; HCT, hydroxycinnamoyl transferase; PAL, phenylalanine ammonia lyase. (TIF 202 kb)

**Additional file 4:** Summary of the SNPs significantly associated with ADF. (XLSX 11 kb)

**Additional file 5:** Summary of the SNPs significantly associated with NDF. (XLSX 10 kb)

**Additional file 6:** Summary of the SNPs significantly associated with IVDMD. (XLSX 12 kb)

## Abbreviations

ADF: Acid detergent fiber; BLUP: Best linear unbiased prediction; BSSS: The Iowa stiff stalk synthetic; C3H: Coumarate-3-hydroxylase; CAD: Cinnamyl alcohol dehydrogenase; CIMMYT: The international maize and wheat improvement center; COMT: Caffeic acid O-methyltransferase; FPGS: Folylpolylglutamate synthase; GWAS: Genome-wide association studies; IVDMD: In vitro dry matter digestibility; KAPP: Kinase associated protein phosphatase; LD: Linkage disequilibrium; MAF: Minor allele frequency; MTHFR: Methylenetetrahydrofolate reductase; NDF: Neutral detergent fiber; NIRS: Near-infrared reflectance spectroscopy; NSS: Non-stiff stalk; P3D: Population parameters previously determined; QTG: Quantitative trait genes; QTL: Quantitative trait locus/loci; ref: Reduced epidermal fluorescence; RMIP: Resample model inclusion probability; SNP: Single nucleotide polymorphism; SS: Stiff stalk; TST: Tropical-subtropical

## Acknowledgements

We greatly appreciate Dr. Xiaohong Yang at China Agricultural University for providing the maize association panel and its genotypes and giving valuable suggestions to improve this manuscript.

## Funding

This study was supported by the National Natural Science Foundation of China (Grant No.31171561), the Agricultural Science and Technology

Innovation Program at CAAS and the National Basic Research Program of China (973 Program) (Grant No.2014CB138200).

#### Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its Additional files 1, 2, 3, 4, 5 and 6.

#### Authors' contributions

HW and KL carried out the experiments, analyzed the data and wrote the manuscript; XH, ZL and YW assisted in data collection and analysis; and CH designed the study and helped write the manuscript. All authors have read and approved this manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Consent for publication

Not applicable.

#### Ethics approval and consent to participate

Not applicable.

Received: 28 June 2016 Accepted: 12 October 2016

Published online: 21 October 2016

#### References

- Santiago R, Barros-Rios J, Malvar RA. Impact of cell wall composition on maize resistance to pests and diseases. *Int J Mol Sci.* 2013;14(4):6960–80.
- Gall H, Philippe F, Domon J-M, Gillet F, Pelloux J, Rayon C. Cell Wall Metabolism in Response to Abiotic Stress. *Plants.* 2015;4(1):112–66.
- Seifert GJ, Blaukopf C. Irritable walls: the plant extracellular matrix and signaling. *Plant Physiol.* 2010;153(2):467–78.
- Zhong R, Ye ZH. Regulation of cell wall biosynthesis. *Curr Opin Plant Biol.* 2007;10(6):564–72.
- Wang Y, Chantreau M, Sibout R, Hawkins S. Plant cell wall lignification and monolignol metabolism. *Front Plant Sci.* 2013;4:220.
- Barrière Y, Thomas J, Denoue D. QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 × F286. *Plant Sci.* 2008;175(4):585–95.
- Goering HK, Van Soest PJ. Forage Fiber Analysis (apparatus, reagents, procedures and some applications). Washington: ARS-USDA; 1970. Agric[J]
- Krakowsky M, Lee M, Beeghly H, Coors J. Characterization of quantitative trait loci affecting fiber and lignin in maize (*Zea mays* L.). *Maydica.* 2003;48(4):283–92.
- Thomas J, Guillaumie S, Verdu C, Denoue D, Pichon M, Barrière Y. Cell wall phenylpropanoid-related gene expression in early maize recombinant inbred lines differing in parental alleles at a major lignin QTL position. *Mol Breeding.* 2010;25(1):105–24.
- Truntzler M, Barriere Y, Sawkins MC, Lespinasse D, Betran J, Charcosset A, Moreau L. Meta-analysis of QTL involved in silage quality of maize and comparison with the position of candidate genes. *Theor Appl Genet.* 2010;121(8):1465–82.
- Jess DR. Relationships among Soluble Phenolics, Insoluble Proanthocyanidins and Fiber in East African Browse Species. *J Range Manage.* 1986;39(1):5–7.
- Barrière Y, Guillet C, Goffner D, Pichon M. Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. *Anim Res.* 2003;52(3):193–228.
- Casler M, Vogel K. Accomplishments and impact from breeding for increased forage nutritional value. *Crop Sci.* 1999;39(1):12–20.
- Sattler SE, Funnell-Harris DL, Pedersen JF. Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci.* 2010;178(3):229–38.
- Mechin V, Laluc A, Legee F, Cezard L, Denoue D, Barriere Y, Lapierre C. Impact of the brown-midrib bm5 mutation on maize lignins. *J Agric Food Chem.* 2014;62(22):5102–7.
- Halpin C, Holt K, Chojeccki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA. Brown-midrib maize (bm1)—a mutation affecting the cinnamyl alcohol dehydrogenase gene. *Plant J.* 1998;14(5):545–53.
- Barrière Y, Chavigneau H, Delaunay S, Courtial A, Bosio M, Lassagne H, Deroy J, Lapierre C, Méchin V, Tatout C. Different mutations in the ZmCAD2 gene underlie the maize brown-midrib1 (bm1) phenotype with similar effects on lignin characteristics and have potential interest for bioenergy production. *Maydica.* 2013;58(1):06–20.
- Tang HM, Liu S, Hill-Skinner S, Wu W, Reed D, Yeh CT, Nettleton D, Schnable PS. The maize brown midrib2 (bm2) gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. *Plant J.* 2014;77(3):380–92.
- Vignols F, Rigau J, Torres MA, Capellades M, Puigdomènech P. The brown midrib3 (bm3) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. *Plant Cell.* 1995;7(4):407–16.
- Li L, Hill-Skinner S, Liu S, Beuchle D, Tang HM, Yeh CT, Nettleton D, Schnable PS. The maize brown midrib4 (bm4) gene encodes a functional folylpolyglutamate synthase. *Plant J.* 2015;81(3):493–504.
- Penning BW, Sykes RW, Babcock NC, Dugard CK, Held MA, Klimek JF, Shreve JT, Fowler M, Ziebell A, Davis MF, et al. Genetic Determinants for Enzymatic Digestion of Lignocellulosic Biomass Are Independent of Those for Lignin Abundance in a Maize Recombinant Inbred Population. *Plant Physiol.* 2014;165(4):1475–87.
- Marriott PE, Sibout R, Lapierre C, Fangel JU, Willats WG, Hofte H, Gomez LD, McQueen-Mason SJ. Range of cell-wall alterations enhance saccharification in *Brachypodium distachyon* mutants. *Proc Natl Acad Sci U S A.* 2014;111(40):14601–6.
- Guzzo de Carli Poelking V, Giordano A, Ricci-Silva ME, Rhys Williams TC, Alves Pecanha D, Contin Ventrella M, Rencoret J, Ralph J, Pereira Barbosa MH, Loureiro M. Analysis of a Modern Hybrid and an Ancient Sugarcane Implicates a Complex Interplay of Factors in Affecting Recalcitrance to Cellulosic Ethanol Production. *PLoS One.* 2015;10(8):e0134964.
- Cass CL, Lavell AA, Santoro N, Foster CE, Karlen SD, Smith RA, Ralph J, Garvin DF, Sedbrook JC. Cell Wall Composition and Biomass Recalcitrance Differences Within a Genotypically Diverse Set of *Brachypodium distachyon* Inbred Lines. *Front Plant Sci.* 2016;7:708.
- Barrière Y, Gibelin C, Argillier O, Mechin V. Genetic analysis in recombinant inbred lines of early dent forage maize. I: QTL mapping for yield, earliness, starch and crude protein contents from per se value and top cross experiments. *Maydica.* 2001;46(4):253–66.
- Barrière Y, Méchin V, Denoue D, Bauland C, Laborde J. QTL for yield, earliness, and cell wall quality traits in topcross experiments of the F838 × F286 early maize RIL progeny. *Crop Sci.* 2010;50(5):1761–72.
- Barrière Y, Mechin V, Lefevre B, Maltese S. QTLs for agronomic and cell wall traits in a maize RIL progeny derived from a cross between an old Minnesota13 line and a modern Iodent line. *Theor Appl Genet.* 2012;125(3):531–49.
- Barrière Y, Riboulet C, Méchin V, Maltese S, Pichon M, Cardinal A, Lapierre C, Lübberstedt T, Martinant J-P. Genetics and genomics of lignification in grass cell walls based on maize as model species. *Genes Genomes Genomics.* 2007;1:133–56.
- Courtial A, Méchin V, Reymond M, Grima-Pettenati J, Barrière Y. Colocalizations Between Several QTLs for Cell Wall Degradability and Composition in the F288 × F271 Early Maize RIL Progeny Raise the Question of the Nature of the Possible Underlying Determinants and Breeding Targets for Biofuel Capacity. *BioEnergy Res.* 2013;7(1):142–56.
- Fontaine A, Briand M, Barriere Y. Genetic variation and QTL mapping of para-coumaric and ferulic acid. *Maydica.* 2003;48:75–84.
- Krakowsky MD, Lee M, Coors JG. Quantitative trait loci for cell-wall components in recombinant inbred lines of maize (*Zea mays* L.): stalk tissue. *Theor Appl Genet.* 2005;111(2):337–46.
- Krakowsky MD, Lee M, Coors JG. Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.): leaf sheath tissue. *Theor Appl Genet.* 2006;112(4):717–26.
- Lübberstedt T, Melchinger AE, Fäh S, Klein D, Dally A, Westhoff P. QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits. *Crop Sci.* 1998;38(5):1278–89.
- Lübberstedt T, Melchinger AE, Klein D, Degenhardt H, Paul C. QTL mapping in testcrosses of European flint lines of maize: II. Comparison of different testers for forage quality traits. *Crop Sci.* 1997;37(6):1913–22.
- Lübberstedt T, Melchinger AE, Schön CC, Utz HF, Klein D. QTL mapping in testcrosses of European flint lines of maize: I. Comparison of different testers for forage yield traits. *Crop Sci.* 1997;37(3):921–31.
- Méchin V, Argillier O, Hébert Y, Guingo E, Moreau L, Charcosset A, Barrière Y. Genetic analysis and QTL mapping of cell wall digestibility and lignification in silage maize. *Crop Sci.* 2001;41(3):690–7.

37. Papst C, Melchinger A, Eder J, Schulz B, Klein D, Bohn M. QTL mapping for resistance to European corn borer (*Ostrinia nubilalis* Hb.) in early maturing European dent maize (*Zea mays* L.) germplasm and comparison of genomic regions for resistance across two populations of F3 families. *Maydica*. 2001;46(3):195–205.
38. Riboulet C, Fabre F, Dénoue D, Martinant J, Lefevre B, Barrière Y. QTL mapping and candidate gene research from lignin content and cell wall digestibility in a top-cross of a flint maize recombinant inbred line progeny harvested at silage stage. *Maydica*. 2008;53(1):1.
39. Roussel V, Gibelin C, Fontaine A, Barriere Y. Genetic analysis in recombinant inbred lines of early dent forage maize. II. QTL mapping for cell wall constituents and cell wall digestibility from per se value and top cross experiments. *Maydica*. 2002;47(1):9–20.
40. Yu J, Buckler ES. Genetic association mapping and genome organization of maize. *Curr Opin Biotechnol*. 2006;17(2):155–60.
41. Yan J, Warburton M, Crouch J. Association Mapping for Enhancing Maize (*Zea mays* L.) Genetic Improvement. *Crop Sci*. 2011;51(2):433.
42. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet*. 2013;45(1):43–50.
43. Albanell E, Plaixats J, Ferret A, Bosch L, Casañas F. Evaluation of near-infrared reflectance spectroscopy for predicting stover quality trait in semi-exotic populations of maize. *J Sci Food Agr*. 1995;69(3):269–73.
44. Lambert RJ, Alexander DE, Mollring EL, Wiggins B. Selection for increased oil concentration in maize [*Zea mays*] kernels and associated changes in several kernel traits. *Maydica*. 1997;42(1):39–43.
45. Shenk JS, Westerhaus MO. Populations structuring of near infrared spectra and modified partial least squares regression. *Crop Sci*. 1991;31(6):1548–55.
46. Fu J, Cheng Y, Linghu J, Yang X, Kang L, Zhang Z, Zhang J, He C, Du X, Peng Z, et al. RNA sequencing reveals the complex regulatory network in the maize kernel. *Nat Commun*. 2013;4:2832.
47. Knapp S, Stroup W, Ross W. Exact confidence intervals for heritability on a progeny mean basis. *Crop Sci*. 1985;25(1):192–4.
48. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet*. 2006;38(2):203–8.
49. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007;23(19):2633–5.
50. Olukolu BA, Tracy WF, Wissler R, De Vries B, Balint-Kurti PJ. A Genome-Wide Association Study for Partial Resistance to Maize Common Rust. *Phytopathology*. 2016;106(7):745–51.
51. Valdar W, Holmes CC, Mott R, Flint J. Mapping in structured populations by resample model averaging. *Genetics*. 2009;182(4):1263–77.
52. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013. ISBN 3-900051-07-0; 2014.
53. Yang X, Gao S, Xu S, Zhang Z, Prasanna BM, Li L, Li J, Yan J. Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol Breeding*. 2010;28(4):511–26.
54. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA. The B73 maize genome: complexity, diversity, and dynamics. *Science*. 2009;326(5956):1112–5.
55. Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. Lignin biosynthesis and structure. *Plant Physiol*. 2010;153(3):895–905.
56. Bajsa J, Pan Z, Duke SO. Serine/threonine protein phosphatases: Multi-purpose enzymes in control of defense mechanisms. *Plant Signal Behav*. 2011;6(12):1921–5.
57. Andersen JR, Zein I, Wenzel G, Darnhofer B, Eder J, Ouzunova M, Lubberstedt T. Characterization of phenylpropanoid pathway genes within European maize (*Zea mays* L.) inbreds. *BMC Plant Biol*. 2008;8:2.
58. Andersen JR, Zein I, Wenzel G, Krutzfeldt B, Eder J, Ouzunova M, Lubberstedt T. High levels of linkage disequilibrium and associations with forage quality at a phenylalanine ammonia-lyase locus in European maize (*Zea mays* L.) inbreds. *Theor Appl Genet*. 2007;114(2):307–19.
59. Guillet-Claude C, Birolleau-Touchard C, Manicacci D, Fourmann M, Barraud S, Carret V, Martinant JP, Barriere Y. Genetic diversity associated with variation in silage corn digestibility for three O-methyltransferase genes involved in lignin biosynthesis. *Theor Appl Genet*. 2004;110(1):126–35.
60. Guillet-Claude C, Birolleau-Touchard C, Manicacci D, Rogovsky PM, Rigau J, Murigneux A, Martinant JP, Barriere Y. Nucleotide diversity of the ZmPox3 maize peroxidase gene: relationships between a MITE insertion in exon 2 and variation in forage maize digestibility. *BMC Genet*. 2004;5:19.
61. Kump KL, Bradbury PJ, Wissler RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, et al. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat Genet*. 2011;43(2):163–8.
62. Samayoa LF, Malvar RA, Olukolu BA, Holland JB, Butrón A. Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (*Sesamia nonagrioides* L.) in a maize diversity panel. *BMC Plant Biol*. 2015;15(1):35.
63. Bohn M, Schulz B, Kreps R, Klein D, Melchinger A. QTL mapping for resistance against the European corn borer (*Ostrinia nubilalis* H.) in early maturing European dent germplasm. *Theor Appl Genet*. 2000;101(5–6):907–17.
64. Cardinal AJ, Lee M, Moore KJ. Genetic mapping and analysis of quantitative trait loci affecting fiber and lignin content in maize. *Theor Appl Genet*. 2003;106(5):866–74.
65. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*. 2010;8(1):e1000294.
66. Visscher PM. Sizing up human height variation. *Nat Genet*. 2008;40(5):489–90.
67. Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C. The ArabidopsisREF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J*. 2002;30(1):33–45.
68. Ralph J, Akiyama T, Coleman HD, Mansfield SD. Effects on lignin structure of coumarate 3-hydroxylase downregulation in poplar. *Bioenergy Res*. 2012;5(4):1009–19.
69. Ralph J, Akiyama T, Kim H, Lu F, Schatz PF, Marita JM, Ralph SA, Reddy MS, Chen F, Dixon RA. Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *J Biol Chem*. 2006;281(13):8843–53.
70. Reddy MS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA. Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proc Natl Acad Sci U S A*. 2005;102(46):16573–8.
71. Fornale S, Rencoret J, Garcia-Calvo L, Capellades M, Encina A, Santiago R, Rigau J, Gutierrez A, Del Rio JC, Caparros-Ruiz D. Cell wall modifications triggered by the down-regulation of Coumarate 3-hydroxylase-1 in maize. *Plant Sci*. 2015;236:272–82.
72. Rogers LA, Campbell MM. The genetic control of lignin deposition during plant growth and development. *New Phytol*. 2004;164(1):17–30.
73. Zhao Q, Dixon RA. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends Plant Sci*. 2011;16(4):227–33.
74. Zhong R, Lee C, Zhou J, McCarthy RL, Ye Z-H. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell*. 2008;20(10):2763–82.
75. Zhong R, Ye Z-H. Transcriptional regulation of lignin biosynthesis. *Plant Signal Behav*. 2009;4(11):1028–34.
76. Legay S, Sivadon P, Blervacq AS, Pavy N, Baghdady A, Tremblay L, Levasseur C, Ladouce N, Lapiere C, Séguin A. EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *New Phytol*. 2010;188(3):774–86.
77. Patzlaff A, Newman LJ, Dubos C, Whetten RW, Smith C, McClinnis S, Bevan MW, Sederoff RR, Campbell MM. Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol*. 2003;53(4):597–608.
78. Zhou J, Lee C, Zhong R, Ye Z-H. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell*. 2009;21(1):248–66.
79. Kawaoka A, Ebinuma H. Transcriptional control of lignin biosynthesis by tobacco LIM protein. *Phytochemistry*. 2001;57(7):1149–57.
80. Kawaoka A, Kaohien P, Yoshida K, Endo S, Yamada K, Ebinuma H. Functional analysis of tobacco LIM protein Ntlm1 involved in lignin biosynthesis. *Plant J*. 2000;22(4):289–301.
81. Ko JH, Yang SH, Park AH, Lerouxel O, Han KH. ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in Arabidopsis thaliana. *Plant J*. 2007;50(6):1035–48.
82. Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M. NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *Plant Cell*. 2007;19(1):270–80.
83. Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell*. 2005;17(11):2993–3006.

84. Zhong R, Demura T, Ye Z-H. SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell*. 2006;18(11):3158–70.
85. Zhong R, Richardson EA, Ye Z-H. Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta*. 2007;225(6):1603–11.
86. Buendgen M, Coors J, Grombacher A, Russell W. European corn borer resistance and cell wall composition of three maize populations. *Crop Sci*. 1990;30(3):505–10.
87. Vorwerk S, Somerville S, Somerville C. The role of plant cell wall polysaccharide composition in disease resistance. *Trends Plant Sci*. 2004;9(4):203–9.
88. Hüchelhoven R. Cell wall-associated mechanisms of disease resistance and susceptibility. *Annu Rev Phytopathol*. 2007;45:101–27.
89. Wang GF, Balint-Kurti P. Maize Homologs of CCoAOMT and HCT, Two Key Enzymes in Lignin Biosynthesis, Form Complexes with the NLR Rp1 Protein to Modulate the Defense Response. *Plant Physiol*. 2016. pp. 00224.02016.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

