Mining the global diversity for bioenergy traits of barley straw: genomewide association study under varying plant water status

ALI A. NAZ1, STEPHAN REINERT1, CIHAN BOSTANCI2, BAHARE SIEPERI2, JENS LEON1, CHRISTIAN BÖTTGER3, KARL-HEINZ SÜDEKUM3 and MICHAEL FREI2

1Plant Breeding, Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Bonn, Germany, 2Abiotic Stress Tolerance in Crops, INRES, University of Bonn, Bonn, Germany, 3Animal Nutrition, Institute of Animal Science (ITW), University of Bonn, Bonn, Germany

Abstract

Cereal straws constitute a considerable source of biomass that can be used for bioenergy applications. Its composition is crucial for the energy value in biological or thermochemical conversion processes. Therefore, this study aimed at (i) exploring the global diversity in the composition of barley (Hordeum vulgare L.) straw; (ii) testing the effect of drought on straw composition; (iii) correlating compositional traits with energy value; and (iv) identifying loci associated with straw composition through genomewide association study (GWAS). A population of 179 barley accessions was grown in control and drought conditions, and straw was analyzed for thioglycolic acid lignin (TGAL), total phenolics (TP), carbon, crude protein (CP), C/N ratio, and ash. Substantial variability was observed in all traits. Moreover, drought treatment affected all traits leading to significant decreases in carbon, CP, ash, TGAL and TP concentrations, and a significant increase in C/N ratio. In vitro incubations in rumen fluid were used to estimate the energy value in biological energy conversion, while calorimetry was used to estimate the energy yield in thermochemical energy conversion. Thioglycolic acid lignin was singled out as the most influential trait determining energy value, as it was negatively correlated with the digestibility of organic matter and metabolizable energy in in vitro incubations, but positively correlated with gross energy measured by calorimetry. The GWAS yielded four loci significantly associated with TGAL irrespective of plant water status, which explained between 22.5% and 38.7% of the phenotypic variation. In addition, three loci significantly affected the response of TGAL to plant water status, and explained between 11.2% and 16.6% of the phenotypic variation. These loci contained plausible candidate genes that could be associated with lignin biosynthesis based on their annotations. In conclusion, this study illustrated great potential for the molecular breeding of barley varieties with enhanced straw quality for bioenergy applications.

Keywords: by-products, cell wall, cereals, global change, lignocellulose, renewable energy

Received 6 October 2016; revised version received 14 December 2016 and accepted 5 January 2017

Introduction

Every year, around 2 × 1011 tons of biomass are produced globally as agricultural by-products, including straw, roots, husks, shells, and bagasse (Tuck et al., 2012). Cereal straws alone account for an estimated 2.5 to 3 × 109 tons every year (Smil, 1999; Lal, 2005; Sun, 2010). This vast amount of biomass constitutes an enormous resource for the bioeconomy with various potential applications. It can be incorporated into soil as an organic amendment to improve soil structure and fertility (D’hose et al., 2016). Traditionally, cereal straws have always been used as a feed for ruminant herbivores, which are able to derive energy from cellulose and hemicellulose by fermenting them with the aid of their rumen microbial population (Sawatdeenarunat et al., 2015). Similar microbial processes are exploited in the biological conversion of straw into bioenergy in the form of biogas or ethanol (Ghaffar et al., 2015). Alternatively, biomass can be converted into energy via thermochemical conversion processes, which include direct combustion in small-scale applications (Saidur et al., 2011; Mendu et al., 2012), and industrial processing by gasification or pyrolysis to produce liquid fuels (Orts et al., 2008; Tanger et al., 2013). And lastly, biomass constitutes a source of raw materials for the industry, which can be processed using innovative biorefinery concepts (Sawatdeenarunat et al., 2015; Liguori & Faraco, 2016).
In all of these potential uses, the biomass composition is crucial for the efficiency of straw utilization, although different uses require different composition. Cereal straws consist largely of the lignocellulose complex (cellulose, hemicellulose, and lignin), and to a lesser extent of crude protein (CP, nitrogen × 6.25), ash, nonstructural carbohydrates, and secondary metabolites. Most importantly, lignin quantity and quality plays an outstanding role in determining straw quality for various applications (Frei, 2013). As a complex polymer composed of phenylpropanoid units, it is quite recalcitrant to biological decomposition (Tuomela et al., 2000), but is rich in gross energy due to its strongly reduced chemical nature (Petrus & Noordermeer, 2006; Hodgson et al., 2010; Frei, 2013). Lignin is therefore considered as a negative factor in all biological energy conversion processes, including rumen fermentation, biogas, and ethanol production, but it is conducive to thermochemical energy conversion processes, or relatively stable carbon sequestration in soil (Frei, 2013). Another cell wall-related trait that affects biomass degradability is the cross-linking of ferulates (Torres et al., 2015). Moreover, CP can affect biological decomposition of straw, because a minimum amount of nitrogen compounds is required to sustain microbial activity in diverse habitats such as the rumen (Russell et al., 1992) or soil (Henrikson & Breland, 1999). In addition, protein has higher energy density (around 23 MJ kg⁻¹) than cellulose and nonstructural carbohydrates (15–17 MJ kg⁻¹) and thus positively affects the gross energy content of biomass (FAO & WHO, 1985). In contrast, ash (i.e., the mineral fraction of biomass) does not contain any energy at all and should thus negatively affect energy yield on a total mass basis (Tanger et al., 2015). As relatively abundant secondary metabolites, phenolics are another factor affecting the biological decomposition of biomass, because they can inhibit both host and microbial enzyme activity, as shown, for example, in ruminant diets (Getachew et al., 2002) or in the saccharification process of biomass (Kandil et al., 2012). All of these constituents should thus be considered in optimizing straw quality for specific applications.

The composition of biomass depends on multiple environmental factors such as climate, nutrient availability, and exposure to abiotic and biotic stresses (Wang & Frei, 2011; Tanger et al., 2015). Lignin helps plants to defend themselves against various biotic and abiotic stresses and is therefore responsive to environmental stimuli. Abiotic stresses, such as tropospheric ozone or high UV radiation, as well as biotic stresses such as herbivory tend to increase the lignin concentration in agricultural crops (Frei, 2013). In contrast, drought stress decreased lignin concentrations in a number of studies with forage crops such as alfalfa (Wang & Frei, 2011), indicating contrasting responses to different stresses. Similarly, CP concentration is responsive to environmental factors, especially the nitrogen (N) availability. When plants grow in stress conditions without N limitation, they often have higher CP concentrations because less biomass is produced at a given N supply, leading to enhanced CP concentration (Wang & Frei, 2011). Ash constitutes the mineral fraction of biomass, and depends on mineral availability in the soil, but also on water supply, as minerals are taken up into plants and translocated to shoots via the transpiration stream (He & Dijkstra, 2014). And lastly, many phenolic compounds function as stress defense compounds in plants due to their antioxidant activity (Blokhina et al., 2003) and are therefore responsive to stress stimuli (Wang & Frei, 2011).

Apart from environmental factors, straw composition depends on the interspecies and intraspecies genetic variability. Substantial differences in the degree of lignification were reported between different annual crop species, ranging from an estimated 1 to 15 percent (Frei, 2013). In addition, considerable variation also exists within cereal crop species, such as rice (Jahn et al., 2011; Tanger et al., 2015) or maize (Riboulet et al., 2008b). Such intraspecies variation provides the prerequisite for marker-assisted breeding targeting bioenergy-related traits. A number of previous studies have reported quantitative trait loci (QTL) associated with straw composition traits in cereal crops such as maize (Barrière et al., 2008, 2012, 2015; Riboulet et al., 2008a), barley (Grando et al., 2005; Siahsar et al., 2009), sorghum (Murray et al., 2008; Shiringani & Friedt, 2011), and rice (Bao et al., 2007). All of these previous studies have employed biparental mapping populations. This approach has repeatedly proven to have great agronomic potential, but it covers limited genetic variability, and the mapping resolution is restricted by the number of chromosomal recombinations in a biparental cross (Murrell et al., 2012). These limitations are overcome in genome-wide association studies (GWAS), in which diverse populations of unrelated individuals are screened and QTL are determined based on linkage disequilibrium using high-density marker maps (Huang & Han, 2014). No previous GWAS has comprehensively investigated biomass compositional or bioenergy traits in cereal straws.

This study aimed at exploring the variability in bioenergy traits in a global selection of barley accessions, which were grown at different soil moisture levels. More specifically, the following hypotheses were tested: (i) Straw composition and bioenergy traits in barley are affected by both the environment (water supply) and genotype. (ii) Straw composition affects the energy yield in different conversion processes. We used an in vitro
incubation system mimicking rumen fermentation as an example for a biological energy conversion process, and bomb calorimetry as a proxy for energy yield in thermo-chemical processes. (iii) The variability in straw composition can be genetically dissected by GWAS and thus be harnessed for the breeding of novel barley varieties optimized for specific straw applications.

Materials and methods

Plant material

The barley (Hordeum vulgare L.) population used in this experiment has previously been described (Reinert et al., 2016). It consisted of 179 different accessions collected from 38 countries across the globe and thus represented substantial genetic diversity. It included 48 Hordeum vulgare ssp. spontaneum (wild) accessions and 131 Hordeum vulgare ssp. vulgare (cultivar) accessions, the latter including 72 landraces and 59 modern cultivars. The seeds had originally been provided by the Leibniz Institute for Plant Genetic and Crop Science (IPK, Gatersleben, Germany), Nordgen (NGB, Alnarp, Sweden), and ICARDA (Beirut, Lebanon).

Experiment

A phenotyping experiment using the above panel was conducted in the year 2014 in a foil tunnel located at the University of Bonn, Germany, with natural climate conditions except for precipitation. Individual accessions were replicated four times and arranged in a split-plot design with two treatments (control and drought) with subplots. The subplots were separated in lines in which they were arranged randomly. Seeds of individual accessions were sown in plastic pots (19.5 cm × 25.5 cm) containing a mixture of topsoil (40%) and natural sand (60%) (Cordel & Sohn, Salm, Germany). A drip water irrigation system (Netafilm, Adelaide, Australia) was installed, and water was supplied three times a day. To determine the volumetric moisture content (VMC), a soil moisture sensor equipped with a data logger was used. At plant development stage BBCH31–34, which refers to the bolting stage from the first to fourth detectable node on the main stem (Brockerhoff, 2015), the water supply was reduced in the drought treatment until reaching a VMC of 5% within 2 weeks, and maintained at this level for another two weeks. Control plants were irrigated without interruption. Thereafter, shoots were harvested, dried at 70 °C for 72 h, and stored until analysis.

Phenotyping

For biochemical analysis, the four replicate plants of each genotype and treatment were pooled, leading to 358 samples (179 for control and drought treatment, respectively). These pooled samples were first ground to pass a 1-mm mesh, and representative subsamples were pulverized using a bead mill. All subsequent analyses were carried out at least in technical duplicate on dry matter (DM) basis.

Carbon (C) and nitrogen (N) analyses were conducted using a CN analyzer (Eurovector EA 3000, HEKAtech, Wegberg, Germany) using 5–5.5 mg of pulverized sample material. The N values were converted into crude protein (CP) by multiplying the values with the factor 6.25.

Ash was determined by burning 0.5 g of sample material at 550 °C for 6 h in porcelain crucibles and determining the difference in weight (VDLUF, 1976).

Thioglycolic acid lignin (TGAL) was determined using a procedure described by Suzuki et al. (2009). In brief, 20 mg of dried and pulverized samples was extracted with water and methanol to obtain crude cell wall fractions. Then, the samples were digested with thioglycolic acid to obtain the TGAL complexes. Finally, the TGAL was dissolved in 1 M NaOH and the absorbance was read at 280 nm using a microplate reader (Powerwave XSII, BioTec, Bad Reichenhall, Germany). Because lignin is a very heterogeneous polymer, different quantification methods represent different complementary lignin fractions and cannot be quantitatively compared (Hatfield & Fukushima, 2005). We therefore refrained from quantification using a lignin standard curve, and instead expressed lignin values as absorbance values (mOD g⁻¹) as previously employed in TGAL assays (Bonawitz et al., 2014).

Total phenolics (TP) concentration was determined using a microplate-adjusted Folin–Ciocalteu assay (Ainsworth & Gillespie, 2007). In brief, 20 mg of dried and pulverized sample material was extracted with 70% methanol. The extracts were incubated for 40 min with Folin–Ciocalteu reagent and sodium carbonate, and the extinction was read at 725 nm using a microplate reader. A standard curve was made using gallic acid as a reference.

For gross energy analyses and in vitro incubation assays, sixteen accessions were preselected to represent contrasting values for the compositional traits determined earlier, leading to a total of 32 samples in both treatments.

In vitro gas production was determined according to Menke & Steingass (1988). Rumen fluid was obtained from two mature cannulated sheep prior to morning feeding. The sheep received a diet of grass hay, compound feed, and barley (65 : 20 : 15) corresponding to maintenance energy requirements. Rumen fluid was strained through two layers of linen cloth into a pre-warmed, insulated bottle. Samples (200–250 mg) ground to pass a 1-mm mesh were accurately weighed into 100-mL glass syringes, and a piston greased with vaseline was pushed into the cylinder. Syringes were preheated (39 °C) and filled with 30 mL of medium consisting of 10 mL of rumen fluid and 20 mL of buffer solution as described by Menke & Steingass (1988). Three blanks containing 30 mL of medium only were included. The exact volumes were read, and the syringes were placed in the rotor inside a preheated (39 °C) incubator. Gas production was measured after 24 h. Gas production of blanks was subtracted from sample gas production. Additionally, a hay standard and a concentrate standard were incubated along with each batch to correct for variations in gas production. All samples were analyzed in six analytical replicates, with three subreplicates being incubated on separate days, that is, with different rumen fluid.

The gas produced on incubation of 200 mg feed dry matter after 24 h of incubation (GP) together with the levels of other

© 2017 The Authors. Global Change Biology Bioenergy Published by John Wiley & Sons Ltd, 9, 1356–1369
chemical constituents were used to predict digestibility of organic matter determined in vivo (DO, %) and metabolizable energy (ME, MJ kg$^{-1}$ DM) according to the following formulas (Menke & Steingass, 1988):

$$DO = 16.49 + 0.9042GP + 0.0492CP + 0.0387CA$$

and

$$ME = -0.07 + 0.1478DO - 0.0062CA$$

where CP is the crude protein concentration (g N kg$^{-1}$ DM) *6.25), CA is the crude ash concentration (g kg$^{-1}$ DM), and GP, DO, and ME are as described above.

Gross energy (GE) was determined using an adiabatic bomb calorimeter (model C 200; IKA, Heitersheim, Germany).

**Statistical analyses**

Statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA). Correlations were determined using PROC CORR, principal component analysis was conducted using PROC FACTOR, and treatments (control vs. drought stress) were compared by *t*-test.

**Genomewide association study**

For determination of phenotype-genotype associations, we used the SNP marker, population structure, and kinship matrix data reported by Reinert et al. (2016). The population structure and kinship matrix were calculated using 5892 polymorphic SNP in the statistical software R. The SNP markers were selected based on minor alleles frequency >0.05 and a SNP call rate <0.95. Genomewide association mapping was performed following the GRAMMAR method described by Aulchenko et al. (2007), where the population structure was represented by the first principal components and the kinship matrix was included in the marker by trait analysis. The marker main effect and marker by treatment interaction effects were calculated simultaneously using a linear mixed model in PROC MIXED procedure in SAS program (SAS version 9.2). In this analysis, we first obtained the residual by including the principal components and kinship matrix and calculated the QTL using the residuals as new trait values in a linear mixed model as presented below:

$$Y_{ijk} = \mu + M_i + T_j + M_i * T_j + L_k(M_j) + e_{ijk}$$

where $Y_{ijk}$ is the phenotypic value; $\mu$ is the general mean; $M_i$ is the fixed effect of *i*-th marker genotype/haplotype; $T_j$ is the random effect of *j*-th treatment; $M_i * T_j$ is the interaction effect of *i*-th marker with *j*-th treatment; $L_k$ ($M_j$) is the random effect of *k*-th barley line nested within *i*-th marker genotype/haplotype, and $e_{ijk}$ is the residual. To determine traits of interest in the genomewide detection analysis, a log-of-odds (LOD) threshold with P-value ≤0.0001 and 1000 permutations was determined. The QTL model comprises an iterative multilocus procedure. Therefore, the most informative SNP (QTL) was set as a fixed factor during each calculation iteration step. All remaining markers were again incorporated in the next iteration round and re-analyzed. The starting point of each calculation round was determined by the result of the previous iteration. P-values of significant markers were corrected using probability of false discovery rate (PFDR), implemented in the SAS procedure PROC MULTTEST according to Benjamini & Yekutieli (2005). This procedure was repeated until no marker could be detected, which led to a reduction in significant markers and thereby a reduced number of false-positive QTL. A confidence interval of 5 cm was chosen on both sides of the most significant SNP and designated as putative QTL. SNPs were combined to one joint QTL depending on their estimated (significant) P-value from the first iteration of the multilocus procedure. Therefore, the size of the genetic interval was dependent on the significance value of flanking SNPs. A ‘leave-20%-out’ cross-validation procedure was used to increase the validity of all significant SNPs.

**Determination of genes colocalizing with QTL for TGAL**

The most significant SNP markers were localized on the barley physical map using the BLASTn (Altschul et al., 1990) function of DNA sequence analysis in the Ensembl-plant database (http://plants.ensembl.org/index.html). The most significant BLAST hits revealing maximum similarity percentage (>99%) and an E-value cutoff of 1E-15 was chosen as reference point on the barley genome for each QTL region. Five putative candidate genes on the left and right borders of the most significant SNP markers associated were nominated. The gene ontology terms (molecular function and/or biological processes) of the putative candidate genes were assigned using the UniProt database (Bateman et al., 2015). The putative orthologues of colocalizing genes were searched in the model species Arabidopsis thaliana, Brachypodium distachyon, Oryza sativa, and Zea mays using the Gramene database (http://www.gramene.org/).

**Results**

**Phenotypic variation in bioenergy traits and effects of drought treatment**

Ten traits were determined which are related to the energy value of barley straw. Leaf compositional traits were determined in the whole population of 179 barley accessions (Fig. 1a–f), while gross energy and in vitro biological energy value parameters were determined in 16 accessions (Fig. 1g–j), which had been preselected to contrast in leaf compositional traits. All traits showed substantial variation in the association panel, and all except DO and ME were significantly affected by the drought treatment (Fig. 1). Specifically, the drought treatment led to reductions in C concentration, CP, ash, TGAL, TP, and GE, while the C/N ratio and gas production were increased. DO and ME were also slightly but not significantly enhanced in the drought treatment.
The lack of significant increases in these two parameters despite significant increases in gas production were due to the negative effect of drought on CP, which is considered in the calculation of DO and ME.

Next, we plotted the phenotypic values for each accession in the control vs. the corresponding values in the drought treatment and determined linear regression coefficients (Fig. 2). High regression coefficients would indicate that traits are largely determined by the genotype, whereas low coefficients would suggest that traits are more responsive to the environment, that is, the plant water status. For all traits except carbon concentration and GE, R² values indicated significant correlations between values in the control and drought treatment (Fig. 2). The highest R² values were seen in TGAL, gas production, DO, and ME.

To explore interrelations between straw composition and bioenergy traits, we conducted PCA and determined pairwise Pearson’s correlation coefficients (Fig. 3). Some of these variables (DO, ME, gas production, CP, and ash) are autocorrelated due to the calculation procedures described earlier. Nevertheless, correlations between CP and ash with the digestibility parameters DO and ME were not significant, indicating that these were mostly influenced by gas production, which in turn was dominated by other chemical constituents such as TGAL. Indeed, PCA indicated that DO, ME, and gas production formed a cluster, which had the opposite factor loading from TGAL in both principal components 1 and 2 (Fig. 3a). Conclusively, highly significant negative correlations were observed between TGAL and in vitro biological energy value traits (Fig. 3b). In contrast, two straw compositional traits, that is, CP and TGAL, and GE had similar factor loading on principal component 1 (Fig. 3a), and exhibited positive correlation coefficients (Fig. 3b), which reflects the fact that these components have relatively high GE density compared to cellulose and hemicellulose, the main components of barley straw. Ash and TP were positively correlated with each other, but did not seem to affect any of the energy-related traits (Fig. 3).

Together, these data suggest that TGAL was the most important factor negatively affecting in vitro biological energy value, while TGAL and CP affected GE positively.

Loci associated with bioenergy-related traits

Association mapping detected a total of 23 QTL associated with six different traits (Table 1). Thirteen of these QTL were main effect QTL (i.e., independent of the drought treatment), while ten were interaction QTL (i.e., inducible by drought treatment). The most robust and the highest number of QTL were identified for TGAL, a trait which was found to be essential for bioenergy value in the previous section. Four main effect QTL for TGAL explained between 22.45 and 38.7 percent of the phenotypic variation, and had LOD scores ranging from 21.7 to 49.5. In addition, three interaction QTL explained between 11.2 and 16.6 percent of the phenotypic variation and had LOD scores of 10.8–17.2. In three of the main effect QTL, the minor allele was associated with lower TGAL values, indicating scope for lowering lignin concentrations in the majority of barley accessions by marker-assisted breeding. As for the interaction QTL, we plotted TGAL values for the two marker classes in both treatments (Fig. 4). With QTGAL.4H.a, the A allele but not the G allele was responsive to drought (Fig. 4a). In QTGAL.4H.b, the marker classes (C and A) differed in the drought treatment but not in the control (Fig. 4b). The opposite was seen in QTGAL.6H, where marker classes differed in the control but not in the drought treatment (Fig. 4c). Because we considered TGAL as the most influential trait determining the bioenergy value of barley straw, we curated a list of candidate genes located in each of the QTL regions (Tables S1–S7).

CP was identified as another important trait determining GE value of barley straw. However, fewer QTL were determined for this trait. One main effect QTL (QCP3H.a) explained 17.8 percent of the phenotypic variation, had an LOD score of 17.8, and the minor allele exhibited higher CP values. In addition, one interaction QTL explained 8.2 percent of the variation in CP, and had an LOD score of 7.5 (Table 1). The significant interaction occurred, because the marker classes at QCP7H.b differed significantly in the control but not in the drought treatment (Fig. 5). These interaction QTL demonstrate that plant water status needs to be considered in the marker-assisted breeding targeting TGAL and CP.

Discussion

Drought treatment and genotype affect barley straw composition

Abiotic stresses can alter the chemical composition of harvested crops and thus affect their quality in various
uses such as human and animal nutrition, or the processing into bioenergy (Wang & Frei, 2011). All of the straw compositional traits measured in this study were responsive to the drought treatment (Fig. 1a–f). While the carbon concentration was only slightly affected, CP (or N) was strongly reduced, leading to a highly significant shift in the C/N ratio. The substantial reduction in CP concentration contradicts the idea of a ‘concentration’ effect, which would imply that plants producing less biomass at given N supply would exhibit higher CP concentration. Such effects were frequently observed when crops were grown under other abiotic stresses such as high tropospheric ozone concentrations (Frei et al., 2011). However, drought treatment differs from other stresses because water is a prerequisite for N mobility. More specifically, water is required

---

**Fig. 3** Principal component analysis (a) and correlation matrix (b) for straw composition and bioenergy traits. C carbon, CP crude protein TGAL thioglycolic acid lignin, TP total phenolics, GE gross energy, Gas = gas production in *in vitro* incubation assays, DO digestibility of organic matter, ME metabolizable energy. *n* = 32. (a) Factors 1 and 2 explained 40 and 27 of the phenotypic variation, respectively. (b) Shading indicates significant correlations at *P* < 0.001 (dark gray), *P* < 0.01 (medium gray), or *P* < 0.05 (light gray).

|            | TGAL | TP  | C  | CP  | C/N | Ash | Gas | DO  | ME  | GE  |
|------------|------|-----|----|-----|-----|-----|-----|-----|-----|-----|
| TGAL       | -0.15|     |    |     |     |     |     |     |     |     |
| TP         | 0.15 | 0.25|    |     |     |     |     |     |     |     |
| C          | -0.06| -0.01| 0.38|     |     |     |     |     |     |     |
| CP         | -0.07| 0.02 | -0.29| -0.97|     |     |     |     |     |     |
| C/N        | -0.28| 0.48 | 0.24| 0.49| -0.49|     |     |     |     |     |
| Ash        | -0.80| -0.06| -0.21| -0.27| 0.28| -0.07|     |     |     |     |
| DO         | -0.73| -0.08| -0.09| -0.04| 0.04| 0.07| 0.82|     |     |     |
| ME         | -0.69| -0.14| -0.12| -0.10| 0.10| -0.06| 0.83| 0.99|     |     |
| GE         | 0.39 | -0.05| 0.27| 0.69| -0.70| 0.22| -0.43| -0.34| -0.36|     |

---

**Fig. 2** Linear regressions of control vs. drought treatment values for straw composition and bioenergy traits in diverse barley accessions. C carbon, CP crude protein TGAL thioglycolic acid lignin, TP total phenolics, GE gross energy, Gas = gas production in *in vitro* incubation assays, DO digestibility of organic matter, ME metabolizable energy. The value for each accession in the control was plotted against the value of the same accession in the drought treatment. Asterisks indicate statistically significant correlations at ***P* < 0.001, **P* < 0.01, NS not significant. *n* = 179 (a–f) or *n* = 16 (g–j).
| Trait          | QTL name                      | Marker        | Effect | Chromosomal position (cM) | Flanking region | LOD | Var (%) | Major/Minor | Major Het | Minor | H$_2$ (SE) |
|---------------|-------------------------------|---------------|--------|---------------------------|-----------------|-----|---------|-------------|-----------|-------|------------|
| TGAL (mOD g$^{-1}$) | QTGAL.2H                      | SCRI_RS_918   | M      | 2H (106.8)                | 104.2-111.3     | 21.7| 22.5    | T/C         | 18.0      | 23.3  | 26.0       | 0.88 (0.02) |
|               | QTGAL.5H.a                    | BOPA1_3200-242| M      | 5H (80.8)                 | 75.9-85.6       | 40.2| 34.6    | G/A         | 24.9      | 22.7  | 15.4       |
|               | QTGAL.5H.b                    | BOPA1_3696-2238| M    | 5H (128.1)                | 125.5-131.7     | 48.2| 38.7    | A/G         | 24.2      | 20.6  | 13.2       |
|               | QTGAL.7H                      | BOPA1_8582-772| M      | 7H (69.6)                 | 64.8-74.4       | 49.5| 37.6    | A/G         | 23.5      | 17.2  | 10.9       |
|               | QTGAL.4H.a                    | BOPA1_ABC14026-1-2-168 | T*M | 4H (51,4)              | 49.9-56.2       | 11.9| 11.2    | A/G         | 0.93      |       | (0.01)     |
|               | QTGAL.4H.b                    | BOPA2_12_30455| M    | 4H (60.3)                 | 55.7-64.0       | 10.8| 12.8    | C/A         |           |       |            |
|               | QTGAL.6H                      | SCRI_RS_176086| M      | 6H (53.6)                 | 49.2-57.2       | 17.2| 16.6    | G/A         |           |       |            |
| TP (mg g$^{-1}$)   | QPhen.2H                      | BOPA2_12_20027| M    | 2H (132.2)                | 127.2-137.0     | 10.8| 10.7    | G/A         | 24.9      | 24.9  | 22.7       | 0.08 (0.12) |
|               | QPhen.3H                      | BOPA2_12_30924| M    | 3H (108.4)                | 103.7-111.3     | 11.0| 15.4    | A/G         | 25.7      | 26.2  | 21.9       |
|               | QPhen.5H                      | BOPA2_12_21009| M    | 5H (157.4)                | 152.4-161.8     | 16.6| 16.4    | G/A         | 24.9      | 22.8  | 23.9       |
|               | QPhen.7H.b                    | BOPA1_8582-772| M    | 7H (69.6)                 | 64.8-74.4       | 12.5| 11.4    | A/G         | 25.5      | 25.1  | 21.7       |
|               | QPhen.7H.b                    | SCRI_RS_126143| M    | 7H (137.2)                | 132.2-141.4     | 12.9| 11.8    | A/C         | 25.9      | 25.8  | 23.5       |
| C (%)          | QC.3H                         | SCRI_RS_151680| T*M  | 3H (55.8)                 | 50.9-57.2       | 7.4 | 9.6     | T/G         |           |       | 0.05 (0.10) |
|               | QC.4H                         | SCRI_RS_123275| T*M  | 4H (54.6)                 | 50.6-59.5       | 16.8| 16.8    | C/T         |           |       |            |
|               | QC.6H                         | SCRI_RS_138188| T*M  | 6H (115.9)                | 113.2-119.1     | 11.5| 11.8    | C/T         |           |       |            |
|               | QC.7H                         | SCRI_RS_207086| T*M  | 7H (126.3)                | 121.8-129.1     | 8.6 | 9.6     | T/C         |           |       |            |
| CP (%)         | QCP.3H.a                      | SCRI_RS_13871 | M    | 3H (120.7)                | 115.9-124.5     | 17.8| 17.6    | A/C         | 11.3      | 11.3  | 14.4       | 0.24 (0.06) |
|               | QCP.7H.b                      | BOPA2_12_10887| T*M  | 3H (63.0)                 | 63.0-67.9       | 7.5 | 6.2     | G/A         |           |       | 0.00 (0.11) |
| C/N ratio      | QCN.2H                        | SCRI_RS_223119| M    | 2H (120.0)                | 115.2-124.9     | 16.8| 16.7    | G/A         | 26.2      | 24.7  | 22.5       | 0.40 (0.09) |
|               | QCN.4H.a                      | BOPA1_ABC08788-1-1-329 | M  | 4H (57.3)               | 52.3-61.8       | 19.2| 18.7    | A/G         | 26.5      | 25.5  | 22.5       |
|               | QCN.4H.b                      | SCRI_RS_235738| T*M  | 4H (57.5)                 | 54.0-61.4       | 9.4 | 10.0    | C/T         |           |       | 0.96 (0.00) |
|               | QCN.6H                        | SCRI_RS_136848| T*M  | 6H (72.2)                 | 69.3-75.5       | 12.6| 13.6    | A/G         |           |       |            |
| Ash (%)        | QCA.3H                        | BOPA1_ABC13753-1-2-167 | M   | 3H (105.3)               | 100.4-109.8     | 19.0| 20.8    | A/C         | 6.3       | 6.7   | 7.8        | 0.72 (0.04) |

M, marker main effect; T*M, marker by treatment interaction; TGAL, thioglycolic acid lignin; TP, total phenolics; C, carbon; CP, crude protein; Var(%), variation explained by each individual; QTL, Major Het and Minor indicate LS means of phenotypic values for the major allele, heterozygous, and minor allele for marker main effects, respectively; H$_2$, heritability; SE standard error.
for N mineralization in the soil, N uptake into roots via mass flow and diffusion, and translocation into shoots via the transpiration stream (Gonzalez-Dugo et al., 2010; He & Dijkstra, 2014). Previous studies have observed contrasting responses of crop plant CP concentration to drought treatment (Wang & Frei, 2011), which can be explained with differences in drought duration, genotype differences, and the plant tissue investigated (grain vs. straw). In our study, limited N mobility was dominant over a possible N concentration effect in barley exposed to water deficit for four weeks. Similar mechanisms may have been responsible for the slight decrease in average ash concentration in the drought treatment (Fig. 1d).

Another trait that was negatively affected in the drought treatment was TGAL (Fig. 1e). This is consistent with previous studies, for example, in maize leaves (Vincent et al., 2005), where reduced lignification was explained with an inhibitory effect of drought on lignin biosynthetic enzymes. On the other hand, reduced lignification contradicts the idea that phenylpropanoid metabolism and lignin synthesis are generally stimulated in response to diverse abiotic stresses (Cabane et al., 2012). In fact, with the majority of abiotic stresses, increases in lignification of plant materials are observed due to oxidative stress reactions favoring the synthesis of phenylpropanoids having antioxidant potential, and their oxidative coupling to lignin polymers in apoplastic redox reactions (Wang & Frei, 2011). But together with the slightly negative effect of drought on TP concentration (Fig. 1f), our results clearly demonstrate that inhibitory effects on the biosynthesis of phenolics and lignin dominated over positive stress stimuli in barley plants exposed to water deficit for four weeks.

The negative effects of drought on GE can be explained with the decreases in both lignin and CP, two components with relatively high GE density (FAO & WHO, 1985, Frei, 2013). On the other hand, gas production in in vitro assays was increased, presumably because drought inhibited the formation of lignin, a major limiting factor in microbial fermentation (Krause et al., 2003). That drought effects on DO and ME were
not significantly positive was due to the fact that CP concentration (Menke & Steingass, 1988) is considered in the calculation of these parameters, which partly offset the positive effects on gas production.

Apart from the environmental effects on straw composition, all measured compositional traits showed substantial genotypic variation (Fig. 1), which is the prerequisite for exploiting genetic diversity via GWAS (Huang & Han, 2014). When control values were plotted vs. the values from drought stress treatment, the highest $R^2$ was observed in lignin, indicating that genotypes ranked similarly in both treatments. This was consistent with high H2 values for TGAL in the GWAS (Table 1), and suggests high potential for lignin as a target trait in breeding irrespective of the water status. Consistently, lignin has been proposed as an important target trait for improving feed value of forage crops (Barrière et al., 2003), or energy yield from biomass crops (Zhao et al., 2014) by harnessing natural variation.

**Compositional traits are correlated with bioenergy value**

Biomass can be converted into energy using biological or thermochemical conversion processes. In our study, we employed in vitro straw degradation using rumen fluid as a proxy for the energy value in biological energy conversion processes. Originally, this method was developed to estimate the energy value of feed in ruminant diets (Menke et al., 1979; Menke & Steingass, 1988). More recently, such in vitro incubation assays have also been used to estimate bioenergy yield from crops such as maize (Lorenz et al., 2009). High correlations occur between in vitro degradation assays using rumen fluid and bioenergy yield in ethanol or biogas production, because similar microbial processes and enzymes mediate these processes, in which lignin is generally considered as an inhibitory factor (Frei, 2013; Merlin Christy et al., 2014). In addition, we used GE measurements as a proxy for the energy yield obtained in thermochemical processes such as direct combustion, pyrolysis, and gasification. In these processes, lignin is considered as a positive factor as it contains less oxygen but more gross energy (around 22–24 MJ kg$^{-1}$) than carbohydrates (15–17 MJ kg$^{-1}$) (Petrus & Noordermeer, 2006; Hodgson et al., 2010). Correlations between compositional and bioenergy traits observed in this study confirm this conflicting role of lignin in different bioenergy applications. While TGAL was strongly negatively correlated with in vitro biological energy value parameters, it was positively correlated with GE (Fig. 3). These high correlations imply that TGAL is a useful trait that can be targeted in the breeding for high bioenergy value. Due to its very complex and heterogeneous nature of lignin polymers, different analyses assays typically measure complementary lignin fractions and cannot be directly compared (Hatfield & Fukushima, 2005). Gravimetric methods such as Klason lignin and acid detergent lignin (ADL) are widespread, and their correlations with bioenergy yield are well established (Van Der Weijde et al., 2016). However, these methods are laborious and therefore less suitable for high-throughput applications required in the phenotyping for genome mapping and plant breeding, unless they are used to calibrate near infrared spectroscopy (NIRS) models. The TGAL method employed in this study is based on a spectrometric, microplate-adjusted procedure (Suzuki et al., 2009) and allows for high-throughput lignin analyses, which show high correlations with bioenergy traits.

Apart from lignin, CP showed positive correlation with GE, which was due to the fact that CP has higher energy density (around 23 MJ kg$^{-1}$) than carbohydrates (around 15–17 MJ kg$^{-1}$) (FAO & WHO, 1985, Lewandowski & Kicherer, 1997). In contrast, CP was not correlated with any of the in vitro degradation traits, which is surprising at first glance because a sufficient amount of CP (or N) is required to sustain rumen microbial growth and activity (Russell et al., 1992). This lack of correlation can be explained with the fact that the in vitro protocol employed in this study in not N-limited because sufficient N is added to the buffered rumen fluid used for incubations (Menke et al., 1979; Menke & Steingass, 1988).

**Straw compositional traits can be genetically dissected by GWAS**

Phenotypic analyses revealed substantial variation in straw compositional traits, which was genetically dissected by GWAS. The highest number of QTL was identified for TGAL (Table 1), which was found to be crucial in determining the bioenergy value in the previous sections. High heritability, high percentage of phenotypic variation explained, and high LOD score point to the robustness of these QTL. A number of previous studies have employed QTL analysis for lignin concentration in cereal straw. While a GWAS identified only relatively weak marker–trait associations for TGAL in a global population of rice (Ueda et al., 2015), several significant QTL were found in this study with barley, suggesting that this species might offer better perspectives in the breeding for enhanced straw quality. An earlier biparental mapping study with barley (Grando et al., 2005) identified eleven QTL associated with lignin (measured indirectly using infrared spectrometry), using a population derived from a cross between the H. vulgare cultivar Arta and the H. spontaneum line 41-1, which covers 890 cM of the exotic genome. In all except one
QTL, the allele from *H. spontaneum* increased lignin values, and most of these QTL were specific to certain environments, which is consistent with our observation of three QTL interacting with the plant water status (Fig. 4). Colocalization analysis of these previous QTL with the QTL determined in this study is difficult, because of the low resolution of the previous mapping conducted with 188 markers in a biparental cross. With the advent of high-throughput marker genotyping and sequencing, it became possible to establish SNP markers throughout the genome and increase the resolution of QTL analysis manifold. High-density genomewide marker maps and highly polymorphic populations now make it possible to narrow down the QTL regions instantly.

As previously shown in maize (Barrière et al., 2015), candidate genes can be selected by screening QTL regions for genes with putative functions in lignin biosynthesis, such as orthologues of known genes involved in monolignol synthesis, polymerization, MYB, or NAC transcription factors. In the present study, we employed the barley physical map to curate a list genes colocalizing with the QTL for TGAL (Tables S1–S7), which contains several plausible candidates that could be involved in lignin biosynthesis. For example, the putative multi copper oxidase MLOC_24900 located near the QTL QTGAL.5H (Tables S2) could function as an apoplastic laccase, a class of enzymes which is involved in the oxidative coupling of monolignols to forming lignin polymers (Bonawitz & Chapple, 2010). The same QTL region contained a putative NADH dehydrogenase (MLOC_25780), a class of enzymes which has been suggested to act as electron donors to cytochrome P450 enzymes involved in lignin biosynthesis (Śundin et al., 2014). For the same reason, the putative cytochrome 450 enzyme MLOC_72591 located near QTL QTGAL4H.b (Tables S6) is also a plausible candidate gene. In general, the list of candidate genes contains several transcription factors, which can play important roles in lignin biosynthesis (Zhao & Dixon, 2011; Soler et al., 2017). Among others, WRKY family transcription factors such as the gene MLOC_59246 (Tables S3) have recently been shown to regulate lignin biosynthesis (Gallego-Giraldo et al., 2016). Similarly, NAC-type transcription factors such as MLOC_65101, which contains the SNP marker underlying the interaction QTL QTGAL.4H.a (Tables S5), can play important roles in cell wall formation (Sakamoto et al., 2016). And lastly, the region of the interaction QTL QTGAL4H.b (Tables S6) contained a putative AP2-like ethylene responsive transcription factor (MLOC_72591), which could regulate lignin formation in response to drought by interacting with ethylene, a major drought-responsive plant hormone (Farooq et al., 2009). It must be noted that none of the orthologues of the suggested candidate genes (Tables S1–S7) had previously been shown to be directly involved in lignin biosynthesis in other species, whereas *a priori* candidates such as MYB transcription factors (Bonawitz & Chapple, 2010; Soler et al., 2017) were not among the genes colocalizing with the QTL determined in this study. This may be due to species differences such as functional divergences of genes that have developed during the evolutionary history, or the use of different lignin fractions in different studies (e.g., TGAL vs. ADL). Therefore, the potentially novel candidate genes suggested in this study require further characterization, for example, through functional genomics, reverse genetics, and analyses of sequence polymorphisms, to confirm their involvement in TGAL synthesis in barley.

In conclusion, this study demonstrated substantial natural variation in bioenergy traits of barley straw. TGAL was identified as a suitable target trait for marker-assisted breeding due to its high correlation with energy efficiency in both biological and thermochemical conversion processes, its high heritability, and the robust QTL associated with this trait. Further studies are recommended to investigate the candidate genes underlying these QTL. Another question that will have to be addressed is whether straw composition can be altered without compromising agronomic traits of primary interest to crop growers, such as grain yield, resistance to biotic and abiotic stresses, or lodging resistance. Also, the requirements for bioenergy need to be balanced vs. alternative straw uses such as organic amendment to increase soil organic matter and fertility. Together, these efforts can lead to more efficient utilization of a crop by-product, which is abundantly available as a source of bioenergy.

**References**

Ainsworth EA, Gillespie KM (2007) Estimation of total phenolic content and other oxidation substrates in plant tissues using Folin-Ciocalteu reagent. *Nature Protocols*, 2, 875–877.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410.

Aulchenko YS, De Koning DJ, Haley C (2007) Genomewide rapid association using sequence polymorphisms, to confirm their involvement of three QTL interacting with the plant water status (Fig. 4). Colocalization analysis of these previous QTL with the QTL determined in this study is difficult, because of the low resolution of the previous mapping conducted with 188 markers in a biparental cross. With the advent of high-throughput marker genotyping and sequencing, it became possible to establish SNP markers throughout the genome and increase the resolution of QTL analysis manifold. High-density genomewide marker maps and highly polymorphic populations now make it possible to narrow down the QTL regions instantly.

As previously shown in maize (Barrière et al., 2015), candidate genes can be selected by screening QTL regions for genes with putative functions in lignin biosynthesis, such as orthologues of known genes involved in monolignol synthesis, polymerization, MYB, or NAC transcription factors. In the present study, we employed the barley physical map to curate a list genes colocalizing with the QTL for TGAL (Tables S1–S7), which contains several plausible candidates that could be involved in lignin biosynthesis. For example, the putative multi copper oxidase MLOC_24900 located near the QTL QTGAL.5H (Tables S2) could function as an apoplastic laccase, a class of enzymes which is involved in the oxidative coupling of monolignols to forming lignin polymers (Bonawitz & Chapple, 2010). The same QTL region contained a putative NADH dehydrogenase (MLOC_25780), a class of enzymes which has been suggested to act as electron donors to cytochrome P450 enzymes involved in lignin biosynthesis (Śundin et al., 2014). For the same reason, the putative cytochrome 450 enzyme MLOC_72591 located near QTL QTGAL4H.b (Tables S6) is also a plausible candidate gene. In general, the list of candidate genes contains several transcription factors, which can play important roles in lignin biosynthesis (Zhao & Dixon, 2011; Soler et al., 2017). Among others, WRKY family transcription factors such as the gene MLOC_59246 (Tables S3) have recently been shown to regulate lignin biosynthesis (Gallego-Giraldo et al., 2016). Similarly, NAC-type transcription factors such as MLOC_65101, which contains the SNP marker underlying the interaction QTL QTGAL.4H.a (Tables S5), can play important roles in cell wall formation (Sakamoto et al., 2016). And lastly, the region of the interaction QTL QTGAL4H.b (Tables S6) contained a putative AP2-like ethylene responsive transcription factor (MLOC_72591), which could regulate lignin formation in response to drought by interacting with ethylene, a major drought-responsive plant hormone (Farooq et al., 2009). It must be noted that none of the orthologues of the suggested candidate genes (Tables S1–S7) had previously been shown to be directly involved in lignin biosynthesis in other species, whereas *a priori* candidates such as MYB transcription factors (Bonawitz & Chapple, 2010; Soler et al., 2017) were not among the genes colocalizing with the QTL determined in this study. This may be due to species differences such as functional divergences of genes that have developed during the evolutionary history, or the use of different lignin fractions in different studies (e.g., TGAL vs. ADL). Therefore, the potentially novel candidate genes suggested in this study require further characterization, for example, through functional genomics, reverse genetics, and analyses of sequence polymorphisms, to confirm their involvement in TGAL synthesis in barley.

In conclusion, this study demonstrated substantial natural variation in bioenergy traits of barley straw. TGAL was identified as a suitable target trait for marker-assisted breeding due to its high correlation with energy efficiency in both biological and thermochemical conversion processes, its high heritability, and the robust QTL associated with this trait. Further studies are recommended to investigate the candidate genes underlying these QTL. Another question that will have to be addressed is whether straw composition can be altered without compromising agronomic traits of primary interest to crop growers, such as grain yield, resistance to biotic and abiotic stresses, or lodging resistance. Also, the requirements for bioenergy need to be balanced vs. alternative straw uses such as organic amendment to increase soil organic matter and fertility. Together, these efforts can lead to more efficient utilization of a crop by-product, which is abundantly available as a source of bioenergy.

**References**

Ainsworth EA, Gillespie KM (2007) Estimation of total phenolic content and other oxidation substrates in plant tissues using Folin-Ciocalteu reagent. *Nature Protocols*, 2, 875–877.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410.

Aulchenko YS, De Koning DJ, Haley C (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigreed-based quantitative trait loci association analysis. *Genetics*, 177, 577–585.

Bao JS, Jin L, Shen Y, Xie J (2007) Genetic mapping of quantitative trait loci associated with fiber and lignin content in rice. *Cereal Research Communications*, 35, 23–30.

Barrière Y, Guillet C, Goffner D, Pichon M (2003) Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. *Animal Research*, 2, 193–228.

Barrière Y, Thomas J, Denoué D (2008) QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 x F286. *Plant Science*, 175, 585–595.

Barrière Y, Courtial A, Soler M, Grima-Pettenati J (2015) Towards the identification of genes underlying maize QTLs for lignin content, focusing on colocalizations with lignin biosynthetic genes and their regulatory MYB and NAC transcription factors. *Molecular Breeding*, 25, 87.

© 2017 The Authors. *Global Change Biology Bioenergy* Published by John Wiley & Sons Ltd, 9, 1356–1369.
Tanger P, Vega-Sanchez ME, Fleming M et al. (2015) Cell wall composition and bioenergy potential of rice straw tissues are influenced by environment, tissue type, and genotype. Bioenergy Research, 8, 1165–1182.
Torres AF, Visser RGF, Trindade LM (2015) Bioethanol from maize cell walls: genes, molecular tools, and breeding prospects. Global Change Biology Bioenergy, 7, 591–607.
Tuck CO, Perez E, Horvath IT, Sheldon RA, Poliakoff M (2012) Valorization of biomass: deriving more value from waste. Science, 337, 695–699.
Tuomela M, Vikman M, Hatakka A, Itavaara M (2000) Biodegradation of lignin in a compost environment: a review. Bioresource Technology, 72, 169–183.
Ueda Y, Frimpong F, Qi Y et al. (2015) Genetic dissection of ozone tolerance in rice (Oryza sativa L.) by a genome-wide association study. Journal of Experimental Botany, 66, 293–306.
Van Der Weijde T, Torres AF, Dolstra O, Dechesne A, Visser RGF, Trindade LM (2016) Impact of different lignin fractions on saccharification efficiency in diverse species of the bioenergy crop miscanthus. Bioenergy Research, 9, 146–156.
VDLUFA (1976) Methodenhandbuch Band III - die chemische Untersuchung von Futtermitteln. Verband deutscher landwirtschaftlicher Untersuchungs- und Forschungsanstalten, Speyer, Germany. 2190 S., ISBN 978-3-941273-14-6.
Vincent D, Lapierre C, Pollet B, Cordier G, Negroni L, Zivy M (2005) Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation. Plant Physiology, 137, 949–960.
Wang Y, Fret M (2011) Stressed food - The impact of abiotic environmental stresses on crop quality. Agriculture Ecosystems and Environment, 141, 271–286.
Zhao Q, Dixon RA (2011) Transcriptional networks for lignin biosynthesis: more complex than we thought? Trends in Plant Science, 16, 227–233.
Zhao H, Li Q, He Jr, Yu JY, Yang JP, Liu CZ, Peng JH (2014) Genotypic variation of cell wall composition and its conversion efficiency in Miscanthus sinensis, a potential biomass feedstock crop in China. Global Change Biology Bioenergy, 6, 768–776.

Supporting Information
Additional Supporting Information may be found online in the supporting information tab for this article:

Tables S1–S7 Lists of genes co-localizing with QTL that control thioglycolic acid lignin as marker main effects or marker by treatment interaction.