Association analysis of five candidate genes with plant height and dry matter yield in perennial ryegrass

Gražina Statkevičiūtė[1,4], Andrius Aleliūnas[1], Vilma Kemešytė[1], Izolda Pašakinskienė[2], Thomas Lübberstedt[3] and Gintaras Brazauskas[1]

[1]Institute of Agriculture, Lithuanian Research Centre for Agriculture and Forestry, Instituto a. 1, Akademija, LT-58344 Kėdainiai distr., Lithuania; [2]Lithuania Faculty of Natural Sciences, Vilnius University, M.K.Čiurlionio 21, LT-03101 Vilnius, Lithuania; [3]Department of Agronomy, Iowa State University, 1204 Agronomy Hall, Ames, IA 50011, USA; [4]Corresponding author, E-mail: grazinastat@lzi.lt

Abstract

Five candidate genes *LpIAA1*, *LpRUB1*, *LpBRI1*, *LpSHOOT1* and *LpTB1* with putative function in plant architecture were allele sequenced in a panel of 96 diploid perennial ryegrass genotypes of diverse origin. The total length of the non-redundant genomic DNA alignment was 5425 bp and revealed 270 polymorphic sites in total. A negative significant Tajima's D value was detected in *LpTB1* gene, suggesting selection pressure for low frequency alleles in this gene. All 96 genotypes were evaluated for plant height and dry matter

yield over two years. Marker – trait associations were calculated between polymorphic sites and both phenotypic traits. Three indels and three single nucleotide substitutions in *LpTB1* gene were significantly ($p < 0.05$, $q < 0.05$) associated with plant height, while one indel was associated with dry matter yield. The results suggest putative role of *LpTB1* gene in plant height determination in perennial ryegrass and provide means for target allele selection.

Key words: Association mapping – *Lolium perenne* – *tb1* – single nucleotide polymorphism

Introduction

Perennial ryegrass (*Lolium perenne* L.) is the most widely grown grass species in temperate regions used both for turf and forage. It has a high nutritional value, superior digestability and good grazing tolerance as well as adequate seed production (Wilkins 1991, Forster et al. 2008), which lead to it's importance as a pasture and hay grass, while rapid establishment rate and excellent tolerance to traffic make perennial ryegrass one of the most extensively used grass species for turf or amenity purposes (Wang et al. 2001).

Herbage yield is a key productivity trait in perennial ryegrass breeding for forage whereas plant height is of higher importance in breeding for turf. Several QTLs related to plant growth have been published. QTLs for plant height and plant growth rate were identified on LG7 by Pauly et al. (2012), which co-localized with QTL for fresh and dry herbage weight (Anhalt et al. 2009), plant height and leaf length (Studer et al. 2008, Barre et al. 2009, Kobayashi et al. 2011) and QTL for heading date (Armstead et al. 2004, 2008, Studer et al. 2008, Barre et al. 2009, Byrne et al. 2009). QTL studies usually focus on segregating populations, which require laborious population construction; however, only relatively low genetic resolution can be achieved due to limited number of recombination events in such populations. Association mapping has emerged as a tool to resolve complex trait variation down to the sequence level by exploiting historical and evolutionary recombination events at the population level, depending on the extent of linkage disequilibrium (LD) in the respective population. Either a candidate gene approach or a whole genome scan can be applied to identify marker-trait associations (reviewed in Zhu et al. 2008). Rapid LD decay in perennial ryegrass has been reported due to its outbreeding reproductive system (Ponting et al. 2007, Skot et al. 2007, Xing et al. 2007, Auzanneau et al. 2011, Brazauskas et al. 2011) which in combination with highly diverse germplasm resources enables application of the candidate gene

approach. LD level and nucleotide diversity have previously been studied in five perennial ryegrass genes with putative function in plant architecture. Orthologs of these genes were shown to be involved in phytohormone signalling regulation and axillary meristem outgrouth in various plant species (Brazauskas et al. 2010). The objectives of this study were to evaluate a perennial ryegrass association mapping population for variation in plant height and dry matter yield and identify associations between these traits and sequence variants within the *LpIAA1*, *LpRUB1*, *LpBRI1*, *LpSHOOT1,* and *LpTB1* genes.

Materials and methods

Plant material and phenotyping: A collection of 96 perennial ryegrass (*Lolium perenne* L.) genotypes consisted of 44 cultivars, 51 natural ecotype and one colchicine induced mutant (Pašakinskienė 2005). Twenty ecotype accessions were collected in their natural habitats in Lithuania, 24 in Northern and North-Western Ukraine, 6 in Russia (Kaliningrad region) and 1 in Latvia. The cultivar group consisted of 39 cultivars developed in Europe, 4 in the USA and 1 in Japan. Twenty two, twelve and ten cultivars were of turf, forage and intermediate (turf/forage) type, repectively. A single seed per accession was germinated. Plants were established and clonally propagated in the greenhouse. A set of 96 genotypes went through a vernalisation process in the greenhouse at

approx. 5°C from November 2012 to February 2013 (four months in total). Ramet-based field collection was established in April 2013. The second set of un-vernalised genotypes was planted in the field in July 2013 and went through vernalisation under natural conditions during the winter of 2013/2014. In both sets the same panel of genotypes was used and four ramets (replicates) of each genotype were planted at 50 × 50 cm distances using a randomised complete block design in Akademija (Kedainiai distr., Lithuania). The soil of the experimental site was Endocalcari–Epihypogleyic Cambisols (CMg-p-w-can) with pH 7.1 and 7.3, P_2O_5 143 and 175 mg kg^{-1}, K_2O 165 and 175 mg kg^{-1}, humus 1.61 and 1.93 % in 2013 and 2014, respectively. Plant height and dry matter yield (DMY) of the 1st cut were assessed in 2013 and 2014 in the first and second set, respectively. Plant height was measured at full emergence of inflorescences by stretching the plant and recording the height from the ground to the point where most of the inflorescences end. Genotypes that did not flower were discarded from further analysis. Each plant was cut at full ripening stage and dried at 105°C till constant weight for dry matter yield measurement. Pearson correlation coefficients among plant height and DMY were calculated with STATISTICA 7 (StatSoft Inc., USA). Broad sense heritability was estimated by the formula $H^2 = \sigma^2_g/(\sigma^2/re + \sigma^2_{ge}/e + \sigma^2_g)$, where r is the number

of replicates and e is the number of environments, with PLABSTAT software package (Utz 2011).

Candidate gene amplification and sequencing: Gene fragments of *LpIAA1*, *LpBRI1*, *LpRUB1,* and *LpSHOOT1* were amplified with primers described in Brazauskas et al. (2010), whereas for *LpTB1*the sequences of forward and reverse primers were 5'-TGATCTGCTCCTGCTAGTCCT-3' and 5'-TGCAGATTAGAATCCACGCAAGA-3', respectively. PCR was carried out in 10 µl reaction mixtures, containing 1 x Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, Lithuania), 0.5 µM of each forward and reverse primer and 80 ng of DNA. PCR thermal profile was 98°C for 10 sec, followed by 30 cycles of 98°C for 1 sec, annealing temperature for 5 sec, 72°C for 40 sec and final step of 72°C for 60 sec. Reaction was terminated with a 1 min extension at 72°C. The annealing temperature was set to 71°C for *LpIAA1* and *LpBRI1*, 64°C for *LpRUB1* and *LpSHOOT1* and 61°C for *LpTB1*. The cloning of amplified fragments, transformation, plasmid DNA extraction and sequencing were performed as described in Aleliūnas et al. (2014). Five clones per fragment were sequenced. Each fragment was sequenced from both directions with forward and reverse pJET 1.2 sequencing primers. Sequence chromatograms were assembled into contigs using ChromasPro v.1.7.5

(Technelysium Pty Ltd, Australia) and two alleles per genotype were obtained. Sequence alignment was done with MEGA6 software (Tamura et al. 2013).

Linkage disequilibrium and marker–trait association analysis: Linkage disequilibrium (r^2) as well as the level of genetic variation within each gene was calculated with DnaSP v5.10 (Librado and Rozas 2009). Fisher's exact test (Sokal and Rolf 1981) was used to determine the statistical significance of LD. Population structure and relative kinship estimations were based on AFLP markers as described in Aleliūnas et al. (2014). Association analysis between SNPs and phenotypic traits was performed using a mixed linear model (MLM) implemented in the Tassel v.4.3 software (Bradbury et al. 2007). SNPs with minor allele frequency of less than 5 % were excluded from further analysis. The multiple testing correction of p values was performed by positive false discovery rate (q value). q values were estimated using QVALUE package (Storey et al. 2004). The phenotypic means of marker genotype classes were compared using post-hoc Tukey HSD test.

Results

Phenotypic variation

Plant height and dry matter yield (DMY) of each plant were measured in 2013 and 2014. Four genotypes did not produce inflorescences in 2013 and were

discarded from the analysis. High variation for both phenotypic traits was observed, with average height and DMY being higher in 2014 than 2013 (Table 1). The spring of 2013 was late, therefore the first set of genotypes could be re-planted from the greenhouse to the field only in the second half of April. The autumn was warm and vegetative growth continued until the end of December, thus the plants in the second set experienced superior conditions to establish and develop roots. The different growth conditions determined the two-fold difference in average DMY – it was 155 g in 2014 but only 78 g in 2013. The same trend was observed for plant height – 71 cm in 2014 and 50 cm in 2013. The heritability of plant height H^2 was 0.80, but only 0.32 for DMY. Variable correlations for plant height between two years were obtained in the genotype groups (Fig. 1). It was strong in both turf type and intermediate type cultivar groups ($r = 0.83$ and $r = 0.85$ respectively), however low in the ecotype group ($r = 0.34$) and moderate ($r = 0.66$) in the forage type cultivar group. The correlation between height and DMY was moderate ($r = 0.50$ and $r = 0.63$ in 2013 and 2014, respectively). The height (2 years average) of the plants in the ecotype group was higher than in the cultivar groups: 64.8, 61.5, 58.8 and 49.5 cm in ecotype, forage-, intermediate- and turf-type cultivar group, respectively. The same trend was observed with the dry matter yield: 99.0, 102.9 and 84.4 g

in forage-, intermediate- and turf-type cultivars, respectively, while ecotypes yielded 136.7 g of DMY on average.

Sequence variation in candidate genes and linkage disequilibrium

Gene-specific primers were used to amplify fragments of five candidate genes, 5425 bp of genomic sequence in total. Indels and SNPs were detected in each gene. The number of indels varied from 8 in *LpIAA1* to 43 in *LpRUB1* (Table 2). Most of indels (99 out of 103) were located in non-coding sequences, one indel of 3 bp in length was identified in the coding sequence of *LpIAA1* while 3 short indels were also detected in *LpTB1* coding sequence. The longest indel of 105 bp in length was present in non-coding sequence of *LpIAA1*. Tajima's D of the indels was not significant in all five genes.

SNPs were detected in both coding and non-coding regions. The total number of segregating sites in non-coding sequence was higher than in coding (234 and 150, respectively). However, *LpBRI1* and *LpTB1* exhibited higher nucleotide diversity in coding regions. The number of haplotypes for the five genes ranged from 24 to 55 in *LpIAA1* and *LpRUB1*, respectively. Haplotype number as well as haplotype diversity index of coding sequences was higher compared to non-coding sequences, with the exception of *LpBRI1*. Tajima's D value was positive for *LpIAA1* and negative for the rest of the genes. Most Tajima's D values were nonsignificant, thus deviation from neutral expectations was not supported.

However, a negative and significant Tajima's D value was obtained for *LpTB1* indicating an excess of rare alleles.

The average level of LD for the five genes was r^2=0.49 (only pair-wise comparisons significant by Fisher's exact test were used) (Fig. 2). The lowest level of LD (r^2 = 0.35) was detected in *LpSHOOT1* gene while the highest LD (r^2 = 0.71) was present in *LpTB1* gene. However, the number of significant pair-wise comparisons in *LpTB1* was smaller compared to the rest of the genes. High levels of LD were maintained over the sequenced fragment length in all five genes (Supp. Fig. 1).

Marker – trait associations

Six phenotypic traits were used in the analysis: plant height in 2013, in 2014 and the average over both years, DMY in 2013, in 2014 and it's average over both years. The total number of polymorphic sites was 270: 137 in *LpIAA1*, 67 in *LpBRI1*, 55 in *LpRUB1*, 91 in *LpSHOOT1* and 20 in *LpTB1*. Three indels, two in the 5'-UTR and one in the 3'-UTR, and three single nucleotide substitutions in the coding region of *LpTB1* gene were significantly associated with plant height, while one indel in the 3'-UTR of *LpTB1* was associated with dry matter yield (Table 3). One of the three associated SNPs caused an amino acid change from glycine to cysteine, the other two were synonymous. ANOVA analysis revealed that the average height in 2014 of plants harboring

predominant alleles at the three associated SNP positions was significantly lower ($p < 0.05$) compared to the plants homozygous for variant alleles as well as heterozygous plants ($p = 0.053$) (Fig. 3). Fifteen significant associations ($p < 0.05$) were detected within *LpIAA1*, with 3 SNPs being associated with dry matter yield and 12 with plant height. Sixteen SNPs were associated with dry matter yield in *LpSHOOT1*. However, *q* values for these associations in *LpIAA1* and *LpSHOOT1* exceeded the threshold of 0.05. No significant marker–trait associations were detected for *LpBRI1* and *LpRUB1*.

Discussion

Heritability of phenotypic traits

The association mapping panel used in this study consisted of highly diverse perennial ryegrass genotypes. Half of the cultivar genotypes (22 out of 44) were of turf type, 12 were attributed to forage type and 10 cultivars were described as suitable for both purposes. Perennial ryegrass ecotypes usually exhibit large variation in phenotypic traits (Statkevičiūtė et al. 2012) as was observed for plant height and dry matter yield in this study. Plant height usually shows high heritability (Alves de Araujo and Coulman 2004, Yamada et al. 2004, Shortell et al. 2009), which was also the case in this study. The dry matter yield

heritability, on the contrary, was low as this trait is significantly influenced by the environmental factors (Conaghan et al. 2008).

High levels of LD detected

An important factor to be considered in association studies is the level of LD present in the panel of genotypes under study. With high LD values, marker-trait associations can be revealed with low number of markers, whereas with low LD the power of single marker will only extend over a short distance (Salvi and Tuberosa 2005). Rapid LD decay is expected for perennial ryegrass due to an obligate outcrossing nature. However, the level of LD varies across the genome and the estimates may also differ among studies due to different sets of genotypes used. A very rapid LD decay with r^2 values reaching 0.1 within distance of 150 bp was reported in *L. perenne* gibberellic acid insensitive gene in the set of 216 plants of synthetic variety 'Herbie'(Auzanneau et al. 2011) whereas almost complete LD was observed within floral control gene *LpSFT* in a panel of 20 unrelated perennial ryegrass genotypes (Fiil et al. 2011). The level of LD in *LpIAA1*, *LpRUB1*, *LpBRI1*, *LpSHOOT1* and *LpTB1* genes had been investigated in the same panel of 20 genotypes (Brazauskas et al. 2010). Average LD values had reached 0.2 within 263 bp and 955 bp in *LpRUB1* and *LpIAA1*, respectively, while high LD was maintained over the entire alignment length for *LpBRI1*, *LpSHOOT1,* and *LpTB1*. In the present study, high levels of

average LD persisted throughout the entire sequence of all five genes. Underlying population structure could be a major cause for slow LD decay, however the same level of LD decay was detected in both subpopulations. Most likely the studied genes are under selection pressure as was demonstrated for *LpTB1* in this study and higher levels of LD are maintained throughout their length.

Sequence polymorphism indicates selection in *LpTB1*

The average nucleotide diversity (π) was only slightly lower compared to the previous study (Brazauskas et al. 2010) with values of 0.009 and 0.0095, respectively, despite the fact that 5-fold higher number of genotypes was used in the current study. Both studies revealed highest and lowest π values in *LpSHOOT1* and *LpTB1* genes, respectively. The drop in diversity is substantially greater at genes involved in domestication (Buckler et al. 2001). The domestication process of ryegrasses started very recently in comparion to major food crops, therefore high levels of diversity can be expected in cultivars and wild ecotypes alike. Indeed, *LpTB1* π values did not differ between cultivar and ecotype groups, however higher haplotype diversity was detected among ecotypes (ecotype Hd = 0.818, cultivar Hd = 0.717). This difference in Hd values between genotype groups was not observed in the remaining four candidate genes. Furthermore, the significant negative Tajima's D values for

both coding and non-coding regions of the *LpTB1* suggest selection for low frequency alleles in this gene.

Marker – trait associations in *LpTB1*

The *teosinte branched1* is one of the best studied genes affecting plant architecture . It has been demonstrated that *TB1* expression was about two-fold higher in maize than in its wild ancestor teosinte, and this increased expression led to suppression of auxiliary bud outgrowth (Doebley et al. 1997, Wang et al. 1999, Clark et al. 2004). The lack of any fixed amino acid differences between maize and teosinte supports the hypothesis that selection was targeted at regulatory sequences rather than protein structure (Doebley et al. 2006). Similar pattern can be observed in *TB1* homologs of other crop species. The alterations in *TB1* expression resulted in changes in tiller and panicle number in rice (Takeda et al. 2003, Choi et al. 2012). The overexpression of *TB1* in wheat resulted in a reduced tiller number and other changes in phenotype, including reduced plant height (Lewis et al. 2008). In this study significant marker-trait associations were detected between plant height and SNP/indel polymorphism in *LpTB1* gene. All four indels showing significant associations were situated in non-coding region, and two of the associated SNPs were synonymous thus did not alter protein structure. The only associated SNP, causing amino acid change from glycine to cysteine, was detected upstream of conserved TCP domain.

Coding mutations in barley ortholog of *TB1* correlated with lateral spikelet fertility phenotypes, yet the authors hypothesize that exonic variation is not causal but shows complete association with regulatory polymorphism due to extensive LD (Ramsay et al. 2011) as it is in maize, where reduced tillering is caused by a regulatory polymorphism 60 kb upstream of *TB1* gene (Clark et al. 2006). It remains to be determined, if this could be the case with perennial ryegrass as promoter region of *LpTB1* was not sequenced and the level of LD between regulatory and exonic sequences is not known. However due to outbreeding nature of perennial ryegrass LD decay is typically expected within 150-2000 bp (Ponting et al. 2007, Auzanneau et al. 2011, Fiil et al. 2011). Consequently exonic polymorphism observed in *LpTB1* could actually be linked to plant height differences. Possible effects of *TB1* sequence variation rather than expression level on other phenotypic traits remains unclear. No evidence for selection pressure in *LpTB1* was shown in a previous study (Brazauskas et al. 2010), yet the test panel in that study consisted of 20 cultivar accessions and only 4 of them being of turf type.

Short plant stature is a preferred trait in turf cultivar development. Furthermore, forage plant height is also important as it is one of the yield components. Therefore, even if breeders do not specifically target height as a main trait, taller plants might still be selected intuitively as this trait correlates with dry

matter yield. *LpTB1* was sequenced in 75 genotypes in total. Thirty-two genotypes were cultivars, and nearly half of them (15) were of turf type. The predominant allele at the SNP position 371 was significantly associated with a reduced plant height compared to heterozygous and homozygous variant genotypes, and it was identified in 10 out of 15 turf cultivars. The height of plants bearing predominant haplotype for all three associated SNPs was also significantly lower compared to homozygous for the less common haplotype, yet the LD r^2 values were 0.9 between each pair of SNPs. Therefore, it is likely that synonymous SNPs were in linkage with a causal SNP rather than alter the protein function themselves.

Conclusions

Plant height and dry matter yield associations with sequence polymorphism have been investigated in five *Lolium perenne* genes with putative role in shoot morphology. Three SNPs in *LpTB1* gene were associated with plant height and the alleles associated with reduced plant height were shown to be prevalent in turf type cultivars. The role of *TB* gene in tiller number determination has been well documented in many studies. In this study, we demonstrate a possible role of *LpTB1* in perennial ryegrass height determination. Identified SNPs could be of interest in ryegrass breeding for turf, especially in combination with molecular markers for leaf morphology.

Acknowledgments

This study was funded by the Research Council of Lithuania, grant No. MIP-032/2012 (FUMAG). The authors acknowledge Vidmantas Feiza for his assistance in sequence analysis.

References

Aleliūnas, A., K. Jonavičienė, G. Statkevičiūtė, D. Vaitiekūnaitė, V. Kemešytė, T. Lübberstedt, and G. Brazauskas, 2014: Association of single nucleotide polymorphism in LpIRI1 gene with freezing tolerance traits in perennial ryegrass. Euphytica, DOI:10.1007/s10681-014-1330-y.

Alves de Araújo, M.R., and C., Bruce, 2004: Genetic variation and correlation of agronomic traits in meadow bromegrass (*Bromus riparius* Rehm) clones. Ciência Rural **34**, 505-510.

Anhalt, U.C.M., J. Heslop-Harrison, S. Byrne, A. Guillard and S. Barth, 2008: Segregation distortion in Lolium: evidence for genetic effects. Theor. Appl. Genet. **117**, 297-306.

Armstead, I.P., L.B. Turner, M. Farrell, L. Skøt, P. Gomez, T. Montoya, I.S. Donnison, I.P. King, and M.O. Humphreys, 2004: Synteny between a major

heading-date QTL in perennial ryegrass (*Lolium perenne* L.) and the *Hd3* heading-date locus in rice. Theor. Appl. Genet. **108**, 822-828.

Armstead, I.P., L.B. Turner, A.H. Marshall, M.O. Humphreys, I.P. King, and D. Thorogood, 2008: Identifying genetic components controlling fertility in the outcrossing grass species perennial ryegrass (*Lolium perenne*) by quantitative trait loci analysis and comparative genetics. New Phytol. **178**, 559-571.

Auzanneau, J., C. Huyghe, J. Escobar-Gutiérrez, B. Jullier, F. Gastal, and P. Barre, 2011: Association study between the gibberellic acid insensitive gene and leaf length in a *Lolium perenne* L. synthetic variety. BMC Plant Biol. **11**, 183.

Barre, P., L. Moreau, F. Mi, L. Turner, F. Gastal, B. Julier, and M. Ghesquiére, 2009: Quantitative trait loci for leaf length in perennial ryegrass (*Lolium perenne* L.). Grass Forage Sci. **64**, 310-321.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler, 2007: TASSEL: Sotware for association mapping of complex traits in diverse samples. Bioinformatics **23**, 2633-2635.

Brazauskas, G., I. Lenk, M.G. Pedersen, B. Studer, and T. Lübberstedt, 2011: Genetic variation, population structure, and linkage disequilibrium in European elite germplasm of perennial ryegrass. Plant Sci. **181**, 421-420.

Brazauskas, G., I. Pašakinskienė, T. Asp, and T. Lübberstedt, 2010: Nucleotide diversity and linkage disequilibrium in five *Lolium perenne* genes with putative role in shoot morphology. Plant Sci. **179**, 194-201.

Buckler, E.S., J.M. Thornsberry, and S. Kresovich, 2001: Molecular diversity, structure and domestication of grasses. Genet. Res. Camb. **77**, 213-218.

Byrne, S., E. Guiney, S. Barth, I. Donnison, L.A.J. Mur, and D. Milbourne, 2009: Identification of coincident QTL for days to heading, spike length and spikelets per spike in *Lolium perenne* L. Euphytica **166**, 61-70.

Choi, M.S., M.O. Woo, E.B. Koh, J. Lee, T.H. Ham, K.S. Seo, and H.J. Koh, 2012: Teosinte Branched 1 modulates tillering in rice plants. Plant Cell Rep. **31**, 57-65.

Clark, R.M., E. Linton, J. Messing, and J.F. Doebley, 2004: Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc. Natl. Acad. Sci. USA **101**, 700-707.

Clark, R.M., T. Nussbaum Wagler, P. Quijada, and J. Doebley, 2006: A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. Nat. Genet. **38**, 594-597.

Conaghan, P., M.D. Casler, D.A. McGilloway, P. O'Kiely, and L.J. Dowley, 2008: Genotype × environment interactions for herbage yield of perennial ryegrass sward plots in Ireland. Grass Forage Sci. **63**, 107-120.

Doebley, J., A. Stec, and L. Hubbard, 1997: The evolution of apical dominance in maize. Nature **386**, 485-488.

Doebley, J.F., B.S. Gaut, and B.D. Smith, 2006: The molecular genetics of crop domestication. Cell **127**, 1309-1321.

Fiil, A., I. Lenk, K. Petersen, C.S. Jensen, K.K. Nielsen, B. Schejbel, J.R. Andersen, and T. Lübberstedt, 2011: Nucleotide diversity and linkage disequilibrium of nine genes with putative effects on flowering time in perennial ryegrass (*Lolium perenne* L.). Plant Sci. **180**, 228-237.

Forster, J.W., N.O.I. Cogan, M.P. Dobrowolski, M.G. Francki, G.C. Spangenberg, and K.F. Smith, 2008: Functionally associated molecular genetic markers for temperate pasture plant improvement. In: R.J. Henry (ed.), Plant genotyping II: SNP technology, 154-186. CABI Publishing, Cambridge.

Kobayashi, S., M.O. Humphreys, K. Tase, Y. Sanada, and T. Yamada, 2011: Molecular marker dissection of ryegrass plant development and its response to growth environments and foliage cuts. Crop Sci. **51**, 600–611.

Lewis, J.M., C.A. Mackintosh, S. Shin, E. Gilding, S. Kravchenko, G. Baldridge, R. Zeyen, and G.J. Muehlbauer, 2008: Overexpression of the maize *Teosinte Branched1* gene in wheat suppresses tiller development. Plant Cell Rep. **27**, 1217-1225.

Librado, P., and J. Rozas, 2009: DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**, 1451-1452.

Pašakinskienė, I., 2005: Creeping stem and barren inflorescence: a new mutation in Lolium perenne 'VIROIZ'. In: Z. Zwierzykowski, and S. Kosmala (eds.), Recent advances in genetics and breeding of the grasses, 209-213. Institute of Plant Genetics, Poznan.

Pauly, L., S. Flajoulot, J. Garon, B. Julier, V. Béguier, and P. Barre, 2012: Detection of favorable alleles for olant height and crown rust tolerance in three connected populations of perennial ryegrass (*Lolium perenne* L.). Theor. Appl. Genet. **124**, 1139-1153.

Ponting, R.C., M.C. Drayton, N.O. Cogan, M.P. Dobrowolski, G.C. Spangenberg, K.F. Smith, and J.W. Forster, 2007: SNP discovery, validation, haplotype structure and linkage disequilibrium in full-length herbage nutritive quality genes of perennial ryegrass (*Lolium perenne* L.). Mol. Genet. Genomics **278**, 585-597.

Ramsay, L., J. Comadran, A. Druka, D.F. Marshall, W.T.B. Thomas, M. Macaulay, K. MacKenzie, C. Simpson, J. Fuller, N. Bonar, P.M. Hayes, U. Lundqvist, J.D. Franckowiak, T.J. Close, G.J. Muehlbauer, and R. Waugh, 2011: *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is

an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. Nat. Genet. **43**, 169-172.

Salvi, S., and R. Tuberosa, 2005: To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci. **10**, 297-304.

Shortell, R.R., W.A. Meyer, and S.A. Bonos, 2009: Classification and inheritance of morphological and agronomic characteristics in Kentucky bluegrass (*Poa pratensis* L.). HortScience **44**, 274-279.

Skot, L., J. Humphreys, M.O. Humphreys, D. Thorogood, J. Gallagher, R. Sanderson, I.P. Armstead, and I.D. Thomas, 2007: Association of candidate genes with flowering time and water-soluble carbohydrate content in *Lolium perenne* (L.). Genetics **177**, 535-547.

Sokal, R.R., and F.J. Rohlf, 1981: Biometry, 2nd Ed. Freeman, San Francisco.

Statkevičiūtė, G., V. Kemešytė, N. Lemežienė, G.S. Konik, and M. Chomiak, 2012: Genetic diversity of perennial ryegrass wild ecotypes based on agro-biological traits. Proc. Latv. Acad. Sci., B Nat. exact appl. sci., **66**, 20-30.

Storey, J.D., J.E. Taylor, and D. Siegmund, 2004: Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: A unified approach. J. R. Stat. Soc. Ser. B Stat. Methodol. **66**, 187-205.

Studer, B., L.B. Jensen, S. Hentrup, G. Brazauskas, R. Kölliker, and T. Lübberstedt, 2008: Genetic characterisation of seed yield and fertility traits in perennial ryegrass (*Lolium perenne* L.). Theor. Appl. Genet. **117**, 781-791.

Takeda, T., Y. Suwa, M. Suzuki, H. Kitano, M. Ueguchi-Tanaka, M. Ashikari, M. Matsuoka, and C. Ueguchi, 2003: The OsTB1 gene negatively regulates lateral branching in rice. Plant J. **33**, 5138-5520.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar, 2013: MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. **30**, 2725-2729.

Utz, H.F., 2011: PLABSTAT-A computer program for statistical analysis of plant breeding experiments. Version 3A. Universität Hohenheim, Germany.

Wang, R.L., A. Stec, J. Hey, L. Lukens, and J. Doebley, 1999: The limits of selection during maize domestication. Nature **398**, 236-239.

Wang, Z., A. Hopkins, and R. Mian, 2001: Forage and Turf Grass Biotechnology. Crit. Rev. Plant Sci. **20**, 573–619.

Wilkins, P.W., 1991: Breeding perennial ryegrass for agriculture. Euphytica **52**, 201-214.

Xing, Y., U. Frei, B. Schejbel, T. Asp, and T. Lübberstedt, 2007: Nucleotide diversity and linkage disequilibrium in full-length herbage nutritive quality

genes of perennial ryegrass *(Lolium perenne* L.). Mol. Genet. Genomics **278**, 585-597.

Yamada, T., E.S. Jones, N.O.I. Cogan, A.C. Vecchies, T. Nomura, H. Hisano, Y. Shimamoto, K.F. Smith, M.D. Hayward, and J.W. Forster, 2004: QTL analysis of morphological, developmental, and winter hardiness-associated traits in perennial ryegrass. Crop Sci. 44, 925-935.

Zhu, C., M. Gore, E.S. Buckler, and J. Yu, 2008: Status and prospects of association mapping in plants. Plant Genome **1**, 5-20.

Table 1: Variation and heritability for plant height and dry matter yield in 96 perennial ryegrass genotypes.

	Plant Height (cm)			Dry matter yield (g)		
	2013	2014	2 year average	2013	2014	2 year average
Minimum	25.7	38.0	33.4	15.2	22.1	23.9
Maximum	69.3	91.3	76.8	147.6	337.2	241.6
Mean	49.8	70.5	60.1	78.3	154.9	116.3
SD[1]	9.4	10.3	9.5	33.1	74.4	44.1
Heritability			0.80			0.32

[1]Standard deviation

Table 2: Diversity in five perennial ryegrass genes.

	Alignment length bp	Segregating site number	Haplotype number	Hd[2]	π[3]	Tajima's D	Indel number	Total Indel length
LpIAA1 n[1]=94								
Total	1129	39	24	0.804	0.0075	0.1559	8	165
Coding	465	12	11	0.688	0.0046	0.0471	1	3
Non-coding	664	27	20	0.793	0.0103	0.1912	7	162
LpBRI1 n=94								
Total	1168	75	27	0.715	0.0066	-1.4924	18	84
Coding	507	45	10	0.243	0.0090	-1.3397	0	–
Non-coding	661	29	22	0.678	0.0043	-1.5198	18	84
LpRUB1 n=96								
Total	1042	95	55	0.838	0.0102	-1.5099	43	181
Coding	306	16	13	0.694	0.0055	-1.0974	0	–
Non-coding	736	79	50	0.817	0.0129	-1.5391	43	181
LpSHOOT1 n=92								
Total	1097	126	49	0,908	0,0188	-0.4493	19	55
Coding	357	38	28	0.845	0.0185	0.1949	0	–
Non-coding	740	88	40	0.893	0.0189	-0.5499	19	55
LpTB1 n=74								
Total	1509	50	38	0.776	0.0017	-2.2245**	15	33
Coding	1095	39	32	0.736	0.0019	-2.1019*	3	8
Non-coding	414	11	10	0.236	0.0009	-2.0286*	12	25

[1]number of genotypes.

[2]haplotype diversity.

[3]nucleotide diversity.

* significant at $p < 0.05$, ** $p < 0.01$.

Table 3: Significant associations in *LpTB1* gene.

Locus	Coding/ Noncoding[1]	Trait[2]	*p*	*q*	Allele	Effect	n
23	N	Height 2014	0.023	0.049	-:T	4.06	22
					-:-	7.04	28
					T:T	0	24
24	N	Height 2014	0.046	0.049	-:C	-9.56	22
					-:-	-6.89	27
					C:C	-13.67	24
					-:T	0	1
371	C(m) Gly→Cys	Height 2014	0.009	0.049	G:T	7.12	20
					T:T	6.76	25
					G:G	0	29
663	C(s)	Height 2014	0.018	0.049	A:C	6.97	18
					C:C	5.53	26
					A:A	0	30
911	C(s)	Height 2014	0.041	0.049	T:G	1.06	20
					T:T	-4.97	31
					G:G	0	23
1329-1332	N	DMY 2014	0.017	0.049	-:-	62.41	52
					-:TCGA	62.50	12
					TCGA:TCGA	0	10
		DMY 2 years average	0.035	0.049	-:-	36.69	52
					-:TCGA	46.77	12
					TCGA:TCGA	0	10
1501-1508	N	Height 2014	0.049	0.049	-:-	-5.09	30
					-:CATGTAAT	0.64	21
					CATGTAAT:CATGTAAT	0	23

[1]N: non-coding region; C(m): missens mutation; C(s): synonymous substitution.

[2]DMY: dry matter weight.

Figures

Fig. 1: Plant height correlation between 2013 and 2014.

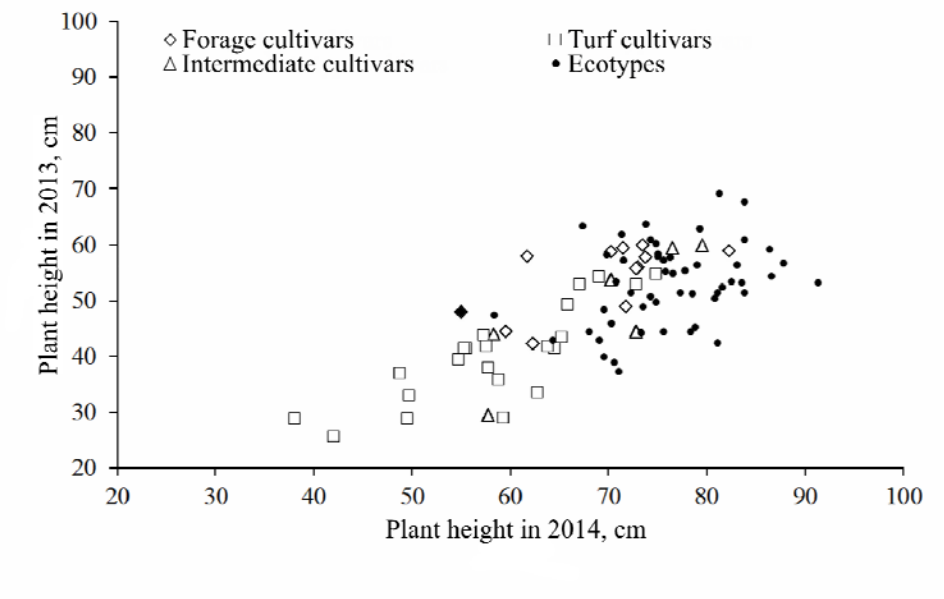


Fig. 2: Pattern of linkage disequilibrium (LD) in five pooled perennial ryegrass genes. The logarithmic trend line shows average LD decay.

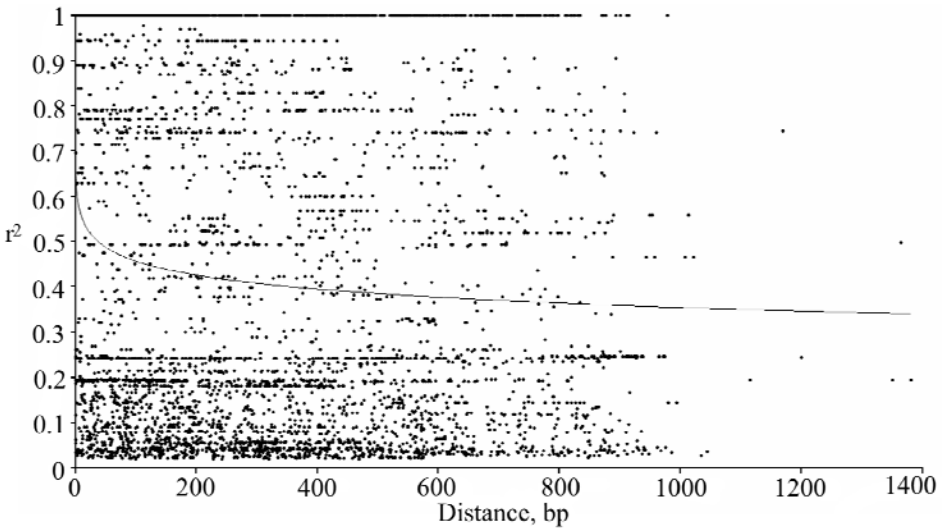
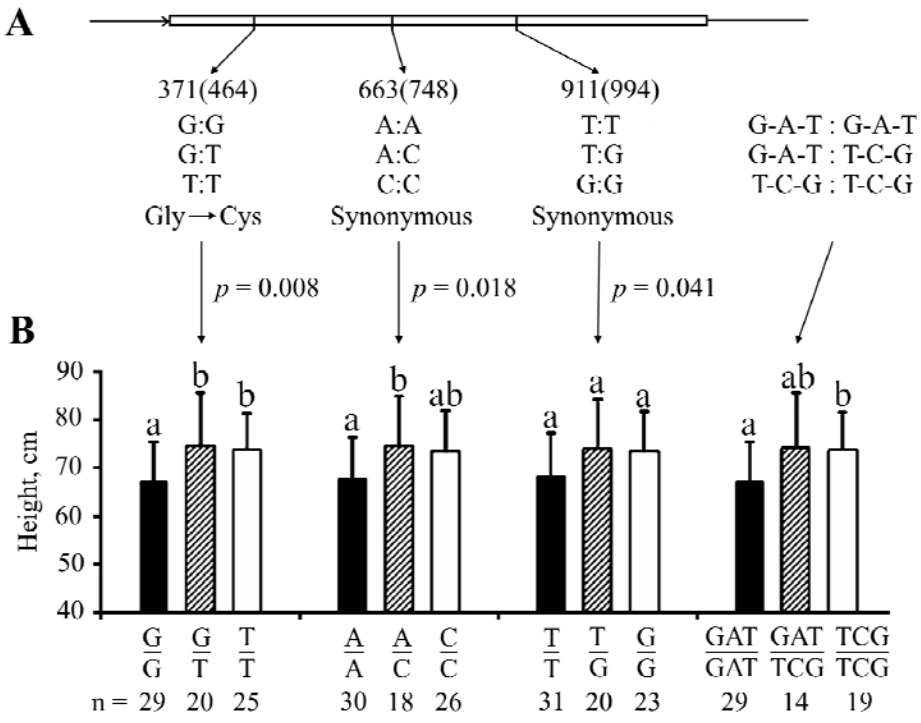


Fig. 3: Marker-trait association and variant alleles of *LpTB1* gene. A – gene structure and relative positions of significant SNPs. The SNP numbers were given according to their position in the alignment of 74 genotypes. SNP positions relative to the reference sequence GenBank GU987123.1 are given in parenthesis. B – plant height in 2014 against the significant associated markers. The *p* values for each SNP are given above the chart. Height values with the same letter are not significantly different ($p < 0.05$). The error bars represent standard deviation of the group mean.



Supporting Information

Supp. Fig. 1: Pattern of linkage disequilibrium (LD) in five perennial ryegrass genes. The logarithmic trend line shows average LD decay.