Theor Appl Genet (2013) 126:1775–1782
DOI 10.1007/s00122-013-2091-z

ORIGINAL PAPER

Diversity of sequences and expression patterns among alleles of a sugarcane loading stem gene

Richard L. Moyle · Robert G. Birch

Received: 20 August 2012 / Accepted: 21 March 2013 / Published online: 2 April 2013

Abstract Modern sugarcane cultivars are highly polyploid and aneuploid hybrids, which are propagated as clones. Their complex genome structure comprises 100–130 chromosomes and 10–13 hom(e)ologous copies of most loci. There is preliminary evidence of very high heterozygosity, with implications for genetic improvement approaches ranging from marker-assisted selection to transgenics. Here, we report that sugarcane cultivar Q200 has at least nine alleles at the *Loading Stem Gene* (*ScLSG*) locus. Exon–intron structure is identical and the predicted protein products show at least 92 % identity, across sugarcane alleles and the *Sorghum* homologue Sb07g027880. There is substantial variation in the 5′ UTR and promoter regions including numerous allele-specific nucleotide polymorphisms, insertions and deletions. We developed an allele-specific qRT-PCR method to undertake the first compelling test of allele-specific expression in polyploid sugarcane. Seven alleles distinguished by this method all showed peak expression in the sucrose-loading zone of the stem, but there was apparent variability in expression patterns across other tissues. The *ScLSG2* and *ScLSG5* alleles appear promising for specificity of expression in stems, relative to leaf, meristem, emerging shoot and root tissues. Within the stem, there was activity in parenchyma, vascular and rind tissues. This expression pattern is of interest in basic research and biotechnology aimed at enhanced sucrose content, engineering value-added products, and manipulation of stem biomass composition.

Communicated by A. Schulman.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2091-z) contains supplementary material, which is available to authorized users.

R. L. Moyle · R. G. Birch (✉)
Hines Plant Science Building, University of Queensland, Brisbane 4072, Australia
e-mail: r.birch@uq.edu.au

Introduction

Sugarcane (*Saccharum* spp. hybrid) is a major crop of tropical and sub-tropical regions, contributing more than 75 % of world sugar production (OECD-FAO 2010). High biomass accumulation makes sugarcane a strong candidate among potentially sustainable bioenergy feedstocks (Waclawovsky et al. 2010). It has long been selected by humans for high sugar concentration in stems, and there is substantial interest in the potential for metabolic engineering aimed at enhancing value as a biomaterial and biofuel crop (Arruda 2012; Birch 2007; Dal-Bianco et al. 2012). *Saccharum* species are wind-pollinated autopolyploids in the wild, and modern cultivars are based primarily on high-sucrose accessions of *S. officinarum*, with introgression from *S. spontaneum* for enhanced disease resistance, vigour and fibre content. The monoploid chromosome set is estimated at 10, and cultivars are typically complex aneuploids with 100–130 chromosomes totalling approximately 10 Gb of DNA (D'Hont et al. 1996). About 70–80 % of the chromosomes are derived from *S. officinarum*, 10–20 % from *S. spontaneum* with the remainder from interspecific recombinations (Cuadrado et al. 2004; D'Hont 2005; Piperidis et al. 2010). Thus, modern sugarcane cultivars have a complex genome structure with 10–13 hom(e)ologous copies of most loci. Molecular marker studies and sequencing of genomic PCR products and BAC library clones indicate high heterozygosity (Bundock et al. 2009; Cordeiro et al. 2000; D'Hont

et al. 1999; Garsmeur et al. 2011), but few loci have been analysed at the transcriptional level.

Polyploidy occurs in many plant systems. Evolutionarily, it may confer advantages in vigour, repertoire of responses to environmental challenges, and opportunities for emergence of new gene functions from redundant copies (Adams and Wendel 2005; Jackson and Chen 2010). There are theoretical possibilities to exploit allele-specific expression patterns or isoform interactions in multi-subunit proteins (Osborn et al. 2003). However, these would add constraints of allele dosage to those of gene-product dosage. Any substantial increase in genetic complexity theoretically necessitates tighter control to avoid disruption from background expression levels. How plants with complex polyploid and aneuploid genomes regulate expression in these contexts is poorly understood. Although many crop species function as diploids, after evolutionary cycles of polyploid formation and diploidization (Soltis et al. 2004); others are stable autopolyploids (e.g. potato and banana) or allopolyploids (e.g. cotton, wheat and canola).

In allopolyploids, the presence of divergent genomes within a common nucleus creates the potential for various intergenomic interactions, including homoeologous recombination (Osborn et al. 2005; Pires et al. 2007), genomic deletions (Feldman et al. 2001; Levy et al. 2002) and modifications of homoeologous gene expression (Bottley et al. 2006; Wendel et al. 2003, 2008). Among 40 homoeologous gene pairs examined in tetraploid cotton, ten exhibited biased expression patterns, including developmental differences and reciprocal silencing in different floral whorls (Wendel et al. 2003). Microarray analyses indicate that homoeologue expression biases are common in cotton fibre cells, and that differential expression can vary temporally and developmentally (Flagel et al. 2008; Wendell et al. 2008). Genes with close physical linkage in one region showed correlated homoeologue expression biases, inferring the action of factors such as a chromatin modification or a shared enhancer affecting a large genomic region. However, closely linked genes from a separate region showed uncorrelated expression patterns, indicating fine-scale regulatory factors such as localised divergence among homoeologous promotors or epigenetic marks, or differential post-transcriptional regulation of alleles (Flagel et al. 2009).

Sugarcane is remarkable for stable cultivars that combine elements of aneuploidy, auto- and allo-polyploidy with high heterozygosity. Expressed sequence tag (EST) resources indicate that multiple alleles are expressed (Grivet et al. 2001, 2003). Among eight alleles of a SHAQYF transcription factor in sugarcane cultivar Q117, one allele was estimated to contribute over 40 % of expressed transcripts in mature internodes whereas four other alleles contributed <1 % to the transcript pool (Mudge et al. 2009). Understanding the allelic variation across more loci should help in the daunting challenge to assemble an annotated sugarcane genome database and indicate the feasibility of marker-assisted selection. An understanding of the control of allelic expression will aid in the selection of suitable promoter sequences to drive desired patterns of transgene expression in biotechnology.

In this study, we examined allelic variation in a gene preferentially expressed in the sucrose-loading zone of the sugarcane stem. We characterised ten alleles of this gene, designated *ScLSG* and developed allele-specific primer pairs that allowed qRT-PCR analysis of expression patterns of seven alleles across different tissues and developmental stages. The results extend the emerging understanding of structural diversity between multiple expressed alleles in a typical sugarcane cultivar, and indicate several candidate promoter sequences for preferential expression of transgenes in the sucrose-accumulating tissues of the sugarcane stem.

Materials and methods

Choice of target locus

Sugarcane EST clones CA081293 and CA260047 were potentially up-regulated in a maturing stem internode, based on microarray data (Casu et al. 2007). These EST clones were BLASTN searched against the entire sugarcane EST sequence collection in GenBank. Alignment of the retrieved EST sequences revealed that CA081293 and CA260017 are derived from the same gene. The consensus sequence from sugarcane EST alignment was BLASTX searched against the GenBank protein database, revealing high similarity with proteins of unknown function from other plant species. Primers were designed to the sugarcane EST consensus sequence (5′ primer ATGGGCGCAAGAG ATAATGGGGAG, 3′ primer TCACCGTTGCTCCAACC ATACTTC) and used to amplify full-length coding sequences from a maturing internode of sugarcane cultivar Q200. The amplified cDNA was used as a probe in northern blot analysis, which confirmed that this gene is up-regulated in the sucrose-loading zone of sugarcane stems (see "Results"). We designated this gene *ScLSG* (Sugarcane Loading Stem Gene).

BAC library screening

A previously described sugarcane Q200 BAC library (Mudge et al. 2009) was screened using the amplified *ScLSG* cDNA as a probe. DNA was extracted from BAC clones using Plasmid Maxiprep or Midiprep kits

(QIAGEN), and sequenced by primer walking (AGRF and AEGRC, Brisbane, Australia). Sequences from each BAC clone were assembled using the ContigExpress module of Vector NTI (Invitrogen).

Plant tissues, RNA isolation and cDNA synthesis

Leaf, meristem and internode (IN) samples were harvested from three replicate field-grown Q200 plants, numbering from +1 for the top visible dewlap, attached leaf, node and IN below (Van Dillewijn 1952). Young stems, shoots and roots were each harvested from two replicate glasshouse-grown Q200 plants. RNA was also isolated from rind, vascular and parenchyma tissues dissected from IN5 of glasshouse-grown Q200 plants.

RNA for northern blot analysis and qRT-PCR was isolated using TRIzol® reagent (Invitrogen). RNA for qRT-PCR was treated using 2 U of DNase I (New England Biolabs) at 37 °C for 30 min followed by enzyme inactivation at 70 °C for 10 min. cDNA was then synthesised from 2 μg of RNA using Superscript III reverse transcriptase (Invitrogen) and Oligo(dT)$_{20}$.

Quantitative real-time PCR, northern blot and sugar analysis

qRT-PCR primers were designed to sequences conserved across all of the *ScLSG* alleles (*ScLSG* qRT5′ GTTGGAG CAACGGTGATG, *ScLSG* qRT3′ GGGAATGTACTGT ACAAGAAAGTG) and predicted to yield a 149-bp amplicon. Primers containing allele-specific polymorphisms (Supplementary Table 1) were designed using Vector NTI (Invitrogen). qRT-PCR was performed using FastStart Universal SYBR Green Master (ROX) (Roche

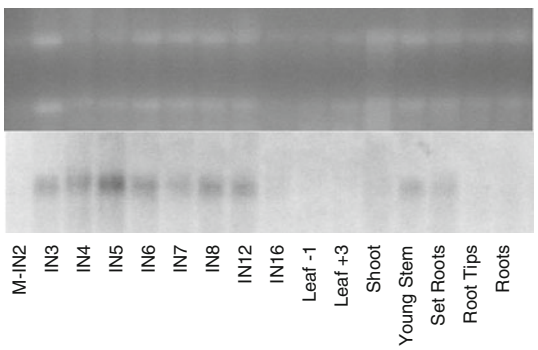

Fig. 1 Northern analysis of total RNA from sugarcane Q200, showing the ethidium–bromide-stained gel above the northern blot probed using *ScLSG* cDNA. Sampled tissues were meristem to internode 2 (M-IN2), numbered internodes (IN#) and leaves from field-grown plants; and the following from glasshouse-grown plants: shoots (2-cm long), young stems (30 cm), set roots (2 weeks after planting), root tips and remaining root tissues (4 months after planting)

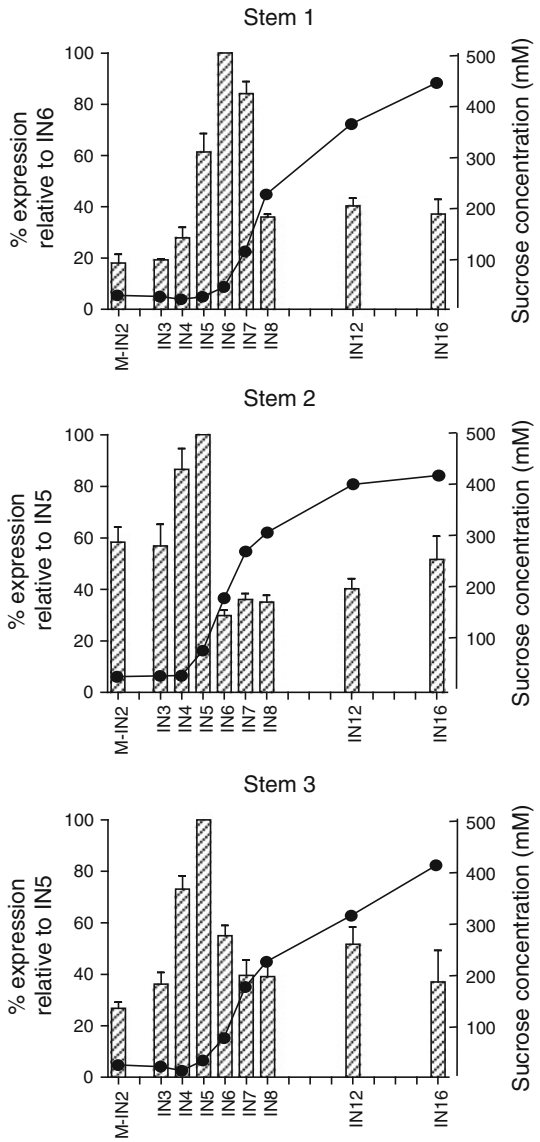

Fig. 2 *ScLSG* expression in relation to sucrose accumulation in three stems of field-grown Q200 (20 internode stage of development). qRT-PCR data were normalised using GAPDH as the reference gene. Relative expression was then converted to a percentage of the level in the highest expressing internode. *Bars* show means with standard errors from three or four qRT-PCRs per sample using a conserved primer pair (left *Y* axis). The *line graphs* show sucrose concentration in juice from the same samples (right *Y* axis)

Diagnostics) in a 7900HT thermocycler (Applied Biosystems). After denaturation at 95 °C for 10 min, the qRT-PCR cycles consisted of denaturation at 95 °C for 15 s followed by annealing/extension at 62 °C for 60 s for a total of 45 cycles, with the exception of allele *ScLSG1* for which annealing/extension was at 62 °C for 90 s. Differences in expression were calculated using the comparative C_T method ($2^{-\Delta\Delta C_t}$). GAPDH was used as the reference gene (Iskandar et al. 2004). Statistical analysis was undertaken using GraphPad Prism software. Data from

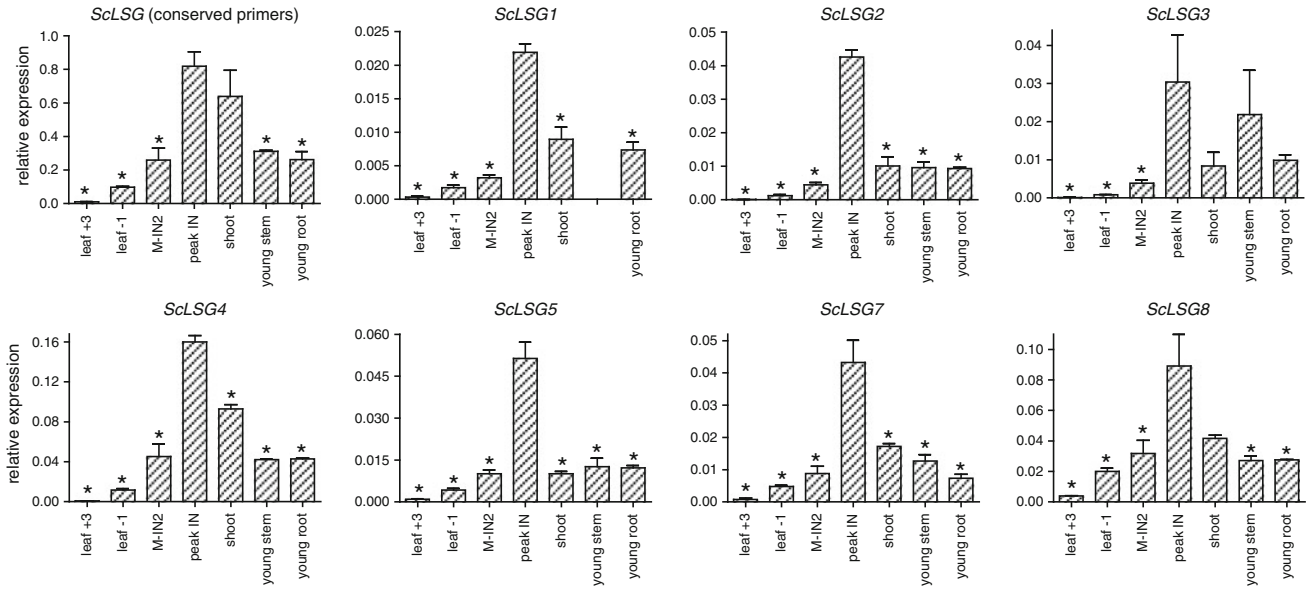

Fig. 3 *ScLSG* expression pattern in sugarcane Q200 indicated by qRT-PCR using conserved or allele-specific primer pairs. Tissues are as detailed in Fig. 1 except that young roots were from glasshouse-grown plants (4 months after planting). Peak IN is the highest expressing internode (IN5–7). The data were normalised to the GAPDH reference gene using the comparative C_t method ($2^{-\Delta\Delta C_t}$). *Bars* show means with standard errors from two (glasshouse-grown) or three (field-grown) replicate plants. *Astersik* indicates significantly different from peak IN ($P < 0.05$)

multiple tissues from the same plants were analysed using ratio *t* tests. Data from different plants were analysed using unpaired *t* tests. In all cases, $P < 0.05$ was applied as the criterion for statistical significance.

Northern blot analysis was undertaken as described by Mudge et al. (2009). Sucrose was quantified by high-performance anion-exchange chromatography with pulsed amperometric detection (Wu and Birch 2007).

Results

Identification of a gene that is preferentially expressed in the sucrose-loading region of sugarcane stems

We commenced with EST clones CA081293 and CA260047, which appeared by microarray analysis to be up-regulated in the maturing stem (Casu et al. 2007). After a BLASTN search against the sugarcane EST collection in GenBank and alignment of the matching EST sequences, consensus primers were designed to amplify full-length coding sequences from sugarcane Q200 IN6 cDNA. Northern analysis using this cDNA as a probe indicated strongest expression in stem tissues, particularly the zone of rapid sucrose accumulation (IN3–12 in Fig. 1). The gene was designated *ScLSG*. Multiple alleles were cloned, sequenced and aligned as described below.

qRT-PCR using primers matching sequences conserved across *ScLSG* alleles, applied to RNA from internodes of field-grown Q200 stems, showed peak *ScLSG* expression around the transition zone to rapid sucrose accumulation (Fig. 2). There is slight variation between stalks in the internode at which this peak occurs (between IN5 and IN7). Therefore, we used the peak-expressing internode for comparison with other tissues. This indicated relatively low expression in leaf, meristem or root tissues (Fig. 3, conserved primers). The internodes exhibiting peak *ScLSG* expression had significantly higher *ScLSG* transcript levels than leaves, stem apex regions, young internodes or young roots. Expression in emerging shoots appeared higher by qRT-PCR than by northern analysis on the same RNA samples, possibly reflecting the different basis for generating the primary signals and normalising the results in these methods (northern hybridisation intensity to full-length cDNA probe amplified using conserved primers and normalised through equivalent RNA loading, versus qRT-PCR amplification rate of a gene fragment flanked by conserved primers and normalised relative to GAPDH amplification rate).

qRT-PCR on RNA extracted from dissected rind, vascular and storage parenchyma at IN5 of a glasshouse-grown Q200 stem revealed *ScLSG* expression in each tissue type and highest in the storage parenchyma (Fig. 4).

Allelic diversity at the *ScLSG* locus

In a BAC library estimated to contain about 80 % of the Q200 genome (Mudge et al. 2009), 12 clones hybridised to

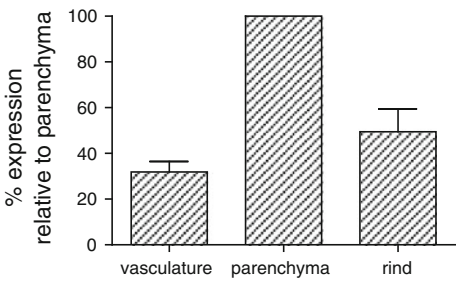

Fig. 4 *ScLSG* expression in tissues dissected from IN5 of glasshouse-grown sugarcane Q200. qRT-PCR data were normalised using GAPDH as the reference gene. Relative expression was then converted to a percentage of the level in parenchyma. *Bars* show means with standard errors from four qRT-PCRs per sample using a conserved primer pair

the *ScLSG* probe. DNA sequence analysis revealed two duplicate pairs, leaving ten unique clones. BAC10 contained partial 3′ *ScLSG* coding sequence and BAC9 contained non-overlapping partial 5′ *ScLSG* coding sequence (Fig. 5); so the possibility that these are derived from the same allele cannot be excluded. BAC7 contained unique partial 3′ *ScLSG* coding sequence, and the full-length allele was obtained by a genomic PCR using an allele-specific 3′ primer (ACCACGCCTCCAACCCAT) and a conserved 5′ primer (AAGTACAAACGTAGACTCTG). The remaining BAC clones contained full *ScLSG* coding sequences, 5′ UTR and at least 2 kb of putative promoter region. The alleles were designated *ScLSG1–ScLSG10* and the nine sequences with promoter regions are in GenBank Accessions JQ920355–JQ920358 and JX514698–JX514702. A BlastN search of the GenBank sugarcane EST database on 25/11/2011 identified 33 *ScLSG* ESTs. Even allowing for the limitations of EST sequence accuracy, it appears that additional *ScLSG* alleles exist in various sugarcane cultivars.

The *ScLSG* genomic sequences from Q200 share a predicted pattern of seven exons in the coding sequence and a large intron in the 5′ UTR (Fig. 5). The full-length alleles encode either 690 amino acids (*ScLSG8*), 691 amino acids (*ScLSG1, 2, 3, 4, 6, 7*) or 695 amino acids (*ScLSG5*). The predicted proteins share between 96 and 100 % amino-acid identity and show 92–95 % identity to their 695 amino-acid *Sorghum* homologue (Sb07g027880). The protein function is unknown. At the nucleotide level, the *ScLSG* coding sequences share 97–99.9 % identity and show 94–96 % identity to Sb07g027880. The *ScLSG* intron sequences (excluding the 5′ UTR intron) share 90–99 % identity and show 85–87 % identity to Sb07g027880.

ScLSG alleles vary more in the 5′ UTR and promoter regions, with multiple single-nucleotide polymorphisms, insertions and deletions (Fig. 5). In the promoter region, *ScLSG7* and *ScLSG8* share a large 1,220-bp deletion relative the other alleles. *ScLSG2* and *ScLSG3* share similar 230 bp insertions. *ScLSG1*, *ScLSG4* and *ScLSG6* each contain at least one unique insertion of >60 bp. *ScLSG6* contains a 1,000-bp inverted repeat of adjacent sequence. The 5′ UTR intron of *ScLSG8* contains an 800-bp insertion.

The origin of the insertion sequences was investigated using Repbase (Kohany et al. 2006). All but one of the insertions longer than 60 bp in the putative promoter and 5′ UTR regions (Fig. 5) showed homology to transposable elements. The 65-bp insertion in *ScLSG1* showed 81 % similarity across 54 bp to an L1-type non-LTR retrotransposon from *Sorghum bicolor*. The 230-bp insertion in *ScLSG2*, *ScLSG3*, *ScLSG5* and the 200-bp insertion in *ScLSG6* showed 71–82 % similarity across 192–222 bp to

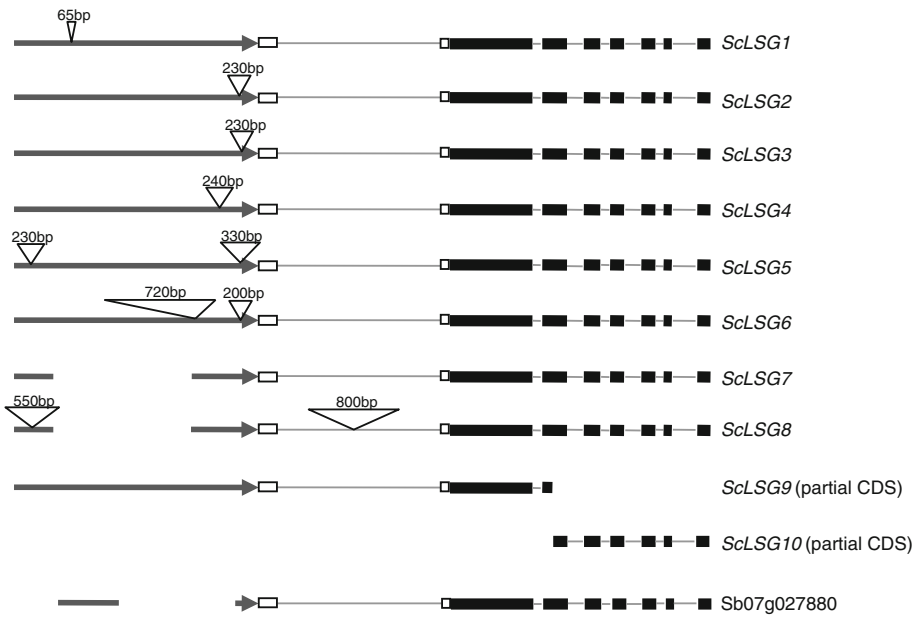

Fig. 5 Structure of *ScLSG* alleles isolated from sugarcane Q200 and *Sorghum* homologue Sb07g027880. *Open blocks* represent 5′ UTR, *closed blocks* represent coding exons and *narrow lines* represent introns. *Arrows* represent promoters, *triangles* represent insertions longer than 65 bp and *gaps* represent deletions. Only partial sequence was obtained for *ScLSG9* and *ScLSG10*

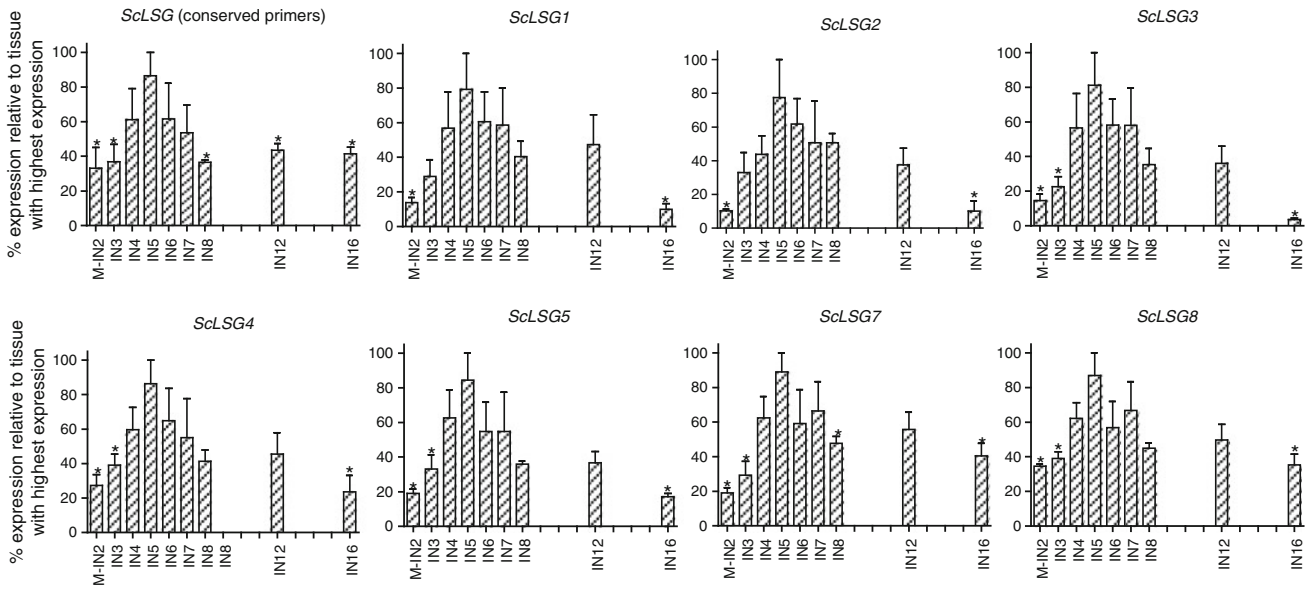

Fig. 6 Stem profiles of *ScLSG* expression in field-grown Q200 stems, based on qRT-PCR using consensus or allele-specific primers. The data were normalised using GAPDH as the reference gene. Relative expression was then converted to a percentage of the level in the highest expressing internode (IN5–7). *Bars* show means with standard errors from three replicate stems. *Asterisk* indicates significantly different from IN5 ($P < 0.05$)

the TSB2 non-autonomous Harbinger-type DNA transposon from *S. bicolor*. The 720-bp insertion in *ScLSG6* showed 86 % similarity across 708 bp to MUDR non-autonomous DNA transposon from *S. bicolor*. The 240-bp insertion in *ScLSG4* and the 550-bp insertion in *ScLSG8* showed 74 and 77 % similarity across 216 bp to the Tourist 1a Harbinger-type DNA transposon from *S. bicolor*. The 550-bp insertion in *ScLSG8* contains a 158-bp region with 84 % similarity to Tourist 1c Harbinger-type DNA transposon from *S. bicolor*. The 800-bp insertion in *ScLSG8* showed 71 % similarity across 288 bp with MUDRN3 non-autonomous DNA transposon from *Oryza sativa*.

Multiple *ScLSG* alleles share similar expression patterns across the maturing sugarcane stem but exhibit different expression patterns among other tissues

The specificity of primer pairs designed to *ScLSG* alleles 1, 2, 3, 4, 5, 7 and 8 (Supplementary Table 1) was confirmed using the suite of *ScLSG* BAC clones as templates in PCRs (Supplementary Fig. 1). We were unable to design allele-specific primer pairs to *ScLSG6* (tested primers gave spurious amplicons) or *ScLSG9* and *ScLSG10* (incomplete coding sequences). The allele-specific qRT-PCR approach cannot be used to determine absolute expression levels (because the relative efficiencies of different primer sets in critical early PCR cycles are unknown), but it can be used to compare expression specificity across tissues.

qRT-PCR using the allele-specific primers, in each case indicated a peak of expression between IN5 and IN7 across three replicate stems (Fig. 6). For each allele, expression in IN5 was significantly higher than in the meristem region or IN16. Comparing other tissues, each allele had significantly lower expression in the leaf, apex and root tissues relative to the peak expressing internode (Fig. 3). All except *ScLSG3* had significantly lower expression in shoots, young stems and young roots relative to the peak-expressing internode. There was apparent variation between alleles in some other respects, such as shoot:root expression ratios.

Discussion

Sugarcane cultivar Q200 is highly heterozygous at the *ScLSG* locus, with nine or ten alleles recovered from an incomplete BAC library. The total number of 12 hybridising clones in this library indicates a single locus in the sugarcane genome (Mudge et al. 2009). The closely related diploid *Sorghum* genome contains a single locus for the orthologous gene Sb07g027880. Combining the detailed analyses of the *ScLSG* and *ScR1MYB1* (Mudge et al. 2009) loci, a recent BAC sequencing study (Garsmeur et al. 2011) and the indications from sugarcane EST (Marconi et al. 2011) and isozyme analyses (Glaszmann et al. 1989); the emerging picture is one of very high heterozygosity at many gene loci in sugarcane. In the loci explored in greatest detail, the number of alleles per cultivar approaches the theoretical limit of around 12 in this complex polyploid genome.

Furthermore, our analysis at the transcript level indicates that most or all of the *ScLSG* alleles are expressed. There were sufficient unique polymorphisms in *ScLSG* sequences to perform qRT-PCRs that distinguish seven alleles, all of which were expressed. The development of validated allele-specific primers permitted the first compelling test for allele-specific developmental expression patterns in sugarcane. Previous explorations based on EST frequencies or sequencing of RT-PCR products have been limited by the available number of sequenced cDNA clones. In the case of *ScLSG*, seven tested alleles showed similar expression profiles along the sucrose-loading zone in field-grown stalks, but varied in other tissues (Figs. 3, 6). Our results indicate that it is prudent to undertake expression analysis on single sugarcane alleles, before selecting for biotechnology applications promoters from a locus with an expression pattern deduced from data integrating multiple alleles.

The 5′ UTR and promoter regions of the identified *ScLSG* alleles exhibit substantial structural and sequence diversity. Taking *ScLSG1* as a reference sequence, all of the eight other sequenced upstream regions contain an insertion or deletion >65 bp in size, in addition to numerous single-nucleotide polymorphisms and small insertions and deletions. All but one of the insertions show high sequence similarity to known transposable elements from *Sorghum* and rice. These structural differences likely contribute to the variation in developmental expression patterns and can be explored further using transgenic approaches with reporter fusions.

The added complexity for marker-assisted selection of traits based on multi-allelic expression has been mentioned previously (Mudge et al. 2009), and is reinforced by the evidence presented here for allele-specific variations in expression patterns. Another practical implication is that understanding of the extent of heterozygosity and the detailed sequence variations at key loci will be important for the design of constructs to reduce expression of some or all alleles of a sugarcane locus by RNAi (Osabe et al. 2009).

Extending our knowledge of the variations between haplotypes will also aid in developing strategies to efficiently sequence and correctly assemble the sugarcane genome, currently one of the greatest challenges in genomics because of complex polyploidy (Dal-Bianco et al. 2012; Souza et al. 2011). Using the *ScLSG* locus as an example, there is insufficient sequence variation in the coding regions to distinguish multiple haplotype sequences with current high-throughput, short-read, whole-genome sequencing methods. However, the presence of multiple allele-specific SNP's and InDels in the 5′ UTR and promoter regions points to the use of mate-paired library strategies to match reads from regions of low allelic variability (such as coding sequences) with reads derived from regions of high variability (such as introns and intergenic regions). The emergence of long-read, deep-sequencing platforms could also help to build sequence scaffolds for individual haplotypes.

Northern blot and qRT-PCR analyses indicate that overall *ScLSG* expression peaks near the sucrose-loading zone of the sugarcane stem and continues in mature internodes, with relatively low expression in most other tissues including leaves, the shoot apical meristem and roots. Furthermore, expression is higher in the sucrose-storage parenchyma than in fibrous rind or vascular tissues at IN5. This expression pattern differs from sugarcane-stem genes characterised in substantial detail for preferential expression in mature-stem internodes (Mudge et al. 2009) or in fibrous cell types (Damaj et al. 2010). The pattern exemplified by *ScLSG*2 and *ScLSG*5 is of interest in sugarcane biotechnology because of the importance of endogenous processes in stems for sucrose and biomass accumulation, and because of the potential through metabolic engineering to accumulate novel biomaterials derived from stored sucrose. We have recently shown that several of the isolated *ScLSG* promoter regions drive useful transgene expression patterns in sugarcane (Moyle and Birch 2013; Mudge et al. 2013).

Acknowledgments The authors acknowledge the excellent technical assistance of Lilian Chou throughout this project. We thank BSES Limited for access to field-grown sugarcane samples. This research was supported through a collaboration between CSR Sugar Limited (Sucrogen) and The University of Queensland under the Australian Research Council's Linkage scheme.

References

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8:135–141

Arruda P (2012) Genetically modified sugarcane for bioenergy generation. Curr Opin Biotechnol 23:315–322

Birch RG (2007) Metabolic engineering in sugarcane: assisting the transition to a bio-based economy. In: Verpoorte RA, Alfermann W, Johnson TS (eds) Applications of plant metabolic engineering. Springer, Berlin, pp 249–281

Bottley A, Xia GM, Koebner RMD (2006) Homoeologous gene silencing in hexaploid wheat. Plant J 47:897–906

Bundock PC, Eliott FG, Ablett G, Benson AD, Casu RE, Aitken KS, Henry RJ (2009) Targeted single nucleotide polymorphism (SNP) discovery in a highly polyploid plant species using 454 sequencing. Plant Biotechnol J 7:347–354

Casu RE, Jarmey JM, Bonnett GD, Manners JM (2007) Identification of transcripts associated with cell wall metabolism and development in the stem of sugarcane by Affymetrix GeneChip Sugarcane Genome Array expression profiling. Funct Integr Genomics 7:153–167

Cordeiro GM, Taylor GO, Henry RJ (2000) Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. Plant Sci 155:161–168

Cuadrado A, Acevedo R, Moreno Diaz de la Espina S, Jouve N, de la Torre C (2004) Genome remodelling in three modern *S. officinarum* × *S. spontaneum* sugarcane cultivars. J Exp Bot 55:847–854

Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann HP, Garcia AA, Souza GM (2012) Sugarcane improvement: how far can we go? Curr Opin Biotechnol 23:265–270

Damaj MB, Kumpatla SP, Emani C, Beremand PD, Reddy AS, Rathore KS, Buenrostro-Nava MT, Curtis IS, Thomas TL, Mirkov TE (2010) Sugarcane DIRIGENT and *O*-methyltransferase promoters confer stem-regulated gene expression in diverse monocots. Planta 231:1439–1458

D'Hont A (2005) Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. Cytogenet Genome Res 109:27–33

D'Hont A, Grivet L, Feldmann P, Rao S, Berding N, Glaszmann JC (1996) Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol Gen Genet 250:405–413

D'Hont A, Jannoo N, Grivet L, Seguin M, Paulet F, Domaingue R, Rao PS, Dookun A, Glaszmann JC (1999) Molecular investigation of the genetic base of sugarcane cultivars. Theor Appl Genet 99:171–184

Feldman M, Ozkan H, Levy AA (2001) Allopolyploidy-induced rapid genome evolution in the wheat (Aegilops-Triticum) group. Plant Cell 13:1735–1747

Flagel L, Udall J, Nettleton D, Wendel J (2008) Duplicate gene expression in allopolyploid Gossypium reveals two temporally distinct phases of expression evolution. BMC Biol 6:16

Flagel LE, Chen LP, Chaudhary B, Wendel JF (2009) Coordinated and fine-scale control of homoeologous gene expression in allotetraploid cotton. J Hered 100:487–490

Garsmeur O, Charron C, Bocs S, Jouffe V, Samain S, Couloux A, Droc G, Zini C, Glaszmann JC, Van Sluys MA, D'Hont A (2011) High homologous gene conservation despite extreme autopolyploid redundancy in sugarcane. New Phytol 189:629–642

Glaszmann JC, Fautret A, Noyer JL, Feldmann P, Lanaud C (1989) Biochemical genetic-markers in sugarcane. Theor Appl Genet 78:537–543

Grivet L, Glaszmann JC, Arruda P (2001) Sequence polymorphism from EST data in sugarcane: a fine analysis of 6-phosphogluconate dehydrogenase genes. Genet Mol Biol 24:161–167

Grivet L, Glaszmann JC, Vincentz M, da Silva F, Arruda P (2003) ESTs as a source for sequence polymorphism discovery in sugarcane: example of the Adh genes. Theor Appl Genet 106:190–197

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM (2004) Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression. Plant Mol Biol Rep 22:325–337

Jackson S, Chen ZJ (2010) Genomic and expression plasticity of polyploidy. Curr Opin Plant Biol 13:153–159

Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinforma 7:474

Levy AA, Kashkush K, Feldman M (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics 160:1651–1659

Marconi TG, Costa EA, Miranda HR, Mancini MC, Cardoso-Silva CB, Oliveira KM, Pinto LR, Mollinari M, Garcia AA, Souza AP (2011) Functional markers for gene mapping and genetic diversity studies in sugarcane. BMC Res Notes 4:264

Moyle R, Birch RG (2013) Sugarcane *Loading Stem Gene* promoters drive transgene expression preferentially in the stem. Plant Mol Biol. doi:10.1007/s11103-013-0034-3

Mudge SR, Osabe K, Casu RE, Bonnett GD, Manners JM, Birch RG (2009) Efficient silencing of reporter transgenes coupled to known functional promoters in sugarcane, a highly polyploid crop species. Planta 229:549–558

Mudge SR, Basnayake SW, Moyle RL, Osabe K, Graham MW, Morgan TE, Birch RG (2013) Mature-stem expression of a silencing-resistant sucrose isomerase gene drives isomaltulose accumulation to high levels in sugarcane. Plant Biotechnol J. doi:10.1111/pbi.12038

OECD-FAO (2010) OECD-FAO agricultural outlook 2010. OECD Publishing, Paris

Osabe K, Mudge SR, Graham MW, Birch RG (2009) RNAi mediated down-regulation of *PDS* gene expression in sugarcane (*Saccharum*), a highly polyploid crop. Trop Plant Biol 2:143–148

Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, Comai L, Madlung A, Doerge RW, Colot V, Martienssen RA (2003) Understanding mechanisms of novel gene expression in polyploids. Trends Genet 19:141–147

Osborn TC, Udall JA, Quijada PA (2005) Detection of chromosomal rearrangements derived from homeologous recombination in four mapping populations of *Brassica napus* L. Genetics 169:967–979

Piperidis G, Piperidis N, D'Hont A (2010) Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. Mol Genet Genom 284:65–73

Pires JC, Gaeta RT, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403–3417

Soltis DE, Soltis PS, Tate JA (2004) Advances in the study of polyploidy since plant speciation. New Phytol 161:173–191

Souza GM, Berges H, Bocs S, Casu R, D'Hont A, Ferreira JE, Henry R, Ming R, Potier B, Van Sluys M-A, Vincentz M, Paterson AH (2011) The sugarcane genome challenge: strategies for sequencing a highly complex genome. Trop Plant Biol 4:145–156

Van Dillewijn C (1952) Botany of sugarcane. Chronica Botanica Co, Waltham

Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. Plant Biotechnol J 8:263–276

Wendel JF, Adams KL, Cronn R, Percifield R (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc Natl Acad Sci USA 100:4649–4654

Wendel JF, Hovav R, Udall JA, Chaudhary B, Rapp R, Flagel L (2008) Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. Proc Natl Acad Sci USA 105:6191–6195

Wu LG, Birch RG (2007) Doubled sugar content in sugarcane plants modified to produce a sucrose isomer. Plant Biotechnol J 5:109–117