Genome **Biology**

REVIEW

Biofuel and energy crops: high-yield Saccharinae take center stage in the post-genomics era

Savio de Siqueira Ferreira[1], Milton Yutaka Nishiyama Jr[1], Andrew H Paterson[2] and Glaucia Mendes Souza*[1]

Abstract

The Saccharinae, especially sugarcane, *Miscanthus* and sorghum, present remarkable characteristics for bioenergy production. Biotechnology of these plants will be important for a sustainable feedstock supply. Herein, we review knowledge useful for their improvement and synergies gained by their parallel study.

Biofuels and sustainability

The replacement of fossil fuels by biofuels is an ongoing effort in many countries. With decreasing oil reserves and increasing fossil fuel prices, bioenergy is a promising alternative. Advantages of biofuels can include a positive energy balance, reduction of greenhouse gas emissions and indirect effects, such as rural development. Studies based on life-cycle analysis conclude that when ethanol from sugarcane is used to replace fossil fuels in transportation, a substantial reduction in net greenhouse gas emissions may result (from 80% to greater than 100% savings [1]). Biomass can also be used to generate electricity, with electric vehicles presenting several advantages over combustion engines. Wood, cellulose and biofuel generation of electricity and stationary generation of energy can be very efficient and are also being implemented as options. In the last 5 years we have seen a 109% increase in global biofuel production. World projections provided by the Organisation for Economic Co-operation and Development (OECD)/Food and Agriculture Organization (FAO) indicate further increases in bioethanol and biodiesel production from the present 140 billion liters to 221 billion liters in 2021, corresponding to an additional 60% increase.

Increased biofuel production, and the associated increase in production of energy feedstocks, raises sustainability concerns over issues such as changes in land use, competition between energy crops and food and feed crops, and impacts on ecosystem services, including soil and water resources. Mandates in several countries to substitute gasoline for bioethanol require a substantial contribution from advanced fuels (sugar-derived and/or lignocellulosic bioethanol) to guarantee a reduction of greenhouse gas emissions.

Which plants are best suited to the requirements of future bioenergy feedstocks? To produce energy from plant-fixed C-bonds, crops should be high yielding, fast growing, with C-bonds that are easy to convert to useful forms, and require relatively small energy inputs for growth and harvest. To achieve sustainability, energy crops should not require extensive use of prime agricultural lands and they should have a low cost of energy production from biomass. Both the realities of agriculture in environments that are always heterogeneous and energy security require that feedstocks include a portfolio of diverse crops rather than merely a single crop.

A strong case can be made that members of the Saccharinae subtribe, particularly *Saccharum* (sugarcane and energy cane), *Miscanthus* and *Sorghum* species (Figure 1), best encompass these requirements. For commercial markets to develop, these crops are being evaluated with respect to their productivity as perennial crops (ratoon) in short growing seasons under different conditions, such as periodic drought, low temperatures and low nutrient inputs [2]. A recent development includes breeding efforts to produce an 'energy cane' (*Saccharum* species or interspecific hybrid) more amenable for hydrolysis of the bagasse and straw lignocellulosic fibers. The high yield of *Saccharum* (sugarcane) in tropical climates is particularly well documented, and *Miscanthus* and sorghum show similar promise in temperate climates. Herein, we review the merits of these grasses as a complementary package of bioenergy feedstock crops, the state of knowledge useful for their study and improvement, and synergies that might be gained by their parallel study.

The Saccharinae group: many energy crop traits combined

Within the Andropogoneae grasses, sugarcane (*Saccharum*) and *Miscanthus* belong to the Saccharinae subtribe.

*Correspondence: glmsouza@iq.usp.br
[1]Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-000 São Paulo, SP, Brazil
Full list of author information is available at the end of the article

BioMed Central

Figure 1. The Saccharinae plants. (a) Glaucia Souza's group collecting photosynthetic data from sugarcane plants in Brazil. **(b)** A sorghum field in Mali; all plants are over 3 m high.

Sorghum has been considered a member of the Sorghinae subtribe, although more recently a good case has been made (that we will accept herein) for expanding the Saccharinae to include the Sorghinae [3].

Sugarcane is a common name of a group of predominantly tropical species that originated in Southeast Asia (Table 1). Modern varieties result from crosses of the sucrose-accumulating relative *Saccharum officinarum* and the wild relative *Saccharum spontaneum*, with contributions from *Saccharum robustum*, *Saccharum sinense*, *Saccharum barberi*, *Erianthus* and *Miscanthus* [4,5]. Commercial varieties have the remarkable capability of storing high sucrose levels in the stem that can reach 40% of dry weight [6]. In a study of sugarcane yields across the world, commercial maximum cane yield averaged 69 t ha^{-1} year^{-1} and the experimental maximum averaged 98 t ha^{-1} year^{-1} in the countries with the highest sunlight [7]. Today, commercial yields closer to the experimental maximum are frequently reported. Sugarcane average annual production per hectare (39 t ha^{-1} of dry

stalks and trash) compares favorably with other high-yield bioenergy crops such as *Miscanthus* (29.6 t ha^{-1}) and switchgrass (10.4 t ha^{-1}) [8] (Table 2). Estimates from field trials show an average yield of 22.8 and 12.2 t ha^{-1} for sugarcane ancestral species *S. spontaneum* and *S. officinarum*, respectively [9].

Complementing the tropical adaptation of *Saccharum*, with most species native to eastern or southeastern Asia [10], its adaptability to continental Europe [11-13] shows the feasibility of producing *Miscanthus* in temperate latitudes (Table 1). *Miscanthus × giganteus*, a sterile, vegetatively propagated hybrid (2n = 3x = 57) believed to originate from crosses between tetraploid *Miscanthus sacchariflorus* and diploid *Miscanthus sinensis* [14], generally produces high yields, similar to (and in some cases better than) other biomass crops [8,15]. Considerable leveraging of breeding, production and processing infrastructure might be gained by the close relationship of *Miscanthus* to *Saccharum* - thought to be the closest relatives of one another, and polyphyletic [16].

Table 1. Origin, production and method of bioenergy production of the Saccharinae plants

	Miscanthus	*Saccharum*	*Sorghum*
Center of diversity	South to temperate east Asia	South Asia	Africa
Production	Limited: Asia, Europe, USA	Extensive: Brazil, India, China, others	Extensive: USA, Africa, central
America, south Asia			
Ethanol production	Cellulosic	Sugar; cellulosic	Starch (grain); sugar; cellulosic
Electricity production	Biomass burning	Biomass burning	Biomass burning

Table 2. Summary of the average biomass yield of the Saccharinae plants

Crop	Dry biomass (t ha^{-1})	Reference(s)
Sugarcane commercial hybrids (*Saccharum*)	39	[2,7]
Sorghum	15-25	[111]
Miscanthus	29.6	[8]

Saccharum × *Miscanthus* hybrids ('Miscanes') have been used for sugarcane improvement [17-19], and also show promise as a highly productive cellulosic biomass crop.
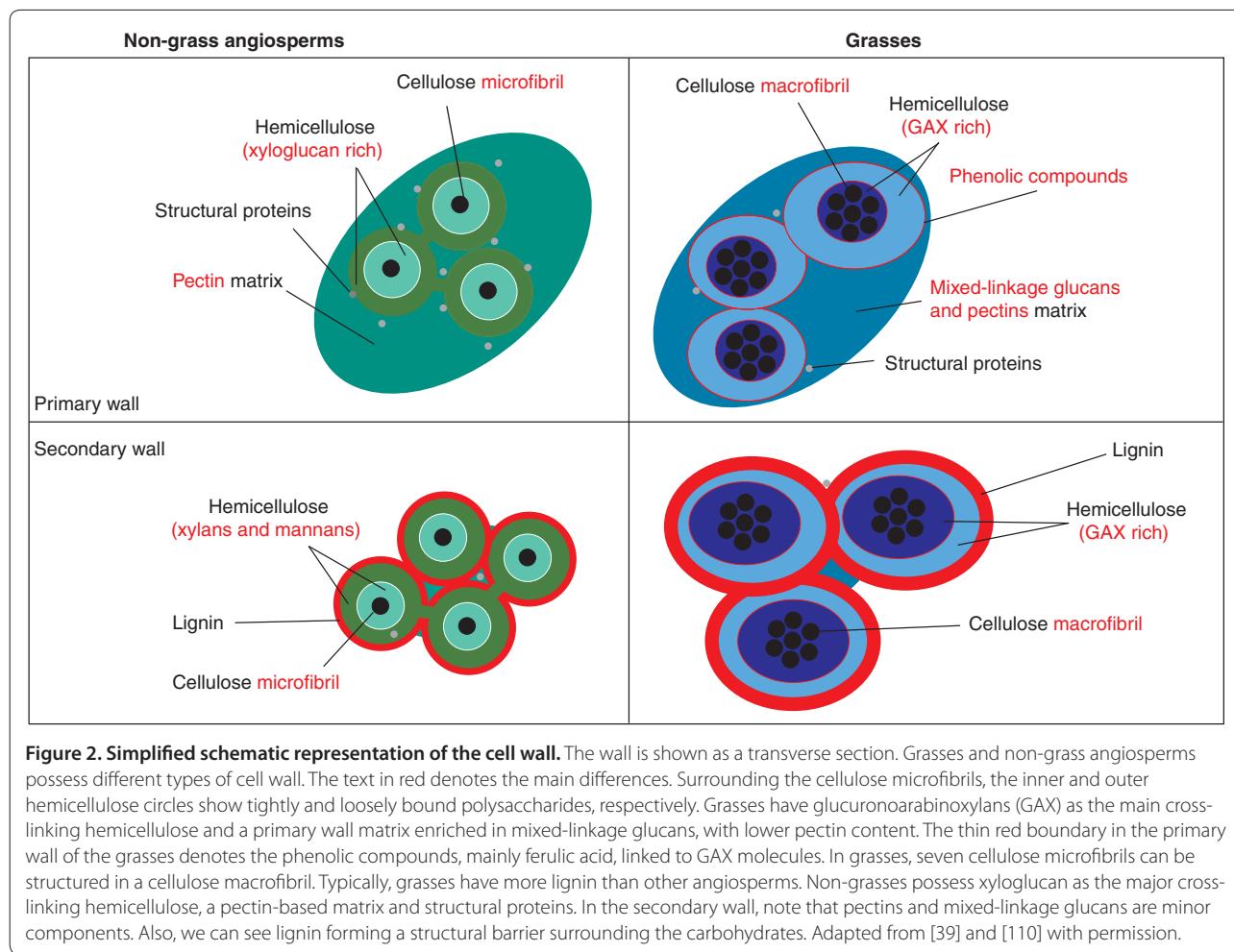
Increased demand for limited fresh water, along with rising global temperatures and aridity, suggest that sustainable future biomass production will have to occur using little or no irrigation, highlighting an important role of sorghum in a portfolio of bioenergy crops. One of the most drought tolerant of cereal crops thanks to its origins in Sudan and Ethiopia [20], the multifaceted history of sorghum improvement offers a wider range of genetic variations than found in many crops, exemplified by the fact that sorghum is one of the few crops suited to all proposed approaches for renewable fuel production (such as from starch, sugar, and/or cellulose; Table 1). About 30% of the US sorghum crop is presently used as feedstock in the grain-to-ethanol process, which has also been commercialized in India and China. The completely sequenced genome of sorghum, which has the further advantages of being relatively small and with minimal gene duplication [21], together with transformation potential, knowledge of cell wall composition and architecture and other features ([22] and references therein), make sorghum an important model for research concerning bioenergy grasses [22,23].

Plants in the Andropogoneae use C4 photosynthesis (Box 1), which avoids photorespiration, leading to higher maximal photosynthetic energy conversion efficiency than the C3 pathway used by rice, wheat and many other grasses [5,24], resulting in more biomass accumulation. In elevated CO_2 conditions, the C4 grasses sugarcane [25], maize and sorghum [26] show better responses to drought stress than C3 grasses. Plants in the Saccharinae have some further advantages in comparison with other C4 grasses, such as maize. First, many routinely produce a 'ratoon' crop, regrowing after harvest and thus

Box 1. C4 photosynthesis

Many of the most productive agricultural crops use the C4 photosynthetic pathway to increase net carbon assimilation at high temperature (Figure 3, adapted from [97]). Discovered in sugarcane [98], C4 photosynthesis may have been an adaptation to hot, dry environments or CO_2 deficiency [99-102], and appears to have evolved repeatedly from ancestors that used C3 photosynthesis [103,104], including multiple origins within some angiosperm families [105,106]. Most C4 plants are grasses, including the entire Andropogoneae tribe (including sorghum, sugarcane and *Miscanthus*), and it has been inferred that C4 photosynthesis first arose in grasses during the Oligocene epoch (24 to 35 million years ago) [107,108]. The high photosynthetic capacity of C4 plants is achieved by CO_2 assimilation in mesophyll cells (by phosphoenolpyruvate carboxylase together with carbonic anhydrase to facilitate rapid equilibrium between CO_2 and HCO_3^-) then diffusion of the resulting C4 acids into bundle sheath cells, where CO_2 is discharged by various decarboxylases at up to 10-fold higher than atmospheric level at the site of ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco), the primary enzyme of C3 photosynthesis. This high CO_2 concentration mitigates wasteful fixation of oxygen by Rubisco, reducing photorespiration, or CO_2 loss during C3 photosynthesis, at high temperatures [109]. C4 plants are classified in part based on the type of decarboxylases used in the bundle sheath: NADP malic enzyme, NAD malic enzyme or phosphoenolpyruvate carboxykinase.

eliminating the need for replanting each year. Indeed, the *Sorghum* genus, with annual and perennial species that are genetically compatible, has become a botanical model for study of attributes related to perenniality [27-29]. Second, sugarcane and *Miscanthus* have lower nitrogen-input requirements [13,30], and the latter can relocate some nutrients from aerial parts to the roots and/or rhizomes at the end of the growing season [31]. Third, some reports show better photosynthetic features of Saccharinae plants than other Andropogoneae. Light interception by the leaves is higher in *Miscanthus* than in maize [15] and *Miscanthus* can sustain higher levels of CO_2 assimilation than maize in lower temperatures [32]. Sugarcane photosynthesis is enhanced in elevated CO_2 in open-top chambers, increasing biomass productivity [33], which does not occur in maize grown in open-air elevation of CO_2 [34]. However, this finding is

Figure 2. Simplified schematic representation of the cell wall. The wall is shown as a transverse section. Grasses and non-grass angiosperms possess different types of cell wall. The text in red denotes the main differences. Surrounding the cellulose microfibrils, the inner and outer hemicellulose circles show tightly and loosely bound polysaccharides, respectively. Grasses have glucuronoarabinoxylans (GAX) as the main cross-linking hemicellulose and a primary wall matrix enriched in mixed-linkage glucans, with lower pectin content. The thin red boundary in the primary wall of the grasses denotes the phenolic compounds, mainly ferulic acid, linked to GAX molecules. In grasses, seven cellulose microfibrils can be structured in a cellulose macrofibril. Typically, grasses have more lignin than other angiosperms. Non-grasses possess xyloglucan as the major cross-linking hemicellulose, a pectin-based matrix and structural proteins. In the secondary wall, note that pectins and mixed-linkage glucans are minor components. Also, we can see lignin forming a structural barrier surrounding the carbohydrates. Adapted from [39] and [110] with permission.

controversial since enclosure and open-air studies give different results for the same crop, and some authors argue that enclosed studies are not the best scenario to mimic future increases in CO_2 concentration [35]. Moreover, experiments with *Miscanthus* in ambient and open-air elevation of CO_2 show no differences in yield [36].
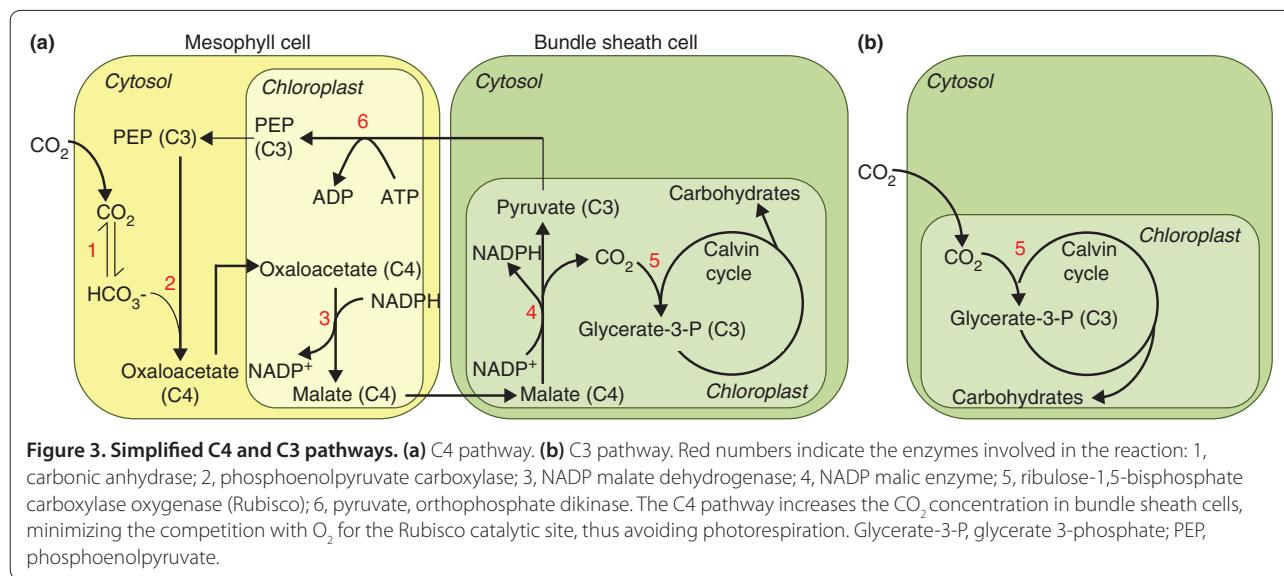
Since lignocellulosic biofuels use the plant cell wall as a source for fermentable sugars, it is important to understand the composition and architecture of the cell wall to develop strategies to degrade it efficiently. Grasses present a particular cell wall structure and composition (Figure 2), making a 'type II' cell wall that differs substantially from the 'type I' cell walls of other feedstocks, such as wood species [22,37,38]. This also implies the evolution of different gene families involved in the synthesis of the cell wall [22]. Recently, a model for sugarcane cell wall architecture and for hierarchical enzymatic hydrolysis was proposed [39]. By understanding the structure of the cell wall, it is possible to choose the best method to improve hydrolysis yield, and design

breeding strategies or develop improved procedures to recover the released carbohydrates.

Genomics meets biotechnology for the improvement of Saccharinae biofuel grasses

Improvements in sorghum are characteristic of many other major food and feed crops, and *Miscanthus* improvement is just beginning; examining sugarcane improvement therefore exemplifies the methods and approaches likely to be employed in biofuel grasses.

Sugarcane improvement efforts follow both molecular-assisted breeding and transgenic routes [40]. *S. officinarum* is a plant with high sugar content in its stems but low productivity, and *S. spontaneum* has high tillering and biomass yield but low sugar accumulation. Modern sugarcane cultivars derive from a few crosses between *S. officinarum* and *S. spontaneum* and have been shown to be genetically very similar [41]. Breeding programs have been able to increase yield and sucrose content by crossing cultivars but gains are becoming slimmer. To continue the improvement of yield it may be necessary to

Figure 3. Simplified C4 and C3 pathways. (a) C4 pathway. **(b)** C3 pathway. Red numbers indicate the enzymes involved in the reaction: 1, carbonic anhydrase; 2, phosphoenolpyruvate carboxylase; 3, NADP malate dehydrogenase; 4, NADP malic enzyme; 5, ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco); 6, pyruvate, orthophosphate dikinase. The C4 pathway increases the CO_2 concentration in bundle sheath cells, minimizing the competition with O_2 for the Rubisco catalytic site, thus avoiding photorespiration. Glycerate-3-P, glycerate 3-phosphate; PEP, phosphoenolpyruvate.

turn back to ancestral genotypes and broaden the genetic basis of crosses. *S. spontaneum* and *S. robustum* are also being used as parents, with the goal of designing a crop more amenable for cellulosic biofuel production, with increased stress tolerance and increased yield but less emphasis on stalk sugar concentration, the so-called 'energy cane'. World collections of *Saccharum* germplasm are held in Florida [42] and India [43], which keep ancestral genotypes and cultivars, and many private collections are also kept and used for crosses in specific breeding programs. Each world collection has over 1,500 accessions of ancestral genotypes, most of them *S. officinarum* (about 750), *S. spontaneum* (about 600) and *S. robustum* (about 130), and 500 to 1,000 hybrids or cultivars. Sorghum, like sugarcane, has large germplasm collections held by the US National Plant Germplasm System and at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT, the CGIAR center with a sorghum improvement mandate). Only a few small *Miscanthus* collections are held publicly, but several private collections associated with breeding programs are similar in size to the *Saccharum* collections.

Crosses between members of the Saccharinae are viable. In fact, sugarcane has been crossed to both *Miscanthus* and sorghum, generating viable progenies, and the strategy has been used to incorporate cold and drought resistance traits from *Miscanthus* into sugarcane [19].

The transformation of sugarcane is becoming an interesting and growing field. Methods for transformation are already established with efforts aimed mostly at sugar yield and quality [44-46], disease resistance [47,48], and the use of sugarcane as a biofactory to produce high-value bioproducts [49,50]. For biofuel production, some approaches show interesting results, with lower biomass recalcitrance [51] and expression and accumulation of microbial cellulolytic enzymes in sugarcane leaves [52] to improve biomass hydrolysis. The most widely used promoters are the constitutive *CaMV 35S* and maize *ubi1*, but sugarcane promoters have already been used or characterized, including tissue-specific [46,47] and responsive promoters [53]. However, sugarcane transformation is not a trivial task since problems such as transgene silencing frequently occur ([40,54] and references therein). Sorghum transformation is also routine (although at lower efficiency than in some crops [55]), and *Miscanthus* transformation methods have been established [56].

Advantages of a reference genome
For both molecular-assisted and transgenic strategies outlined above, the availability of a reference genome sequence is highly desirable, as well as the definition of the complete complement of genes and proteins. For the Saccharinae, the relatively small (740 Mb) and diploid genome of sorghum, which has not experienced genome duplication in about 70 million years [21], has become the best reference for genomics and transcriptomics in sugarcane [57]. Nonetheless, the sugarcane genome itself is being sequenced using a combination of approaches. In a first phase, researchers are sequencing bacterial artificial chromosomes (BACs) combined with whole-genome shot-gun sequencing to produce a reference genome [58]. Currently, three sugarcane BAC libraries are available; from variety R570 [59], selfed progenies of R570 [60] and SP80-3280 [61]. The two former libraries have 103,000 to 110,000 clones comprising about 12 times coverage of the basic genome complement but only about 1.3 to 1.4 times coverage of the individual alleles. The latter library has

about 36,000 clones, and all three have inserts of about 125 to 130 kb. BAC sequencing has enabled research to deduce synteny and collinearity of much of the sugarcane genome with other grasses, especially sorghum, particularly in genic regions [61-63]. Unaligned regions between sorghum and sugarcane genomes are largely repetitive [62], enriched in transposon-related sequences [61,63]. Consistent with several genetic mapping efforts, the sequencing of BAC clones revealed high levels of gene structure/sequence conservation and collinearity among hom(oe)ologous haplotypes of the sugarcane genome [64], and several putative sugarcane-specific genes/sequences [61-63]. Groups from Australia, Brazil, France, South Africa and the USA are advancing these efforts in genome sequencing, increasing the number of BACs sequenced and producing shot-gun data of several cultivars. It is expected that reference genome sequences will be made available for both cultivars and ancestral genotypes [65] and, to that end, researchers are developing statistical models using SNPs where homology groups with any ploidy level may be estimated [66]. This will be essential to obtain a saturated genetic map of the sugarcane genome that may aid genome assembly. The greatest challenge that distinguishes the sequencing of *Saccharum* and *Miscanthus* from the more tractable genomes of sorghum and other cereal models is large physical size (approximately 10 Gb) and large copy numbers of even 'low-copy' elements (8 to 12 in sugarcane; 4 to 6 in *Miscanthus*). During assembly of such genomes, many closely related alleles 'collapse' into single gene/element models that fail to capture allelic and perhaps also paralogous diversity within even a single genotype. The sorghum genome will greatly help in the assembly, but around 20% of the sugarcane expression sequence tags (ESTs) from the SUCEST project [67] appear to be specific to sugarcane, since they do not match sorghum, *Miscanthus*, maize, rice or *Brachypodium* [68], requiring other strategies in the assembly. Linkage maps based on molecular markers have shown synteny and collinearity of sorghum and sugarcane genomes, but are complicated to make in sugarcane due to the polyploidy and absence of inbred lines ([69] and references therein). This problem was partly overcome with the use of single-dose markers [70], which segregate in a 1:1 ratio in the gametes of a heterozygous genotype, and account for approximately 70% of polymorphic loci in sugarcane [71]. However, among 20 to 30 linkage maps based on a few thousand markers available for sugarcane ([71,72] and references therein), it remains true that only 33% to 60% of the sugarcane genome is represented on these maps [71]. A recent development that may help breeders in marker-assisted selection efforts has been the development of an algorithm and software (ONEMAP) for constructing linkage maps of outcrossing plant species that has been successfully applied to sugarcane [73]. Enriched mapping of DNA polymorphisms that also provide for deconvolution of closely related sequences may also aid in assembly of such highly polyploid genomes.

Saccharinae transcriptomics

Changes in gene expression associated with allopolyploidy are well known, but sugarcane functional genomics is a challenge due the complexity of its largely autopolyploid and aneuploid genome and the absence of a reference sequence. Again, the sorghum genome has been serving as a reference to define putative transcripts. The sorghum transcriptome has been studied by different high-throughput technologies such as cDNA microarrays and massively parallel sequencing (Tables 3 and 4) to understand the expression profiling and biological function of genes in response to herbivory, biotic and abiotic stress in different tissues and treatments [68], and how the genes and their structural/functional changes contribute to the morphological variations between sorghum lines integrating genome evolution and expression divergence [74]. Deep RNA sequencing methods have overcome many limitations of microarray technologies and have allowed recent studies to reveal sorghum genes, gene networks, and a strong interplay among various metabolic pathways in different treatments [75], as well as the identification of particular paralogs that putatively encode enzymes involved in specific metabolic networks [76].

Despite the absence of a sequenced genome and the complexities associated with the presence of about 8 to 12 copies of each gene, functional genomics has made considerable progress towards understanding unique biological attributes of sugarcane. These studies assist in the development of new applications for bioenergy, biomaterial industries and improved 'energy' cultivars [57]. The fundamental databases and resources for studies of functional genomics in sugarcane have been reviewed recently [57,77,78] and a sugarcane computational environment (SUCEST-FUN Database) has been developed for storage, retrieval and integration of genome sequencing, transcriptome, expression profiling, gene catalogs, physiology measures and transgenic plant data [79]. Studies on sugarcane gene expression have been based mainly on EST information from different tissues, treatments and genotypes. The largest contribution to the available ESTs (>80%) comes from the SUCEST project [67], and most of the remainder comes from Australia, USA, South Africa and India (reviewed by [57,68]). To obtain a less redundant dataset including ESTs not sampled by the SUCEST project, a comparison with SoGI [80] was carried out and 8,106 sequences lacking detectable similarity to SAS (sugarcane assembled sequences)

Table 3. Summary of the microarray data available for sugarcane and *Sorghum*

Platform	Record GEO/NCBI	Features	No. of samples	Sample description	Variety	Reference
Sugarcane ESTs nylon arrays I	GPL210	1,536	12	Plantlet samples exposed to cold for up to 48 h	*Saccharum* hybrid cv. SP80-3280	[112]
SUCAST Sugarcane 2208 v1.0 (cDNA microarray)	GPL3799	2,208	16	Roots from six plants harvested 6, 12, 24 and 48 h after exposure to phosphate starvation	*Saccharum* hybrid cv. SP80-3280	[113]
			8	Response to herbivory by *Diatraea saccharalis*		
			28	Hormone treatment		
			8	Response to N_2-fixing endophytic bacteria association	*Saccharum* hybrid cv. SP70-1143	
			12	Leaves collected 24, 72 and 120 h after exposure to drought conditions for the control and experimental groups	*Saccharum* hybrid cv. SP90-1638	
SUCAST 1920v3 (cDNA microarray)	GPL1376	1,920	26	The abundance of transcripts among six different sugarcane tissues (flowers, roots, leaves, lateral buds, 1st and 4th internodes)	*Saccharum* hybrid cv. SP80-3280	[114]
			4	Leaves from sugarcane population segregated in relation to high and low brix		
Affymetrix Sugarcane Genome Array	GPL3844	6,024	12	Leaves of sugarcane in response to sugar accumulation in meristem, internodes 1 to 3, internode 8, internode 20	*Saccharum* hybrid cv. Q117	[115]
			8	Cold-girdles were attached to sugarcane leaves (*n* = 4) for a period of 56 h prior to harvest	*Saccharum* hybrid cv. N19	[116]
SUCEST-FUN Agilent 44k (oligo array)	GPL14862	21,902	6	Leaves were collected 24, 72 and 120 h after exposure to drought conditions for the control and experimental groups	*Saccharum* hybrid cv. SP90-1638	[86]
cDNA microarray	http://fungen.org/ Sorghum.htm	12,982	118	Seedlings exposed to high salinity, osmotic stress, and ABA for 3 and 27 h	Seedlings of *Sorghum bicolor* L. Moench cv BTx623	[117]
		12,982	>102	Seedlings exposed to salicylic acid, methyl jasmonate, and the ethylene precursor aminocyclopropane for 3 and 27 h	Seedlings of *S. bicolor* L. Moench cv BTx623	[118]
cDNA glass slide	Not available	672	>6	Herbivory with *Schizaphis graminum*	Seedlings of *S. bicolor* ATx399 _ RTx430	[119]
		3,508	>4	Herbivory with *Schizaphis graminum*	Seedlings of *S. bicolor* M627 and Tx7000	[120]
Sorghum Agilent 44k (oligo array)	GPL15369	41,977	12	Leaves were collected 24, 72 and 120 h after exposure to drought conditions for the control and experimental groups	Grain of *S. bicolor* L. sweet Keller cv BTx623	[74]

ABA, abscisic acid; GEO, Gene Expression Omnibus; NCBI, National Center for Biotechnology Information.

were identified. The clustering strategy in SoGI produces redundant clusters and makes the SUCEST assembly more appropriate for gene and orthology-based analysis [81]. The SUCEST-FUN project and SAS sequences have been updated with the whole sugarcane ESTs from the National Center for Biotechnology Information (NCBI) and compared with the SoGI assembly (Table 5). A total of 282,683 ESTs are currently catalogued in the SUCEST-FUN Database. Comparison of ESTs from sorghum with sugarcane, maize and rice has revealed mean sequence identities of 97%, 93% and 86%, respectively, indicating a close relationship between sorghum and sugarcane

Table 4. Summary of the high-throughput sequencing data available for sugarcane, *Sorghum* and *Miscanthus*

Platform	Record GEO/NCBI, SRA	No. of samples	Sample description	Variety	Reference
Illumina Genome Analyzer IIx	GPL16316	8	sRNA and non-coding RNAs of sugarcane in response to drought stress	*Saccharum* hybrid (mixed)	[92]
Illumina Genome Analyzer IIx	GPL15137	3	Examination of small RNAs in the sugarcane leaves that match to sugarcane LTR-RTs	*Saccharum* hybrid cv. SP80-3280	[93]
			Identification of small RNAs associated with auxiliary bud outgrowth	*Saccharum* hybrid cv. SP80-3280	[121]
Illumina Genome Analyzer IIx	GPL16317	6	Screening of small RNA transcriptome of sugarcane plants infected with *Acidovorax avenae* subsp. avenae after 7 days	*Saccharum* hybrid cv. SP70-1143	[92]
			Identification of sugarcane microRNAs after exposure to pathogens and microorganisms	*Saccharum* hybrid cv. SP70-1143	
SOLEXA technology	Not available	8	Identification of microRNA differentially expressed under drought stress, correlation between two cultivars with different drought tolerance	*Saccharum* hybrid cv. RB867515, RB855536	[90]
SOLEXA technology	Not available	2	Increase the understanding of the role of microRNAs in the complex regulation of drought stress in field-grown sugarcane	*Saccharum* hybrid cv. RB867515, RB855536	[91]
Illumina Genome Analyzer II	GPL13779	24	RNA-Seq of *Sorghum bicolor* 9-day seedlings in response to osmotic stress and abscisic acid	*Sorghum bicolor*	[75]
Illumina Genome Analyzer IIx	GPL14633	3	Comparative sequence analysis of plant small RNAs	*S. bicolor*	GEO
			Comparative sequence analysis of plant small RNAs from leaves, flowers and panicle tissues	*S. bicolor*	GEO
Illumina Genome Analyzer II	GPL9983	3	Develop sequence resources of small RNAs from different *Miscanthus × giganteus* tissues (including leaves, flowers and rhizomes)	*Miscanthus × giganteus*	[94]
Illumina Genome Analyzer II	SRX131848/ SRX131845	1	RNAseq from two *Miscanthus sinensis* to define single nucleotide variants and find simple sequence repeat markers for understanding genomic responses to tetraploidy and chromosome fusion	*Miscanthus sinensis*	[96]

Data were catalogued based on MIAME compliant public repositories such as Gene Expression Omnibus/Center for Information Biology Gene (GEO/NCBI) Expression Database (CIBEX), Microarray Gene Expression Data Society (MGED) and ArrayExpress. LTR-RT, long terminal repeat retrotransposon.

Table 5. ESTs corresponding to SUCEST, SoGI, JGI, GenBank (as of 4 March 2013)

	Saccharum			*Sorghum*	
	SUCEST	GenBank	SoGI	JGI	GenBank
Number of ESTs	238,208	283,677	282,683	227,154	232,681
Number of ESTs not in SUCEST	-	45,469	-	-	5,527
Number of clusters	26,303	-	42,377	34,496	-
Number of singletons	16,838	-	78,924	-	-
Number of predicted proteins	39,021	-	-	34,496	-

EST, expression sequence tag; JGI, Joint Genome Institute.

(Figure S7 of [21]). A total of 39,021 sugarcane proteins were predicted from 43,141 clusters [67] using ESTScan [82] and the *Oryza sativa* matrix (Table 5). Putative orthologs and paralogs were identified by pairwise proteome comparisons with InParanoid software [83]. With the aid of MultiParanoid software [84], we found orthology relationships among multiple proteomes (Table 6). The analysis encompassed a comparison among five species: *Saccharum* sp., *Sorghum bicolor*, *Zea mays*, *O. sativa* and *Arabidopsis thaliana*. Proteins were grouped into 18,611 orthologous clusters. A total of 16,723 sugarcane proteins found 13,804 orthologs in sorghum, 22,312

Table 6. Number of putative ortholog/paralog genes between sugarcane, other grasses and *Arabidopsis*

Species	No. of proteins	No. of cane orthologs (pairwise species)	No. of orthologs (multiple species)
Saccharum	39,021	-	17,176
Sorghum	36,338	13,804	19,414
Maize	10,6046	22,312	45,237
Rice	51,258	16,913	29,888
Arabidopsis	35,386	13,998	22,165

in maize, 16,913 in rice and 13,998 in *Arabidopsis*, with a confidence score ≥ 0.05 and group merging cut-off >0.5 using the BLOSUM80 matrix, suitable for closely related sequences (Table 6).

The sugarcane transcriptome has been studied using technologies, including cDNA macroarrays (nylon membranes), cDNA microarrays spotted onto glass slides, and oligonucleotide arrays either spotted or synthesized *in situ*. A summary of the available platforms, samples and related works for sugarcane and sorghum using array technologies is shown in Table 3 and has been reviewed recently [57,68,78,85]. Sugarcane transcriptomics has identified genes associated with sucrose content, biotic and abiotic stresses, photosynthesis, carbon partitioning and roles of phytohormones and signaling pathways in adaptive responses. These studies also allowed for the identification of promoters that can be used to drive transgene components in a tissue-specific or controlled manner. Several other methods to study sugarcane expression profiles at a moderate scale have been used to confirm the expression patterns observed in large-scale transcript studies [57].

More recently, the use of oligoarrays has included studies on the regulation of antisense gene expression in sugarcane, pointing to a role for these transcripts in drought responses [86]. Some years ago, serial analysis of gene expression (SAGE) in sugarcane revealed an unexpectedly high proportion of antisense transcripts and chimeric SAGE [87]. High-throughput sequencing (Table 4) is useful for assessing transcriptomes, providing detailed information for transcript variants, particularly SNPs, assessment of the expression of hom(oe)ologous alleles in the polyploid genome, spliced isoforms and so on [88]. Using this strategy, some sugarcane genes were characterized for SNP density and gene haplotypes across varieties [89]. In recent studies, it has become apparent that small RNAs, particularly microRNAs, have important regulatory roles in sugarcane, playing a key role in development and responses to biotic and abiotic stresses [90-92]. Evidence suggests that long terminal repeat retrotransposon (LTR-RT) families may affect nearby genes by generating a diverse set of small RNAs that trigger gene-silencing mechanisms [93].

In contrast to sorghum and sugarcane, genomic and transcriptomic studies on *Miscanthus* are just beginning.

The recent high-throughput sequencing of its genome and transcriptome identified the presence of repeats that are actively producing small RNAs [94], and the construction of a genetic map identified informative simple sequence repeats in sugarcane and a genome-wide duplication in *Miscanthus* relative to *S. bicolor* [95]. These studies will increase the understanding of complex genomes [96].

Conclusions

The Saccharinae grasses sugarcane, *Miscanthus* and sorghum are promising and complementary elements of a portfolio of bioenergy feedstocks. As sustainability criteria take dominant roles in the commercialization of biomass sources, these plants are likely to contribute to provide cheap, reliable and politically viable options for bioenergy production. Biotechnology for these crops is less advanced than in food crops such as maize and rice, but it is progressing quickly. Many efforts are underway to define genes associated with traits of interest such as sucrose content, drought tolerance, yield and adaptation to climate changes, and much is known about genes and markers for the improvement of these crops. Breeding programs are improving germplasm collections and defining routes to speed up selection of progenies and choice of ideal parents for crossing. It is expected that prudent integration of conventional breeding methods with marker-assisted and transgenic options may increase the (currently slow) rates of yield improvement, decreasing the amount of land required for large-scale biofuel production, as well as the need for inputs such as water, herbicides and fertilizers to maintain economical levels of production. Finally, the transition to a more biobased economy may be expedited by the increased value of biobased chemicals that might be harvested from the production chain through the adoption of integrated biorefinery systems. Better understanding of and greater control over carbon partitioning in these plants may greatly increase the number of co-products, including bioethanol, biodiesel, biokerosene, bioplastics and bio-electricity to name a few.

Abbreviations
BAC, bacterial artificial chromosome; EST, expressed sequence tag; SAGE, serial analysis of gene expression; SAS, sugarcane assembled sequence; SNP, single nucleotide polymorphism.

Competing interests

Acknowledgements

Author details

[1]Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-000 São Paulo, SP, Brazil. [2]Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602, USA.

References

1. Macedo IC, Seabra JEA, Silva JEAR: **Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: The 2005/2006 averages and a prediction for 2020.** *Biomass Bioenerg* 2008, **32:**582-595.
2. Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM: **Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content.** *Plant Biotechnol J* 2010, **8:**263-276.
3. Kellogg EA: **Phylogenetic relationships of Saccharinae and Sorghinae.** In *Genomics of the Saccharinae. Volume 11.* Edited by Paterson AH. New York, Berlin: Springer; 2013:3-22.
4. Daniels J, Roach BT: **Taxonomy and evolution.** In *Sugarcane Improvement Through Breeding.* Edited by Heinz DJ. Amsterdam: Elsevier; 1987:7-84.
5. Ming R, Moore PH, Wu K, D'Hont A, Glaszmann JC, Tew TL, Mirkov E, Silva J, Jifon J, Rai M, Schnell RJ, Brumbley SM, Lakshmanan P, Comstock JC, Paterson AH: **Sugarcane improvement through breeding and biotechnology.** *Plant Breeding Rev* 2006, **27:**15-118.
6. Moore PM: **Temporal and spatial regulation of sucrose accumulation in the sugarcane stem.** *Aust J Plant Physiol* 1995, **22:**661-679.
7. Irvine JE: **Sugarcane [physiological characteristics, yield potential and climatic and environment factors].** In *Symposium on Potential Productivity of Field Crops Under Different Environments: 1983; Los Baños, Laguna (Philippines).* Los Baños: International Rice Research Institute; 1983:361-281.
8. Heaton EA, Dohleman FG, Long SP: **Meeting US biofuel goals with less land: the potential of Miscanthus.** *Glob Change Biol* 2008, **14:**2000-2014.
9. Roach BT: **Utilisation of** *Saccharum spontaneum* **in sugarcane breeding.** *Proc Int Soc Sugar Cane Technol* 1977, **16:**43-58.
10. Hodkinson TR, Renvoize SS, Chase MW: **Systematics of Miscanthus.** *Aspects Appl Biol* 1997, **49:**189-198.
11. Lewandowski I, Clifton-Brown JC, Scurlock JMO, Huisman W: **Miscanthus: European experience with a novel energy crop.** *Biomass Bioenerg* 2000, **19:**209-227.
12. Lewandowski I, Scurlock JMO, Lindvall E, Christou M: **The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe.** *Biomass Bioenerg* 2003, **25:**335-361.
13. Sacks EJ, Juvik JA, Lin Q, Stewart JR, Yamada T: **The gene pool of Miscanthus species and its improvement.** In *Genomics of the Saccharinae. Volume 11.* Edited by Paterson AH. New York, Berlin: Springer; 2013:73-104.
14. Hodkinson TR, Chase MW, Takahashi C, Leitch IJ, Bennett MD, Renvoize SA: **The use of dna sequencing (ITS and trnL-F), AFLP, and fluorescent in situ hybridization to study allopolyploid Miscanthus (Poaceae).** *Am J Bot* 2002, **89:**279-286.
15. Dohleman FG, Long SP: **More productive than maize in the Midwest: How does Miscanthus do it?** *Plant Physiol* 2009, **150:**2104-2115.
16. Hodkinson TR, Chase MW, Lledo MD, Salamin N, Renvoize SA: **Phylogenetics of Miscanthus, Saccharum and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences from ITS nuclear ribosomal DNA and plastid trnLintron and trnL-F intergenic spacers.** *J Plant Res* 2002, **115:**381-392.
17. Chen YH, Lo CC: **Disease resistance and sugar content in Saccharum-Miscanthus hybrids.** *Rep Taiwan Sugar Res Inst* 1989:1-7.
18. Sobral BWS, Braga DPV, LaHood ES, Keim P: **Phylogenetic analysis of chloroplast restriction enzyme site mutations in the Saccharinae Griseb. subtribe of the Andropogoneae Dumort. tribe.** *Theoret Appl Genetics* 1994, **87:**843-853.
19. Burner DM, Tew TL, Harvey JJ, Belesky DP: **Dry matter partitioning and quality of Miscanthus, Panicum, and Saccharum genotypes in Arkansas, USA.** *Biomass Bioenerg* 2009, **33:**610-619.
20. Kimber CT, Dahlberg JA, Kresovich S: **The gene pool of Sorghum bicolor and its improvement.** In *Genomics of the Saccharinae. Volume 11.* Edited by Paterson AH. New York, Berlin: Springer; 2013:23-42.
21. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, *et al.*: **The Sorghum bicolor genome and the diversification of grasses.** *Nature* 2009, **457:**551-556.
22. Carpita NC, McCann MC: **Maize and sorghum: genetic resources for bioenergy grasses.** *Trends Plant Sci* 2008, **13:**415-420.
23. Saballos A: **Development and utilization of sorghum as a bioenergy crop.** In *Genetic Improvement of Bioenergy Crops.* Edited by Vermerris W. New York: Springer; 2008:211-248.
24. Zhu XG, Long SP, Ort DR: **What is the maximum efficiency with which photosynthesis can convert solar energy into biomass?** *Curr Opin Biotechnol* 2008, **19:**153-159.
25. Vu JC, Allen LH Jr: **Growth at elevated CO(2) delays the adverse effects of drought stress on leaf photosynthesis of the C(4) sugarcane.** *J Plant Physiol* 2009, **166:**107-116.
26. Allen LH Jr, Kakani VG, Vu JC, Boote KJ: **Elevated CO_2 increases water use efficiency by sustaining photosynthesis of water-limited maize and sorghum.** *J Plant Physiol* 2011, **168:**1909-1918.
27. Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, Yang Y, McNally K, Khush GS, Paterson AH, Li ZK: **Convergent evolution of perenniality in rice and sorghum.** *Proc Natl Acad Sci U S A* 2003, **100:**4050-4054.
28. Kong W, Jin H, Franks CD, Kim C, Bandopadhyay R, Rana MK, Auckland SA, Goff VH, Rainville LK, Burow GB, Woodfin C, Burke JJ, Paterson AH: **Genetic analysis of recombinant inbred lines for Sorghum bicolor × Sorghum propinquum.** *G3 (Bethesda)* 2013, **3:**101-108.
29. Paterson AH, Schertz KF, Lin YR, Liu SC, Chang YL: **The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of Johnsongrass, Sorghum halepense (L) Pers.** *Proc Natl Acad Sci U S A* 1995, **92:**6127-6131.
30. Paterson AH, Moore PH, Tew TL: **The gene pool of Saccharum species and their improvement.** In *Genomics of the Saccharinae. Volume 11.* Edited by Paterson AH. New York, Berlin: Springer; 2013:43-72.
31. Somerville C, Youngs H, Taylor C, Davis SC, Long SP: **Feedstocks for lignocellulosic biofuels.** *Science* 2010, **329:**790-792.
32. Wang D, Portis AR Jr, Moose SP, Long SP: **Cool C4 photosynthesis: pyruvate Pi dikinase expression and activity corresponds to the exceptional cold tolerance of carbon assimilation in** *Miscanthus × giganteus.* *Plant Physiol* 2008, **148:**557-567.
33. De Souza AP, Gaspar M, Da Silva EA, Ulian EC, Waclawovsky AJ, Nishiyama MY Jr, Dos Santos RV, Teixeira MM, Souza GM, Buckeridge MS: **Elevated CO_2 increases photosynthesis, biomass and productivity, and modifies gene expression in sugarcane.** *Plant Cell Environ* 2008, **31:**1116-1127.
34. Leakey AD, Uribelarrea M, Ainsworth EA, Naidu SL, Rogers A, Ort DR, Long SP: **Photosynthesis, productivity, and yield of maize are not affected by open-air elevation of CO_2 concentration in the absence of drought.** *Plant Physiol* 2006, **140:**779-790.
35. Long SP, Ort DR: **More than taking the heat: crops and global change.** *Curr Opin Plant Biol* 2010, **13:**241-248.
36. Souza AP, Arundale RA, Dohleman FG, Long SP, Buckeridge MS: **Will the exceptional productivity of** *Miscanthus × giganteus* **increase further under rising atmospheric CO2?** *Agr Forest Meteorol* 2013, **171-172:**82-92.
37. Carpita NC: **Structure and biogenesis of the cell walls of grasses.** *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47:**445-476.
38. Sarkar P, Bosneaga E, Auer M: **Plant cell walls throughout evolution: towards a molecular understanding of their design principles.** *J Exp Bot* 2009, **60:**3615-3635.
39. Souza A, Leite DC, Pattathil S, Hahn M, Buckeridge M: **Composition and structure of sugarcane cell wall polysaccharides: implications for second-generation bioethanol production.** *Bioenerg Res* 2012:1-16.
40. Hotta CT, Lembke CG, Domingues D, Ochoa E, Cruz GQ, Melotto-Passarin D,

Marconi T, Santos M, Mollinari M, Margarido GA, Crivellari AC, dos Santos WD, de Souza AP, Hoshino AA, Carrer H, Souza AP, Garcia AAF, Buckeridge MS, Menossi M, Van Sluys MA, Souza GM: The biotechnology roadmap for sugarcane improvement. *Tropical Plant Biol* 2010, **3**:75-87.

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

42. USDA Germplasm Resources Information Network [http://www.ars-grin.gov/]

43. Sugarcane Breeding Institute [http://www.sugarcane.res.in]

44. van der Merwe MJ, Groenewald JH, Stitt M, Kossmann J, Botha FC: Downregulation of pyrophosphate: D-fructose-6-phosphate 1-phosphotransferase activity in sugarcane culms enhances sucrose accumulation due to elevated hexose-phosphate levels. *Planta* 2010, **231**:595-608.

45. Basnayake SW, Morgan TC, Wu L, Birch RG: Field performance of transgenic sugarcane expressing isomaltulose synthase. *Plant Biotechnol J* 2012, **10**:217-225.

46. Mudge SR, Basnayake SWV, Moyle RL, Osabe K, Graham MW, Morgan TE, Birch RG: Mature-stem expression of a silencing-resistant sucrose isomerase gene drives isomaltulose accumulation to high levels in sugarcane. *Plant Biotechnol J* 2013, **11**:502-509.

47. Khan MS, Ali S, Iqbal J: Developmental and photosynthetic regulation of delta-endotoxin reveals that engineered sugarcane conferring resistance to 'dead heart' contains no toxins in cane juice. *Mol Biol Rep* 2011, **38**:2359-2369.

48. Zhu YJ, McCafferty H, Osterman G, Lim S, Agbayani R, Lehrer A, Schenck S, Komor E: Genetic transformation with untranslatable coat protein gene of sugarcane yellow leaf virus reduces virus titers in sugarcane. *Transgenic Res* 2011, **20**:503-512.

49. Barros GOF, Ballen MAT, Woodard SL, Wilken LR, White SG, Damaj MB, Mirkov TE, Nikolov ZL: Recovery of bovine lysozyme from transgenic sugarcane stalks: extraction, membrane filtration, and purification. *Bioprocess Biosyst Eng*, in press.

50. Petrasovits LA, Zhao L, McQualter RB, Snell KD, Somleva MN, Patterson NA, Nielsen LK, Brumbley SM: Enhanced polyhydroxybutyrate production in transgenic sugarcane. *Plant Biotechnol J* 2012, **10**:569-578.

51. Jung JH, Fouad WM, Vermerris W, Gallo M, Altpeter F: RNAi suppression of lignin biosynthesis in sugarcane reduces recalcitrance for biofuel production from lignocellulosic biomass. *Plant Biotechnol J* 2012, **10**:1067-1076.

52. Harrison MD, Geijskes J, Coleman HD, Shand K, Kinkema M, Palupe A, Hassall R, Sainz M, Lloyd R, Miles S, Dale JL: Accumulation of recombinant cellobiohydrolase and endoglucanase in the leaves of mature transgenic sugar cane. *Plant Biotechnol J* 2011, **9**:884-896.

53. Prabu G, Prasad DT: Functional characterization of sugarcane MYB transcription factor gene promoter (PScMYBAS1) in response to abiotic stresses and hormones. *Plant Cell Rep* 2012, **31**:661-669.

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

55. Howe A, Sato S, Dweikat I, Fromm M, Clemente T: Rapid and reproducible Agrobacterium-mediated transformation of sorghum. *Plant Cell Rep* 2006, **25**:784-791.

56. Jakob K, Zhou F, Paterson A: Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. In *Biofuels*. Edited by Tomes D, Lakshmanan P, Songstad D. New York: Springer; 2011:113-138.

57. Manners J, Casu R: Transcriptome analysis and functional genomics of sugarcane. *Tropical Plant Biol* 2011, **4**:9-21.

58. Mendes Souza G, Berges H, Bocs S, Casu R, D'Hont A, Ferreira J, Henry R, Ming R, Potier B, Van Sluys MA, Vincentz M, Paterson AH: The sugarcane genome challenge: strategies for sequencing a highly complex genome. *Tropical Plant Biol* 2011, **4**:145-156.

59. Tomkins JP, Yu Y, Miller-Smith H, Frisch DA, Woo SS, Wing RA: A bacterial artificial chromosome library for sugarcane. *Theor Appl Genet* 1999, **99**:419-424.

60. Le Cunff L, Garsmeur O, Raboin LM, Pauquet J, Telismart H, Selvi A, Grivet L, Philippe R, Begum D, Deu M, Costet L, Wing R, Glaszmann JC, D'Hont A: Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance gene (Bru1) in highly polyploid sugarcane (2n approximately 12x approximately 115). *Genetics* 2008, **180**:649-660.

61. Figueira TR, Okura V, Rodrigues da Silva F, Jose da Silva M, Kudrna D, Ammiraju JS, Talag J, Wing R, Arruda P: A BAC library of the SP80-3280 sugarcane variety (saccharum sp.) and its inferred microsynteny with the sorghum genome. *BMC Res Notes* 2012, **5**:185.

62. Wang J, Roe B, Macmil S, Yu Q, Murray JE, Tang H, Chen C, Najar F, Wiley G, Bowers J, Van Sluys MA, Rokhsar DS, Hudson ME, Moose SP, Paterson AH, Ming R: Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. *BMC Genomics* 2010, **11**:261.

63. Kim C, Lee TH, Compton RO, Robertson JS, Pierce GJ, Paterson AH: A genome-wide BAC end-sequence survey of sugarcane elucidates genome composition, and identifies BACs covering much of the euchromatin. *Plant Mol Biol* 2013, **81**:139-147.

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

65. Sugarcane Genome Project [http://sugarcanegenome.org]

66. Garcia AA, Mollinari M, Marconi TG, Mancini MC, Costa EA, Henry R, Bundock PC, Pinto LR, Souza AP: Bulding a SNP-based multipoint sugarcane linkage map using loci with all possible dosages [abstract]. In *International Plant and Animal Genome Conference XX:* (conference program). 2012. [https://pag.confex.com/pag/xx/webprogram/Paper1490.html]

67. Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MI, Henrique-Silva F, Giglioti EA, Lemos MV, Coutinho LL, Nobrega MP, Carrer H, França SC, Bacci Júnior M, Goldman MH, Gomes SL, Nunes LR, Camargo LE, Siqueira WJ, Van Sluys MA, Thiemann OH, Kuramae EE, Santelli RV, Marino CL, Targon ML, Ferro JA, Silveira HC, Marini DC, Lemos EG, Monteiro-Vitorello CB, *et al.*: Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res* 2003, **13**:2725-2735.

68. Nishiyama-Jr MY, Vicente F, Sato PM, Ferreira SS, Feltus FA, Souza GM: Transcriptome analysis in the Saccharinae. In *Genomics of the Saccharinae. Volume 11.* Edited by Paterson AH. New York, Berlin: Springer; 2013:121-140.

69. Pastina MM, Malosetti M, Gazaffi R, Mollinari M, Margarido GR, Oliveira KM, Pinto LR, Souza AP, van Eeuwijk FA, Garcia AA: A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. *Theor Appl Genet* 2012, **124**:835-849.

70. Wu KK, Burnquist W, Sorrells ME, Tew TL, Moore PH, Tanksley SD: The detection and estimation of linkage in polyploids using single-dose restriction fragments. *Theoret Appl Genet* 1992, **83**:294-300.

71. Pastina MM, Pinto LR, Oliveira KM, Souza AP, Garcia AA: Molecular mapping of complex traits. In *Genetics, Genomics and Breeding of Sugarcane.* Edited by Henry R, Kole C. Enfield, NH: Science Publishers; 2010:117-148.

72. Palhares AC, Rodrigues-Morais TB, Van Sluys MA, Domingues DS, Maccheroni W, Jr., Jordao H, Jr., Souza AP, Marconi TG, Mollinari M, Gazaffi R, Garcia AA, Vieira ML: A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. *BMC Genet* 2012, **13**:51.

73. Margarido GRA, Souza AP, Garcia AAF: OneMap: software for genetic mapping in outcrossing species. *Hereditas* 2007, **144**:78-79.

74. Jiang S-Y, Ma Z, Vanitha J, Ramachandran S: Genetic variation and expression diversity between grain and sweet sorghum lines. *BMC Genomics* 2013, **14**:18.

75. Dugas DV, Monaco MK, Olsen A, Klein RR, Kumari S, Ware D, Klein PE: Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. *BMC Genomics* 2011, **12**:514.

76. Mizuno H, Kawahigashi H, Kawahara Y, Kanamori H, Ogata J, Minami H, Itoh T, Matsumoto T: Global transcriptome analysis reveals distinct expression among duplicated genes during sorghum-Bipolaris sorghicola interaction. *BMC Plant Biol* 2012, **12**:121.

77. Casu R: Role of bioinformatics as a tool for sugarcane research. In *Genetics, Genomics and Breeding of Sugarcane.* Edited by Henry R, Kole C. Enfield, NH: Science Publishers; 2010:229-248.

78. Menossi M, Silva-Filho MC, Vincentz M, Van Sluys MA, Souza GM: Sugarcane functional genomics: gene discovery for agronomic trait development. *Int J Plant Genomics* 2008, **2008**:458732.

79. Nishiyama MY Jr, Vicente F, Lembke CG, Sato PM, Dal-Bianco ML, Fandino RA, Hotta CT, Souza GM: The SUCEST-FUN Regulatory Network Database: designing and energy grass. *Int Sugar J* 2012, **114**:821-826.

80. Computational Biology and Functional Genomics Laboratory: DFCI Sugarcane Gene Index [http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=s_officinarum]

81. Vicentini R, Bem LEV, Sluys MA, Nogueira FTS, Vincentz M: Gene content

analysis of sugarcane public ESTs reveals thousands of missing coding-genes and an unexpected pool of grasses conserved ncRNAs. *Tropical Plant Biol* 2012, **5:**199-205.

82. Iseli C, Jongeneel CV, Bucher P: ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc Int Conf Intell Syst Mol Biol* 1999:138-148.

83. Ostlund G, Schmitt T, Forslund K, Kostler T, Messina DN, Roopra S, Frings O, Sonnhammer EL: InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res* 2010, **38**(Database issue):D196-203.

84. Alexeyenko A, Tamas I, Liu G, Sonnhammer EL: Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics* 2006, **22:**e9-15.

85. Casu R, Hotta CT, Souza GM: Functional genomics: transcriptomics of sugarcane - current status and future prospects. In *Genetics, Genomics and Breeding of Sugarcane*. Edited by Henry R, Kole C. Enfield, NH: Science Publishers; 2010:167-191.

86. Lembke CG, Nishiyama MY, Jr., Sato PM, de Andrade RF, Souza GM: Identification of sense and antisense transcripts regulated by drought in sugarcane. *Plant Mol Biol* 2012, **79:**461-477.

87. Calsa T Jr, Figueira A: Serial analysis of gene expression in sugarcane (Saccharum spp.) leaves revealed alternative C4 metabolism and putative antisense transcripts. *Plant Mol Biol* 2007, **63:**745-762.

88. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A: Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011, **29:**644-652.

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

90. Ferreira TH, Gentile A, Vilela RD, Costa GG, Dias LI, Endres L, Menossi M: microRNAs associated with drought response in the bioenergy crop sugarcane (Saccharum spp.). *PLoS One* 2012, **7:**e46703.

91. Gentile A, Ferreira TH, Mattos RS, Dias LI, Hoshino AA, Carneiro MS, Souza GM, Calsa T, Jr, Nogueira RM, Endres L, Menossi M: Effects of drought on the microtranscriptome of field-grown sugarcane plants. *Planta* 2013, **237:**783-798.

92. Thiebaut F, Grativol C, Carnavale-Bottino M, Rojas CA, Tanurdzic M, Farinelli L, Martienssen RA, Hemerly AS, Ferreira PC: Computational identification and analysis of novel sugarcane microRNAs. *BMC Genomics* 2012, **13:**290.

93. Domingues DS, Cruz GM, Metcalfe CJ, Nogueira FT, Vicentini R, Alves Cde S, Van Sluys MA: Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns. *BMC Genomics* 2012, **13:**137.

94. Swaminathan K, Alabady M, Varala K, De Paoli E, Ho I, Rokhsar D, Arumuganathan A, Ming R, Green P, Meyers B, Moose SP, Hudson ME: Genomic and small RNA sequencing of Miscanthus x giganteus shows the utility of sorghum as a reference genome sequence for Andropogoneae grasses. *Genome Biol* 2010, **11:**R12.

95. Kim C, Zhang D, Auckland S, Rainville L, Jakob K, Kronmiller B, Sacks E, Deuter M, Paterson A: SSR-based genetic maps of Miscanthus sinensis and M. sacchariflorus, and their comparison to sorghum. *Theoret Appl Genet* 2012, **124:**1325-1338.

96. Swaminathan K, Chae W, Mitros T, Varala K, Xie L, Barling A, Glowacka K, Hall M, Jezowski S, Ming R, Hudson M, Juvik JA, Rokhsar DS, Moose SP: A framework genetic map for Miscanthus sinensis from RNAseq-based markers shows recent tetraploidy. *BMC Genomics* 2012, **13:**142.

97. Wang X, Paterson AH: Comparative genomic analysis of C4 photosynthesis pathway evolution in grasses. In *Genomics of the Saccharinae. Volume 11*. Edited by Paterson AH. New York, Berlin: Springer; 2013:447-477.

98. Hatch MD, Slack CR: Photosynthesis by sugar-cane leaves. A new carboxylation reaction and the pathway of sugar formation. *Biochem J* 1966, **101:**103-111.

99. Seemann JR, Sharkey TD, Wang J, Osmond CB: Environmental effects on photosynthesis, nitrogen-use efficiency, and metabolite pools in leaves of sun and shade plants. *Plant Physiol* 1987, **84:**796-802.

100. Hattersley PW: The distribution of C3 and C4 grasses in australia in relation to climate. *Oecologia* 1983, **57:**113-128.

101. Ehleringer JR, Bjorkman O: A comparison of photosynthetic characteristics of encelia species possessing glabrous and pubescent leaves. *Plant Physiol* 1978, **62:**185-190.

102. Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, Ehleringer JR: Global vegetation change through the Miocene/Pliocene boundary. *Nature* 1997, **389:**153-158.

103. Sage RF: The evolution of C4 photosynthesis. *New Phytologist* 2004, **161:**341-370.

104. Muhaidat R, Sage RF, Dengler NG: Diversity of Kranz anatomy and biochemistry in C4 eudicots. *Am J Bot* 2007, **94:**362-381.

105. Giussani LM, Cota-Sanchez JH, Zuloaga FO, Kellogg EA: A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C4 photosynthesis. *Am J Bot* 2001, **88:**1993-2012.

106. Pyankov VI, Artyusheva EG, Edwards GE, Black CC Jr, Soltis PS: Phylogenetic analysis of tribe Salsoleae (Chenopodiaceae) based on ribosomal ITS sequences: implications for the evolution of photosynthesis types. *Am J Bot* 2001, **88:**1189-1198.

107. Vicentini A, Barber JC, Aliscioni SS, Giussani LM, Kellogg EA: The age of the grasses and clusters of origins of C4 photosynthesis. *Glob Change Biol* 2008, **14:**2963-2977.

108. Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen V, Salamin N: Oligocene CO_2 decline promoted C4 photosynthesis in grasses. *Curr Biol* 2008, **18:**37-43.

109. Sheen J: C4 gene expression. *Annu Rev Plant Physiol Plant Mol Biol* 1999, **50:**187-217.

110. Buckeridge MS, Cavalari AA, Silva GB: Parede celular. In *Fisiologia Vegetal*. Edited by Kerbauy GB. Rio de Janeiro: Guanabara Koogan; 2008:165-181.

111. Rooney WL, Blumenthal J, Bean B, Mullet JE: Designing sorghum as a dedicated bioenergy feedstock. *Biofuels Bioprod Bioref* 2007, **1:**147-157.

112. Nogueira FT, De Rosa VE Jr, Menossi M, Ulian EC, Arruda P: RNA expression profiles and data mining of sugarcane response to low temperature. *Plant Physiol* 2003, **132:**1811-1824.

113. Rocha FR, Papini-Terzi FS, Nishiyama MY Jr, Vêncio RZ, Vicentini R, Duarte RD, de Rosa VE Jr, Vinagre F, Barsalobres C, Medeiros AH, Rodrigues FA, Ulian EC, Zingaretti SM, Galbiatti JA, Almeida RS, Figueira AV, Hemerly AS, Silva-Filho MC, Menossi M, Souza GM: Signal transduction-related responses to phytohormones and environmental challenges in sugarcane. *BMC Genomics* 2007, **8:**71.

114. Papini-Terzi FS, Rocha FR, Vêncio RZ, Oliveira KC, Felix Jde M, Vicentini R, Rocha Cde S, Simões AC, Ulian EC, di Mauro SM, da Silva AM, Pereira CA, Menossi M, Souza GM: Transcription profiling of signal transduction-related genes in sugarcane tissues. *DNA Res* 2005, **12:**27-38.

115. Casu RE, Jarmey JM, Bonnett GD, Manners JM: 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* 2007, **7:**153-167.

116. McCormick AJ, Cramer MD, Watt DA: Regulation of photosynthesis by sugars in sugarcane leaves. *J Plant Physiol* 2008, **165:**1817-1829.

117. Buchanan CD, Lim S, Salzman RA, Kagiampakis I, Morishige DT, Weers BD, Klein RR, Pratt LH, Cordonnier-Pratt MM, Klein PE, Mullet JE: Sorghum bicolor's transcriptome response to dehydration, high salinity and ABA. *Plant Mol Biol* 2005, **58:**699-720.

118. Salzman RA, Brady JA, Finlayson SA, Buchanan CD, Summer EJ, Sun F, Klein PE, Klein RR, Pratt LH, Cordonnier-Pratt MM, Mullet JE: Transcriptional profiling of sorghum induced by methyl jasmonate, salicylic acid, and aminocyclopropane carboxylic acid reveals cooperative regulation and novel gene responses. *Plant Physiol* 2005, **138:**352-368.

119. Zhu-Salzman K, Salzman RA, Ahn JE, Koiwa H: Transcriptional regulation of sorghum defense determinants against a phloem-feeding aphid. *Plant Physiol* 2004, **134:**420-431.

120. Park SJ, Huang Y, Ayoubi P: Identification of expression profiles of sorghum genes in response to greenbug phloem-feeding using cDNA subtraction and microarray analysis. *Planta* 2006, **223:**932-947.

121. Zanca AS, Vicentini R, Ortiz-Morea FA, Del Bem LE, da Silva MJ, Vincentz M, Nogueira FT: Identification and expression analysis of microRNAs and targets in the biofuel crop sugarcane. *BMC Plant Biol* 2010, **10:**260.