A REVIEW OF GENETIC GAIN IN SUGARCANE BREEDING USING GENOMIC SELECTION IN DIFFERENT COUNTRIES

Muhammad Ehsan Khan
Sugarcane Research and Development Board Faisalabad Pakistan
Email: ehsansrdb@gmail.com

ABSTRACT

Sugarcane is an industrial crop cultivated in tropical and subtropical regions of the world. It is an emerging source of sustainable bioenergy, accounting for more than 70% of world sugar consumption. The increase in productivity from sugarcane has been small compared to other major crops, and the rate of genetic gains from current breeding programs tends to be plateauing. In this review, some of the main contributors for the relatively slow rates of genetic gain are discussed, including (i) breeding cycle length and (ii) low narrow-sense heritability for major commercial traits, possibly reflecting strong non-additive genetic effects involved in quantitative trait expression. A general overview of genomic selection (GS), a modern breeding tool that has been very successfully applied in plant breeding, is given. This review discusses key elements of GS and its potential to significantly increase the rate of genetic gain in sugarcane, mainly by (i) reducing the breeding cycle length, (ii) increasing the prediction accuracy for clonal performance, and (iii) increasing the accuracy of breeding values for parent selection. GS approaches that can accurately capture non-additive genetic effects and potentially improve the accuracy of genomic estimated breeding values are particularly promising for the adoption of GS in sugarcane breeding. Finally, different strategies for the efficient incorporation of GS in a practical sugarcane breeding context are presented. These proposed strategies hold the potential to substantially increase the rate of genetic gain in future sugarcane breeding.

Key words: Genomic selection, sugarcane breeding, genetic gain and selection breeding

INTRODUCTION

1. Importance and production trends

Sugarcane is a C4 plant commercially grown in tropical and subtropical regions worldwide. It is one of the oldest cultivated plants in the worldwith ancient history. Sugarcane accounts for more than 70% of the total sugar produced globally, mostly consumed as refined sugar. Recently, sugarcane has received attention as an energy crop; in many countries, including Australia, Brazil, India bagasse (the fibrous part after juice extraction) is burnt by sugar mills to produce electricity to power the mills' operations.

Sugarcane is also used for animal feed (green leaves and top portion), alcoholic beverages, and as a fertilizer (trash) in crop production across the globe. Sugarcane is the world's most produced crop (total production) and ranks among the ten most widely grown crops worldwide. The total global production of sugarcane in 2021–2022 was 2.3 billion tons, and it was grown in approximately 100 countries, covering an area of ~28 million hectares. The largest sugarcane producer is Brazil (40% of the total production), followed by India, China, Thailand and Pakistan. Other major sugarcane producing countries are Mexico, United States, Colombia, Australia, Cuba, and the Philippines. In the past 60 years, world sugarcane production increased almost three-half fold, mainly because of the rising demand for sugar and ethanol.

Production gains are partly attributed to the genetic improvement of sugarcane varieties that are adapted to particular target environments. Concurrently, improvements in management techniques, fertilization, and irrigation have all played a role in increasing sugarcane productivity. The main driver to the total increase in production is the dramatic increase in cultivated land area.

The occurrence of new diseases and pests could cause increased losses. Continuing monoculture cropping can build up soil pathogens and nematode pressure, which might be partly responsible for a lack of sugarcane yield increase worldwide (Stirling *et al.,* 2001). Additionally, diseases have been observed to substantially impact sugarcane yield. Red rot of sugarcane is one of the most economically important sugarcane diseases worldwide. Reported yield losses due to red rot are 15–50% in irrigated and rainfed conditions in Pakistan and 29% in Fiji (Johnson and Tyagi 2010). Red rot primarily affects yield, while key quality characteristics like sugar content are also affected.

Another major disease that affects sugarcane crops worldwide is sugarcane smut, which can have devastating impacts on yield. The estimated average potential losses due to sugarcane smut in the Punjab region and some losses in Sindh region also reported. Nearly 70% of the sugarcane cultivars were susceptible to smut (Sundar *et al.,* 2012); sugarcane smut resistance is now one of the primary breeding objectives for Pakistan sugarcane.

Extreme weather can also have significant impacts on sugarcane yield. In Pakistan, favorable growing conditions in 1994 resulted in 5.2M tons of national production. In subsequent years, sugarcane production was reported to be reduced by half in the same region because of extreme climatic fluctuation Gawander. 2007). Similar observations were reported in China in 2003–2004, where drought decreased average cane yields by around 18% (Li *et al.,* 2006).

However, as there is no evidence that these negative impacts have increased over the periods of low productivity improvement, the impact of environment-management is not sufficient to explain the continuous slow rate of improvement in sugarcane yield over time. In addition to improving management practices, the genetic improvement of modern cultivars is a main avenue to enhance productivity in sugarcane. To overcome static yield trends, intensified breeding efforts are needed to develop new, improved varieties.

2. Development of Modern Cultivars and Inherent Challenges

Sugarcane (*S. officinarum*) has been cultivated in India, China, and Papua New Guinea for sugar production for 10,000 years. The first sugarcane breeding programs were established in Java and Barbados in the late 1800s after the discovery that sugarcane can produce viable seeds (A.J. Mangelsdorf, 1995; Ming *et al.,* 2010). Until the first quarter of the 20th century, sugarcane varieties used in industrial-scale production of sugar were *S. officinarum* clones, also known as a noble cane, originating from New Guinea.

It is reported that *S. officinarum* species were domesticated from wild *S. robustum* in New Guinea around 8,000 years ago (Ming *et al.,* 2010). Unlike *S. officinarum* Indian cane (S. barberi) and Chinese cane (*S. sinense*) are derived from interspecific hybridization between octoploid *S. officinarum* (2n = 80) and *S. spontaneum* (2n = 40–128) with varying ploidy levels (D'Hont, 2002). Historically, *S. officinarum* species had good commercial milling characteristics such as high sugar content, low impurity levels, and low fiber. However, this species lacked vigor, ratooning performance, and was susceptible to several diseases (Stevenson, 1965). *S. spontaneum* is a genetically diverse wild species that is characterized by a lower commercial merit than *S. officinarum*, because of thin stalks and low sucrose content. Conversely, compared to *S. officinarum*, *S. spontaneum* has an increased ratooning capacity, a higher fiber level, and an overall superior adaptive capacity, characterized by an ability to perform better in unfavorable environmental conditions, such as drought, flood, or high salinity (Mohan and Sreenivasan, 1986). The genetic improvement of sugarcane can be divided in three main phases (Roach, 1989). The first phase began with screening and intercrossing among *S. officinarum* clones.

The major limitation of this approach was that noble canes, and hence progeny created from intercrossing, were susceptible to biotic and abiotic stresses. This led to the second phase, which involved the development of cultivars derived from interspecific hybridization between *S. officinarum* and *S. spontaneum*, and continuous backcrossing efforts with *S. officinarum* clones.

Interspecific hybrids between *S. officinarum* and *S. spontaneum* were able to combine a high cane yield potential with increased disease resistance and improved ratooning ability. The sugar yield in Colombia increased from 5t sugar/ha-year at the end of the 1950s to 8 t sugar/ha-year in the 1970s and recorded 12 t/ha-year at the end of 2000 (Cock, 2001).

Sugarcane production in Brazil and India increased throughout the same period and reached nearly 64–70 t/ha by the end of 2000. Results of a long-term study investigating productivity trends from 1968 to 2000 in Florida demonstrated significant improvements in cane and sucrose yield across the plant cane in first and second-ratoon crops. The positive impacts of genetic gain increases on Florida's sugarcane industry played a significant role in the country's economy across those years (Edme, 2005). However, the observed increases in sucrose yield for the most recent varieties in Florida (unpublished data from a 2011 study) were associated with an increase in total cane yield, rather than improvements in CCS (Zhao and Yang-Rui, 2015). Similar results were reported from three small scale studies conducted in Australia where no significant differences for CCS could be found between older and new varieties (Jackson, 2005). Thus, genetic gain for key traits, particularly sucrose content and, to some extent, cane yield, has been stagnating in the past ten years in some countries. Conversely, genetic improvements for disease resistance achieved through traditional breeding programs have been very substantial.

One consequence of the foundation bottleneck is strong genome-wide linkage disequilibrium (LD) patterns observed in elite germplasm (Aitken, *et al.,* 2006) and a narrow genetic base in modern sugarcane germplasm (Raboin*et al.,* 2008). Commercial hybrids originate from the initial hybrid (*S. officinarum* × *S. spontaneum*), which would have 2n transmission from the *S. officinarum* parent and n transmission from the *S. spontaneum*Price, 1963; Bremer, 1961]. The hybrid is then crossed back to other hybrids to recover the high sugar phenotype, which breaks down the hybrid into n + n transmission (Bremer, 1961). Because of the narrow genetic base of important traits, genetic diversity could be reintroduced in sugarcane by utilizing the potential of wild relatives that are considered reservoirs of potentially useful alleles for important economic traits that might have been lost during domestication and breeding. Such practices of continual introgression of wild material into commercial breeding programs are used intensively in some breeding programs, e.g., in Louisiana. New commercial hybrid cultivars have a complicated chromosome set, ranging between 2n = 100−130; 80% of the chromosomes are of *S. officinarum* origin, 10–15% of the chromosomes are of *S. spontaneum* origin, and the rest of the chromosomes are a combination of the two species Sreenivasan*et al.,* 1987; Garsmeur, *et al.,* 2018). Eight to 14 homologous copies of alleles at a given locus in the hybrid genome are reported in the literature (Grivet and Arruda,2002; Souza, *et al.,* 2005). While the haploid genome of sugarcane is estimated at 1 Gb, the total size of sugarcane nuclear genome is approximately 10Gb (D'Hont, 2001; Le-Cunff, 2008), making it ten times larger than the closest related genome sequenced species, which is sorghum. The extreme polyploid genome of interspecific hybrids possesses irregular genetic characteristics that are passed from both parental species, making it more complicated than that of its precursors (D'Hont*et al.,* 1996, Le Cunff*et al.,* 2008). This phenomenon contributes substantially to the high level of heterozygosity observed between sugarcane cultivars. Because of the random sorting of chromosomes in each crossing, the number of chromosomes varies between genotypes. The complex genetic composition of modern hybrids which are referred to as poly-aneuploids also results in inherent polygenic control of important agronomic traits. This complex genetic structure potentially makes the selection procedure slower and more complicated than in other major crop species.

3. Genomic Selection:

A Powerful New Breeding Tool Genomic selection (GS) is a relatively new breeding method in which individuals are selected based on their predicted breeding values that are calculated from genome-wide DNA marker profiles. Decreasing costs of DNA marker screening methods such as high-density SNP arrays and genotyping by sequencing (GBS) approaches, and the development of statistical methods that can accurately predict marker effects are the main reasons why GS has increasingly been implemented in modern animal and plant breeding programs. Two main avenues by which GS can accelerate the rate of genetic gain is by improving the accuracy at which individuals are selected and by reducing the length of the breeding cycle. However, the incorporation of GS into a breeding program is not a trivial task. It highly depends on several factors, such as the mating type, the genetic architecture and heritability of the target traits, the availability of genotyping platforms, and the total financial budget of the program to build large reference populations that are necessary to accurately estimate the typically small effects of DNA markers that are associated with the underlying causal mutations that affect the traits. Conceptually, GS involves two main steps (Figure-1). The first step is to develop a prediction equation based on a training population (TP) that consists of individuals for which both high-quality phenotypes and genome-wide DNA marker profiles have been obtained. The fundamental requirement for GS to work is that quantitative trait loci (QTL, the actual mutations) that are affecting the expression of the target trait are in LD with the DNA markers that are used for genotyping. If this requirement is met, trait effects for DNA markers can be estimated Agronomy 2020, 10, 585 8 of 21 and used in the prediction equation. In the second step, these marker effects are used to calculate the genomic estimated breeding values (GEBVs) of selection candidates (prediction population; PP) for which only genome-wide marker data (but no phenotypic data) are available. Genotypes can then be ranked based on their GEBVs to support selection decisions in a breeding program.

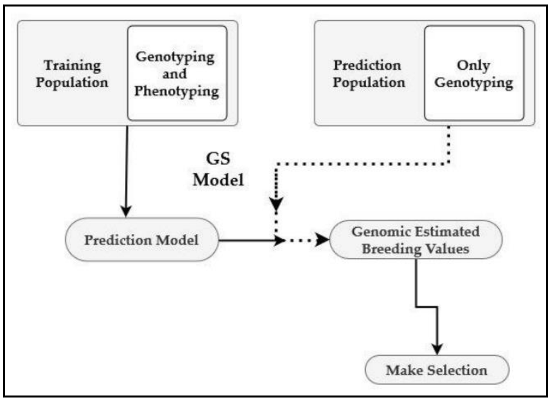

Figure-1: General overview of genomic selection (GS). A GS scheme starts with the training population (TP) that is used to estimate marker effects. These effects are used to calculate genomic estimated breeding values (GEBV) of clones in the prediction population.

4. Implementation of Genomic Selection in Sugarcane Breeding

Increasing the rate of genetic gain is a big challenge in sugarcane breeding, as implied by the static or slowly increasing yield trends in most countries. Several reasons for the observed yield plateaus have been proposed, such as a narrow genetic base of modern elite germplasm (Raboin*et al.,* 2008), highly complex genetic architectures for agronomically important quantitative traits for which non-additive gene action is likely playing a significant role, and very long breeding cycle lengths (Wei and Jackson, 2016). The use of molecular markers has become a standard practice in most important crop species. Traditionally, plant breeders have incorporated molecular markers in phenotypic selection for monoor oligogenic traits to increase the efficiency of the breeding program. For instance, marker-assisted selection (MAS) has proven to be a practical approach for single gene introgression or pyramiding multiple genes in elite cultivars, to improve disease resistance or grain quality. Despite the fact that a range of QTL mapping studies has been undertaken in sugarcane, the size and complexity of the sugarcane genome have limited DNA marker-based selection in this crop (Grivet and Arruda, 2002). Generally, MAS has been largely ineffective for the improvement of highly quantitative traits because of several technical reasons that have been discussed extensively in the literature. Polygenic traits are typically controlled by a huge number of QTL, each having infinitesimal small effects, or possibly with interactions among them as well as with environmental factors.

5. Recurrent Genomic Selection and Reciprocal Recurrent Genomic Selection:

Two Strategies for the Incorporation of Genomic Selection in Sugarcane Breeding Regarding the implementation of GS in sugarcane breeding, a key question is how to incorporate the technology into an existing breeding program. The first critical step in any breeding program is to create new genetic variation. In conventional sugarcane breeding, a large number of seedlings is created through targeted crossing, followed by several selection stages that aim to determine the relative genetic merit of the new germplasm in designed field trials. From the perspective of increasing genetic gain, a key bottleneck with this conventional approach is that alleles are only recombined in the crossing stage at the beginning of the breeding cycle. This could potentially be overcome by a breeding strategy called recurrent genomic selection (RGS) (Figure-2) which aims to rapidly improve the genetic merit of a population of heterozygous genotypes through rapid, recurrent selection and crossing of elite germplasm, and to simultaneously channel selected clones into advanced testing stages that ultimately develop commercial products

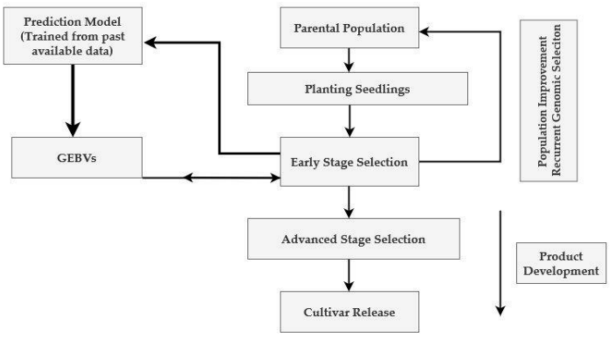

Figure-2: Flow diagram of a Recurrent Genomic Selection (RGS) breeding program for sugarcane.

Conceptually, this can be divided into a population improvement component that uses recurrent genomic selection and a product development component in which clones with high GEBV enter advanced selection stages for variety development. The genomic prediction model is trained using data from previous trials. GEBV = genomic estimated breeding value.

A similar breeding system could be initiated with a small number (2 or 3) parents on one, or both A and B sides (rather than single parents as in Figure-3), and progeny derived from crossing parents on one side would be selected for high predicted breeding values before crossing them with the opposite side. Extending the theory from Cheverud and Routman (1996) to a situation in which a quantitative trait is controlled by many epistatic QTL, in a modified RRGS breeding scheme, the QTL alleles in the opposite heterotic group could be fixed (remain unchanged). This could result in a genetic model with increased additive genetic variance and reduced statistical epistasis. This could contribute to an increase in predictability, leading to improved selection efficiency and higher genetic gain. The proposed GS-based breeding schemes can be advantageous when the desired alleles for the traits of interest are available in the breeding germplasm. However, it could be the case that genetic variation for the trait of interest is limited in the primary gene-pool.

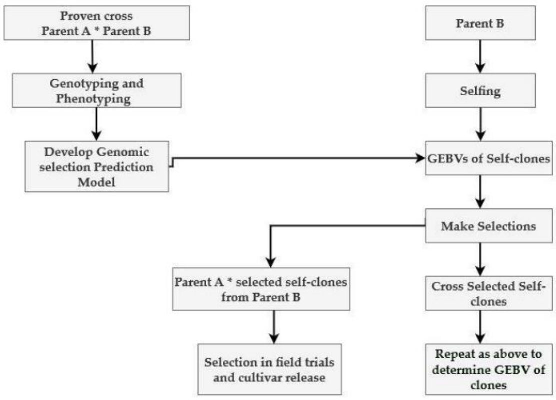

Figure-3: Flow diagram of a modified reciprocal recurrent genomic selection breeding scheme for sugarcane.

The prediction model is trained by generating hundreds of off-springs from a proven cross of unrelated parents that are known to combine well. Either one or both clones in the cross are selfed, and offspring are selected based on their genomic estimated breeding values. If selfing is not feasible, closely related clones (e.g., from the same family) can be used instead. The selfed offspring is crossed with the opposite parent. GEBV = genomic estimated breeding value.

Many modified selection criteria have been proposed to allow balancing genetic gain and maintaining genetic diversity while applying GS. The main idea behind these selection criteria is to determine the exact contribution of an individual to the following generation based on its genetic merit and its genetic relationship with other individuals. Scientists used genomic prediction models, including dominance effects, to predict the performance of offspring generated through mating pairs of individuals. This was followed by an optimization procedure in which a set of mate pairs that can maximize performance in the subsequent generation was selected. In this example, selection and mating were simultaneously performed for improving the management of inbreeding.

The advantage of an adequate mate allocation strategy is particularly relevant for improving complex traits with a high amount of non-additive genetic variance. There are only a few studies that have investigated GS for sugarcane, and the empirical evaluation of different implementation strategies is impractical. Breeding simulations are an elegant way to assess the potential impacts that GS can have on sugarcane breeding efficiency because they require only a few physical resources. Furthermore, simulations can accommodate different genetic models with varying numbers of genes/alleles, dominance, epistatic gene effects, and also handle genotype-environment interaction effects. Empirical validation experiments are then critical to test the most promising strategy in a practical breeding context. Thus, increased simulation efforts could provide valuable information and decision support for the design of empirical validation experiments, and ultimately for the efficient implementation of GS in practical sugarcane breeding. While GS has the potential to tackle fundamental challenges associated with improving important traits in sugarcane, increased research efforts are needed to enable the implementation of the technology. The RGS or RRGS breeding schemes proposed in this paper hold the potential to increase long-term genetic gain for complex quantitative traits in sugarcane, but further investigations are needed.

REFERENCE

Aitken, K.; Jackson, P.; McIntyre, C. Quantitative trait loci identified for sugar related traits in a sugarcane (Saccharum spp.) cultivar× Saccharum officinarum population. *Appl. Genet*. 2006, 112, 1306–1317.

Bremer, G. Problems in breeding and cytology of sugar cane. *Euphytica* 1961, 10, 59–78.
Cock, J.H. Sugarcane growth and development. *Int. SugarJ*. 2001, 105, 5–15.

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

D'Hont, A.; Glaszmann, J.C. Sugarcane genome analysis with molecular markers—A first decade of research. In Proceedings of the International Society of Sugar Cane Technologists: XXIV Congress, Brisbane, Australia, 17–21 September 2001; pp. 556–559.

D'Hont, A.; Paulet, F.; Glaszmann, J. Oligoclonal interspecific origin of 'North Indian' and 'Chinese' sugarcanes. *Chromosome Res*. 2002, 10, 253–262.

Edme, S.J.; Miller, J.D.; Glaz, B.; Tai, P.Y.P.; Comstock, J.C. Genetic contribution to yield gains in the Florida sugarcane industry across 33 years. *Crop Sci*. 2005, 45, 92–97.

Garsmeur, O.; Droc, G.; Antonise, R.; Grimwood, J.; Potier, B.; Aitken, K.; Jenkins, J.; Martin, G.; Charron, C.; Hervouet, C.; et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* 2018, 9, 1–10.

Gawander, J. The impact of climate change on sugar cane production in Fiji. Wmo Bull. 2007, 56, 34–39.

Grivet, L.; Arruda, P. Sugarcane genomics: Depicting the complex genome of an important tropical crop. Curr. Opin. *Plant Biol.* 2002, 5, 122–127.

Jackson, P.A. Breeding for improved sugar content in sugarcane. *Field Crop. Res*. 2005, 92, 277–290.

Johnson, S.S.; Tyagi, A.P. Effect of ratoon stunting disease (RSD) on sugarcane yield in Fiji. South Pac. *J. Nat. Appl. Sci*. 2010, 28, 69–73.

Le Cunff, L.; Garsmeur, O.; Raboin, L.M.; Pauquet, J.; Telismart, H.; Selvi, A.; Grivet, L.; Philippe, R.; Begum, D.; Deu, M.J.G. Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance gene (Bru1) in highly polyploid sugarcane (2n ~12x ~115). *Genetics* 2008, 180, 649–660.

Mangelsdorf, A.J. Sugar-Cane breeding methods. In Proceedings of the International Society of Sugar Cane Technologists: X Congress, Honolulu, HI, USA, 3–22 May 1959.

Ming, R.; Moore, P.H.; Wu, K.K.; D'Hont, A.; Glaszmann, J.C.; Tew, T.L.; Mirkov, T.E.; da Silva, J.; Jifon, J.; Rai, M.; et al. Plant Breeding Reviews; Wiley Blackwell: Hoboken, NJ, USA, 2010; Volume 27, pp. 15–118.

Souza, A.; Grandis, A.; Leite, D.; Buckeridge, M. Sugarcane as a bioenergy source: History, performance, and perspectives for second-generation bioethanol. *Bioenergy Res*. 2014, 7, 24–35.

Stirling, G.; Blair, B.; Pattemore, J.; Garside, A.; Bell, M. Changes in nematode populations on sugarcane following fallow, fumigation and crop rotation, and implications for the role of nematodes in yield decline. *Australas. Plant Pathol*. 2001, 30, 323–335.

Sundar, A.R.; Barnabas, E.L.; Malathi, P.; Viswanathan, R.; Sundar, A.; Barnabas, E.J.B. A mini-review on smut disease of sugarcane caused by SporisoriumScitamineum. In Botany; Mworia, J.K., Ed.; InTech: Rijeka, Croatia, 2012; Volume 16, pp. 107–128.

Li, Y.-R.; Wei, Y.-A. Sugar industry in China: R & D and policy initiatives to meet sugar and biofuel demand of future. *Sugar Tech* 2006, 8, 203–216.

Price, S. Cytogenetics of modern sugar canes. *Econ. Bot*. 1963, 17, 97–106.

Roach, B. Origin and improvement of the genetic base of sugarcane. In Proceedings of the Australian Society of Sugar Cane Technologists: Annual Conference, Tweed Heads, Australia, 2–5 May 1989; pp. 34–47.

Raboin, L.-M.; Pauquet, J.; Butterfield, M.; D'Hont, A.; Glaszmann, J.-C. Analysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. *Appl. Genet.* 2008, 116, 701–714.

Stevenson, G.C. Genetics and Breeding of Sugar Cane; Longmans: London, UK, 1965.

Sreenivasan, T.; Ahloowalia, B. Cytogenetics. In Sugarcane Improvement Through Breeding; Heinz, D.J., Ed.; Elsevier: Amsterdam, The Netherlands, 1987; pp. 211–256.

Souza, G.; Berges, H.; Bocs, S.; Casu, R.; D'Hont, A.; Ferreira, J.; Henry, R.; Ming, R.; Potier, B.; Sluys, M.-A.; et al. The sugarcane genome challenge: Strategies for sequencing a highly complex genome. *Trop. Plant Biol.* 2011, 4, 145–156.

Wei, X.; Jackson, P. Addressing slow rates of long-term genetic gain in sugarcane. In Proceedings of the International Society of Sugar Cane Technologists: XXIX Congress, Chiang Mai, Thailand, 5–9 December 2016; pp. 480–484.

Zhao, D.; Yang-Rui, L. Climate change and sugarcane production: Potential impact and mitigation strategies. *Int. J. Agron*. 2015, 2015, 1–10.