CRC Press
Taylor & Francis Group

Omics Approaches for Sugarcane Crop Improvement

Edited by
Rajarshi Kumar Gaur

Omics Approaches for Sugarcane Crop Improvement

In this book, the information encompasses various researchable biotechnology aspects of sugarcane, its genomic structure, diversity, comparative and structural genomics, data mining, etc. This book explores both the theoretical and practical aspects of sugarcane crops, focusing on innovative processes. This book argues in favor of developing an integrated research and development system to strengthen the research and development capabilities of all the areas of sugarcane. Further, it covers the recent trends of sugarcane biotechnology, especially in the next-generation sequencing (NGS) era. This book will be very useful for professors and scientists who are working in the area of sugarcane crops by using molecular biology and bioinformatics. It is also useful for students to use as a reference for their classes or thesis projects.

Key features:

- Discusses an integral part of molecular biology and pivotal tools for molecular breeding; enables breeders to design cost-effective and efficient breeding strategies for sugarcane
- Discusses the harnessing genomics technologies for genetic engineering and pathogen characterization and diagnosis of sugarcane
- Provides new examples and problems, added where needed
- Provides insight from contributors drawn from around the globe

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Omics Approaches for Sugarcane Crop Improvement

Edited by
Rajarshi Kumar Gaur

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Cover image: ©Shutterstock

First edition published 2023
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

Library of Congress Cataloging-in-Publication Data
Names: Gaur, Rajarshi Kumar, editor.
Title: Omics approaches for sugarcane crop improvement / Rajarshi Kumar Gaur.
Description: First edition | Boca Raton, FL : CRC Press, 2023. | Includes bibliographical references.
Identifiers: LCCN 2022030713 (print) | LCCN 2022030714 (ebook) | ISBN 9781032273686 (hardback) | ISBN 9781032273723 (paperback) | ISBN 9781003292425 (ebook)
Subjects: LCSH: Sugarcane. | Genomics. | Plant molecular genetics. | Bioinformatics.
Classification: LCC SB231 .O45 2023 (print) | LCC SB231 (ebook) | DDC 633.6/1--dc23/eng/20220822
LC record available at https://lccn.loc.gov/2022030713
LC ebook record available at https://lccn.loc.gov/2022030714

ISBN: 978-1-032-27368-6 (hbk)
ISBN: 978-1-032-27372-3 (pbk)
ISBN: 978-1-003-29242-5 (ebk)

DOI: 10.1201/9781003292425

Typeset in Times
by MPS Limited, Dehradun

Contents

Preface

Biotechnological interventions have produced agronomically improved genotypes, yet scientists are currently working to engineer the sugarcane crop as a platform for large-scale production of chemicals with industrial as well as therapeutic significance. Hence, biotechnological interventions hold great promise to develop a better sugarcane crop with improved agronomic traits, sugar content, and biofuel production.

This book is designed to be self-contained and comprehensive, targeting professors and scientists working on sugarcane biotechnology and its related fields, such as genomics, applied data mining, computational biology, disease management, etc. In addition, we believe that this book will serve as a useful reference for sugarcane and crop improvement in the omics era. The 15 chapters described in this book comprise the following theme:

1. Sugarcane improvement through biotechnology
2. Genomics and transcriptomics in sugarcane
3. Mechanism of sugarcane stress tolerance
4. Pathogen identification
5. Breeding and genomics approaches for sugar productivity
6. Sugarcane as a biofuel

Moreover, we have refined the authors' views in a simpler manner that can be easily understood by the readers.

In this book, the information encompasses various researchable biotechnology aspects of sugarcane, its genomic structure, diversity; comparative and structural genomics; data mining; etc. This book explores both the theoretical and practical aspects of sugarcane crops, focusing on the innovation processes. This book argues in favour of developing an integrated research and development system that will strengthen the research and development capabilities of all areas of sugarcane. Furthermore, it covers the recent trends in sugarcane biotechnology, especially in the next-generation sequencing (NGS) era.

To optimise sugarcane improvement, it is necessary to know the impact a selected trait will have on the general physiology of the plant. However, this is not yet possible because there are too many gaps in our knowledge of the unique development and physiology of sugarcane. Such gaps impair our ability to enhance desired agronomical traits; for example, selection for sugarcane varieties with increased photosynthetic capacity may be useless if sugar accumulation is constrained by temperature, water deficit, or nutrient availability. It may prove difficult to consistently increase sucrose levels in the culm without first knowing the factors that affect sugarcane yield and carbon partitioning. Molecular biology has begun to play an important role in agriculture due to its ability to modify microorganisms, plants, animals, and agricultural processes. In particular, it can aid conventional plant-breeding programmes using molecular markers. Several types of DNA markers and molecular-breeding strategies are now available to plant breeders and geneticists to help overcome many problems faced in conventional breeding. The chapters in this book present an overview of the differentially expressed genes in sugarcane under different biotic and abiotic stressors and the use of new biotechniques to develop tolerance in sugarcane against all stresses. This book will be an important resource for providing an excellent reference book for researchers and students working on sugarcane and its improvement.

Rajarshi Kumar Gaur

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Editor Bio

Prof. (Dr.) Rajarshi Kumar Gaur achieved his PhD in 2005 and now serves as professor for the Department of Biotechnology, Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur, Uttar Pradesh, India. His PhD was on molecular characterization of sugarcane viruses, viz., mosaic, streak mosaic, and yellow luteovirus. He has been the recipient of the MASHAV fellowship of the Israeli government for his post-doctoral studies and joined The Volcani Centre, Israel and BenGurion University, Negev, Israel. In 2007, he received the Visiting Scientist Fellowship from the Swedish Institute Fellowship, Sweden, to work in the Umeå University, Umeå, Sweden. Kumar received his post-doctoral fellowship from ICGEB, Italy, in 2008. He has made significant contributions on plant viruses and published 130 national/international papers, authored 18 edited books, and presented about 50 papers in the national and international conferences.

He has been the recipient of a Fellow of Linnean Society, Fellow of Royal Society of Biology, Fellow of Society of Plant Research, Fellow of Society of Applied Biology (FSAB), and Fellow of International Society of Biotechnology (FISBT). He has been awarded many other honors, including: Prof. B.M. Johri Memorial Award, Society of Plant Research (SPR); Excellent Teaching Award by Astha Foundation, Meerut; UGC-Research Teacher Award; Young Scientist Award, 2012, in Biotechnology by Society of Plant Research (SPR), Meerut; Scientific & Applied Research Centre Gold Medal Award, 2011 for Outstanding Contribution in the Field of Biotechnology. He has visited several laboratories of the United States, Canada, New Zealand, UK, Thailand, Sweden, Italy. He is also associated with several international journals as an academic editor and reviewer: Frontiers in Microbiology, Plose One, Scientific Reports, 3Biotech, J. of Plant Growth Regulators, Molecular Biology Report, Plant Disease, Current Genomic, Scientific African, Indian Phytopathology, Virus Disease, etc. Currently, he is involved in many national & international grants and collaborative projects on plant viruses and disease management.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contributors

Niaz Ahmad
National Institute for Biotechnology and Genetic Engineering (NIBGE)
Faisalabad, Pakistan
and
Constituent College Pakistan Institute of Engineering and Applied Sciences
Faisalabad, Pakistan

Harinath Babu
Molecular Biology and Genetic Engineering Laboratory
Vasantdada Sugar Institute
Pune, Maharashtra, India

G Boobalakrsihan
Division of Vegetable Science
Indian Agricultural Research Institute, Pusa Campus
New Delhi, India

F. Budeguer
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross
Las Talitas, Tucumán, Argentina

Rachayya Devarumath
Molecular Biology and Genetic Engineering Laboratory
Vasantdada Sugar Institute
Pune, Maharashtra, India

Suman Devarumath
Vidya Pratishthan's College of Agricultural Biotechnology
Baramati, Maharashtra, India

Abdelaleim Ismail ElSayed
Biochemistry Department
Faculty of Agriculture
Zagazig University
Zagazig, Egypt

R. Enrique
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross
Las Talitas, Tucumán, Argentina

Zeshan Haider
Functional Genomics and Bioinformatics Program
National Institute for Genomics and Advanced Biotechnology
National Agricultural Research Centre
Islamabad, Pakistan

Somnath Kadappa Holkar
ICAR-National Research Centre for Grapes
Pune, Maharashtra, India

Muhammad Ramzan Khan
Functional Genomics and Bioinformatics Program
National Institute for Genomics and Advanced Biotechnology
National Agricultural Research Centre
Islamabad, Pakistan

Haris Khurshid
Oil Seed Research Program
Crop Sciences Institute
National Agricultural Research Center
Islamabad, Pakistan

S. Radhesh Krishnan
T. Stanes & Company Ltd
Coimbatore, Tamil Nadu, India

Atul Kumar
Amity Institute of Biotechnology
Amity University
Uttar Pradesh Campus, Lucknow, India

Ravindra Kumar
ICAR-Indian Agricultural Research Institute
Regional Station
Karnal, Haryana, India

Sarita Kumari
Department of Agricultural Biotechnology and Molecular Biology
CBS&H, RPCAU-Pusa
Samastipur, Bihar, India

P. Malathi
ICAR-Sugarcane Breeding Institute
Coimbatore, India

Ashutosh Kumar Mall
ICAR-Indian Institute of Sugarcane Research
Lucknow, India

Smriti Mall
Molecular Plant Pathology Lab.
Department of Botany
DDU Gorakhpur University
Gorakhpur, Uttar Pradesh, India

Varucha Misra
ICAR-Indian Institute of Sugarcane Research
Lucknow, India

Zahid Mukhtar
National Institute for Biotechnology and Genetic Engineering (NIBGE)
Faisalabad, Pakistan
and
Constituent College Pakistan Institute of Engineering and Applied Sciences
Faisalabad, Pakistan

Pandiyan Muthuramalingam
Department of Horticultural Science
Gyeongsang National University
Jinju, Korea

Pramod N. Nair
Department of Forestry
Fiji National University
Fiji

Gauri Nerkar
Molecular Biology and Genetic Engineering Laboratory
Vasantdada Sugar Institute
Pune, Maharashtra, India

K. Nithya
ICAR-Sugarcane Breeding Institute
Coimbatore, India

A.S. Noguera
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross
Las Talitas, Tucumán, Argentina

Muhammad Noman
Functional Genomics and Bioinformatics Program
National Institute for Genomics and Advanced Biotechnology
National Agricultural Research Centre
Islamabad, Pakistan

Ahmad A. Omar
Biochemistry Department
Faculty of Agriculture
Zagazig University
Zagazig, Egypt
and
Citrus Research and Education Center
University of Florida, IFAS
Lake Alfred, FL, USA

S. Ostengo
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross
Las Talitas, Tucumán, Argentina

A.D. Pathak
ICAR-Indian Institute of Sugarcane Research
Lucknow, India

M.F. Perera
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross Las Talitas, Tucumán, Argentina

N. Prabhakaran
T. Stanes & Company Ltd
Coimbatore, Tamil Nadu, India

Madhavi Purankar
Molecular Biology and Genetic Engineering Laboratory
Vasantdada Sugar Institute
Pune, Maharashtra, India

J. Racedo
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA)
Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
CCT NOA Sur. Av. William Cross Las Talitas, Tucumán, Argentina

Saleem Ur Rahman
National Institute for Biotechnology and Genetic Engineering (NIBGE)
Faisalabad, Pakistan
and
Constituent College Pakistan Institute of Engineering and Applied Sciences
Faisalabad, Pakistan

Hamza Ramzan
National Institute for Biotechnology and Genetic Engineering (NIBGE)
Faisalabad, Pakistan
and
Constituent College Pakistan Institute of Engineering and Applied Sciences
Faisalabad, Pakistan

Ghulam Raza
National Institute for Biotechnology and Genetic Engineering (NIBGE)
Faisalabad, Pakistan
and
Constituent College Pakistan Institute of Engineering and Applied Sciences
Faisalabad, Pakistan

Manuel Díaz de los Ríos
Cuban Research Institute of Sugar Cane By-products
Havana, Cuba

Vimanth S
Department of Agricultural Biotechnology and Molecular Biology
CBS&H, RPCAU-Pusa
Samastipur, Bihar, India

H.K. Sachan
Department of Crop Science
Fiji National University
Fiji

Bilal Saleem
Functional Genomics and Bioinformatics Program
National Institute for Genomics and Advanced Biotechnology
National Agricultural Research Centre
Islamabad, Pakistan

Sumeet Kumar Singh
Department of Seed Science and Technology
PGCA, RPCAU – Pusa
Samastipur, Bihar, India

A. Ramesh Sundar
ICAR-Sugarcane Breeding Institute
Coimbatore, India

M. Swapna
Division of Crop Improvement
ICAR-Indin Institute of Sugarcane Research
Lucknow, Uttar Pradesh, India

Swapan K. Tripathy
Department of Agricultural Biotechnology
Odisha University of Agriculture and Technology
College of Agriculture, Siripur Sq.
Bhubaneswar, Odisha, India

R. Viswanathan
ICAR-Sugarcane Breeding Institute
Coimbatore, India

Kai Wang
College of Life Sciences
Nantong University
Jiangsu, China

Georgina Michelena Álvarez
Cuban Research Institute of Sugar Cane By-products
Havana, Cuba

1 Potential and Advanced Strategies for Sugarcane improvement

M.F. Perera, F. Budeguer, R. Enrique, S. Ostengo, A.S. Noguera, and J. Racedo
Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA), Estación Experimental Agroindustrial Obispo Colombres (EEAOC) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). CCT NOA Sur. Av. William Cross, Las Talitas, Tucumán, Argentina

CONTENTS

1.1 INTRODUCTION

Genetic improvement has played a fundamental role in the development of the sugar industries in almost all the sugarcane-producing countries since it is considered the main method to improve productivity and controls disease (Hogarth et al. 1997). Sugarcane breeding began at the end of the 19th century (Stevenson 1965), and since then, it has been a continuous process mainly oriented toward increasing sugar yield, biomass, resistance to biotic factors, and tolerance to abiotic stresses. Most sugarcane-producing regions of the world have their own breeding programs to develop and improve local varieties adapted to their environments and cultivation practices. Breeding programs make crosses between already improved clones, whose origin usually goes back to a few parents (less than 20 *S. officinarum*; Raboin et al. 2008), whereby the genetic base of modern cultivars is reduced (Arceneaux 1967, Selvi et al. 2006). Conventional breeding applies classical genetic principles based on the phenotype and consists of selecting individuals with

DOI: 10.1201/9781003292425-1

desired characteristics to subject them to cycles of crossing and selection. This process is complex and slow, taking each cycle between 10 and 15 years to obtain a new elite cultivar.

Although genetic-improvement programs may differ in strategies and criteria applied, a general scheme is the basis of most current programs (Ming et al. 2006), and it consists of three stages:

1. Selection of parents from a collection or germplasm bank
2. Hybridization using biparental crosses and / or polycrosses
3. Selection of progeny through clonal multiplication steps

The choice of parents depends on the objectives the program pursues and flower availability. For this stage, it is essential to count with a germplasm bank, which constitutes the source of genetic variability for all the traits to be improved. It is important to maximize the genetic distance between the parents for crossings. Genetic differences of parents were estimated considering genetic relationships, geographical origin, and morphological traits. However, pedigree does not always reflect genetic inheritance because of eventual mixed pollen, selfing, and seed admixture (Nair et al. 2002). Since morphology is highly influenced by environment, genetic component of sugarcane accessions could be misestimated from these traits, requiring extensive field determination to gain accuracy. Currently, numerous molecular marker techniques are successfully used to estimate genetic relationships among germplasm bank accessions (see below). Nowadays, the most important criteria for choosing parents are based on either the performance of the parent in the target environment exclusively or additional information on the behavior of their offspring. Moreover, crossing designs have certain limitations to obtain desired combinations and seed amounts, such as flowering timing, categorization of genotypes as male or female, and availability of glasshouse and photo-period facilities (Zhou 2013).

In the second stage, the production of sexual seed is carried out through crosses for the generation of highly variable populations, which will constitute the basis of selection. Regarding the hybridization methods used, biparental crosses are the main strategy employed in most programs, although polycrosses are also carried out to a lesser extent (Ming et al. 2006). An essential requirement to carry out the crosses is that the cane flourishes, a phenomenon that does not occur in all regions; therefore, many breeding programs have adopted the use of photo-periodic and temperature treatments to induce this process. During crossing, temperature and humidity are subject to strict controls to ensure both good pollen viability and survival, as well as seed set. After careful handling, the obtained seeds are stored at -20°C until required, remaining viable for nearly 10 years.

Finally, the third stage consists of identifying and selecting the superior genotypes within the populations generated from crosses. Sugarcane clones are highly heterozigous, allopolyploids and frequently aneuploidy, and, for this reason, the highly variable first generation is subject to selection. It is essential to know the nature and magnitude of the variability present in breeding populations. Selected clones are reproduced and evaluated in different environments and cutting ages, in successive stages of clonal multiplication quantifying environmental effects on the expression of the traits of interest. Parameters such as phenotypic and genotypic variance, and their additive and dominant components, are relevant for establishing a better performance ranking. The large amount of phenotypic data collected for each trait must be subjected to statistical analyses, which provide accurate estimates of the genetic and environmental effects that impact the development and evaluation of cultivars. Statistical analysis based on modern and appropriate models for each stage of selection for sugarcane clones improves characterization of materials and increases the efficiency of the process (Balzarini 2000; Ostengo 2020). After data analysis, the genotypes that met the best characteristics could be released as new commercial varieties.

Although the genetic improvement of sugarcane has been very efficient in obtaining varieties with outstanding characteristics in biomass production, progress in improving industrial performance has been limited in recent years (Jackson 2005; Dal-Bianco et al. 2012). An apparent plateau has been reached in terms of sugar concentration, so the selection of new high-yield varieties has become very

difficult (Ming et al. 2002). It must be highlighted that the efficient production of biomass from sugarcane arouses great interest due to its potential to generate ethanol as a biofuel and electricity, reducing the dependence on fossil fuels and the emission of greenhouse gases (Goldemberg 2007). Considering the aforementioned reasons, optimizing the efficiency of breeding programs for sugar and biomass is important for meeting the growing demand for sugarcane.

Molecular knowledge does not compete with classical improvement, but it is intended to improve the quality of the estimates required to carry out each of the breeding stages, adding certainties and accuracy to the determinations. Also, other biotechnological tools, such as *in vitro* propagation of new varieties or quarantine sanitation to introduce new germplasm, can be assisted with molecular markers to increase the sensitivity of disease detection and genetic purity tests. Besides, when an outstanding variety is obtained by conventional breeding, one or few characteristics can be added through transformation and / or gene editing. For this process, numerous omic tools are used in the search for genes and regulatory sequences of interest in databases, the *in silico* design of the constructions for genetic transformation, and the determination of the insertion sites in the genome, among other key steps for the successfully commercial release of the new transgenic variety.

1.2 EXPLORATION, CONSERVATION, AND UTILIZATION OF GENETIC RESOURCES OF SUGARCANE GERMPLASM BANK ASSISTED BY MOLECULAR MARKERS

The *Saccharum* genus traditionally comprised six species: *S. spontaneum*, *S. officinarum*, *S. robustum*, *S. edule*, *S. barberi* and *S. sinense* (Daniels and Roach 1987; D'Hont et al. 1998). However, molecular studies suggested that the genus should be reduced to only two species, *S. officinarum* and *S. spontaneum*, considering the four remaining as hybrids between these two species (Irvine 1999; D'Hont et al. 2002).

Modern sugarcane cultivars are hybrids derived from interspecific crossings performed a century ago between *S. officinarum* L. (2n = 80), the noble sugar-producing species, and the wild species, *S. spontaneum* (2n = 40 – 128) with high fiber content and stress tolerance (Piperidis et al. 2010b; Vieira et al. 2018). The introgression in sugarcane breeding that involved the crossings between these two species and a series of backcrossing is known as nobilization process.

The sugarcane hybrids possess highly polyploid and aneuploidy genomes of 53 to 143 chromosomes (Ingelbrecht et al. 1999), with meiotic abnormalities and an estimated size of >10 gigabases (Gb) (Zhang et al. 2012), i.e., it is one of the most complex genomes of any crop plant (Aitken, 2021). Genetic studies have shown that the alloautopolyploid genome of these hybrids contains about 80% of chromosomes from *S. officinarum*, 10% of chromosomes from *S. spontaneum*, and 10% recombinant chromosomes between the two progenitor species (D'Hont et al. 1996).

The complexity and size of the sugarcane genome is a major limitation in its genetic improvement (Singh et al. 2010). For that reason, in recent years, numerous initiatives to understand the complexity of the sugarcane genome had led to different sequencing projects around the world for both parental species and hybrid genotypes. A minimum tiling path of 4,660 sugarcane BACs of the genome in cultivar R570 that best covers the gene-rich part of the sorghum genome was sequenced and then assembled in a 382 megabases (Mb) single-tiling path (Garsmeur et al. 2018), providing an essential genome template for aligning sequencing data. Also, the assembly of one haploid *S. spontaneum* genome derived from SES208 (Zhang et al. 2018) and gene-space assembly of the Brazilian variety SP80–3280 (Souza et al. 2019), were completed. Besides, genomes of Khon Kaen KK3, a Thailand variety (Shearman et al. 2018) and of the Colombian variety CC01–1940 (www.cenicaña.org 2020) were sequenced.

Increasing sugar productivity is the main concern of sugarcane breeding programs (Singh et al. 2010). To develop new varieties, the breeding programs are assembling germplasm collections to increase the possibilities of novel gene combinations. In that sense, there are two sugarcane global

collections held in India and the U.S.A. The Indian collection consists of 1,819 clones, whereas the American collection contains 887 sugarcane clones. Both collections can be used for base-broadening breeding aimed at efficient introgression of desirable alleles (Fickett et al. 2020). Besides, national breeding programs in each country also maintain collections of different numbers of genotypes and exchange their materials to provide germplasm with desirable traits. The safe maintenance of the germplasm collections is a valuable tool for breeders considering the risk of genetic erosion, consisting in a reduction of allelic richness (Van de Wouw et al. 2010).

In terms of diversity, every genus or species is a useful genetic resource for conservation and use, and more efficient introgression programs could be developed with a better knowledge of the donor wild germplasm. Wild sugarcane relatives were employed from the beginning of introgression during the nobilization process, and they are still used in some breeding programs (Cursi et al. 2022). Thus, different breeding programs incorporated genetic material from *Erianthus* species originating from the Old World (*Erianthus Michx*. sect. Ripidium Henrard) and the New World, particularly North America, into *Saccharum* L. (Besse et al. 1998). Among them, *Erianthus arundinaceus* (Retz.) Jeswiet has received special attention due to its high biomass productivity in ratoon crops and its drought tolerance (Jackson and Henry, 2011). However, many sugarcane breeding programs have been crossing it with sugarcane (D'Hont et al. 1995; Besse et al. 1997; Piperidis et al. 2000; Cai et al. 2005a; Nair et al. 2006; Krishnamurthi et al. 2007; Piperidis et al. 2010a; Fukuhara et al. 2013), low production efficiency has been reported among intergeneric hybrids (Piperidis et al. 2000), due to an incompatibility between the *Saccharum* and the *Erianthus* genomes. Several markers were proved useful to follow introgression in *Saccharum* x *Erianthus* hybrids; such as Restriction fragment length polymorphism (RFLP) of the ribosomal DNA (rDNA), (Besse et al. 1996), inter-Alu-like sequences MsCIR2 and EaCIR6 (Alix et al. 1999) and SSR (Chang et al. 2020).

Knowledge of the genetic diversity and population structure among sugarcane parental lines provides valuable information to select the best parental accessions for crossing and broaden the genetic base, i.e., expanding the gene pool, and also contributes to the improved use of genetic resources (Perera et al. 2012a). In this context, molecular markers are powerful tools, both for estimating genetic diversity and understanding the complex genetics of sugarcane (D'Hont et al. 1997). Different types of molecular markers have been used for diversity studies.

In one of the first attempts to estimate diversity, variation of nuclear ribosomal (r) DNA was investigated in cultivated and wild *Saccharum* accessions by digestion with restriction enzymes. Results revealed that clones of the *Erianthus* and *Miscanthus* genera had unique patterns and variation was most extensive in *S. spontaneum*, whereas *S. robustum* was only slightly variable (Glaszmann et al. 1990). After that, 5S rRNA intergenic spacers were amplified and sequenced. Direct sequencing of cloned PCR products revealed different spacer length and sequence polymorphism for *S. officinarum*, *S. spontaneum*, sugarcane cultivars, *S. giganteum*, *Erianthus* spp., *M. sinensis*, Sorghum bicolor, and *Zea mays* (Pan et al. 2000).

RFLP markers with both chloroplast and mitochondrial heterologous probes were useful to study diversity in cultivated (*S. officinarum*, *S. barberi* and *S. sinense*) and wild (*S. spontaneum* and *S. robustum*) genomes in the *Saccharum* complex (D'Hont et al. 1993).

Lu et al. (1994) also performed an RFLP analysis on nuclear DNA of wild and old cultivated sugarcane accessions. Multivariate analyses of the data allowed the separation of the three basic species, *S. spontaneum*, *S. robustum,* and *S. officinarum* and showed that *S. spontaneum* had structure that could be related to the geographic origin of the clones supporting the hypotheses on the origin of secondary species *S. barberi* and *S. sinense*. The authors suggested that RFLPs will be essential in sugarcane for depicting the genomic constitution of modern cultivars.

A collection of *Erianthus* sect. *Ripidium* accessions and *Saccharum* representatives (*S. officinarum* and *S. spontaneum*) were also studied by RFLP analysis using dispersed nuclear single-copy probes from maize (Besse et al. 1997). Results were confirmed by Besse et al. (1998) in an assessment of genetic diversity within and between *Saccharum*, *Erianthus* sect. *Ripidium* and

North American *E. giganteus* (*S. giganteum*) conducted using amplified fragment length polymorphism (AFLP). This study revealed three major groups corresponding to the three genera under study. Within the *Saccharum* group, two subgroups belonged to *S. spontaneum* and *S. officinarum* accessions, respectively, were found.

Moreover, AFLP markers were used to study the genetic relationship among 69 genotypes, representing 16 species from the four genera in the *Saccharum*. Principal component analysis and UPGMA clustering showed that *Saccharum* is closely related to *Miscanthus,* whereas *Narenga* and *Erianthus* were closely related (Cai et al. 2005b).

Selvi et al. (2005) also employed AFLP to understand the genetic constitution of cultivars developed in India suitable for subtropical conditions. Results revealed that the subtropical cultivars retained significantly higher numbers of *S. spontaneum* specific DNA fragments than did the tropical cultivars, reflecting the breeding strategy followed in the development of these cultivars. Additionally, the AFLP-based clustering of the cultivars corresponded well with their pedigree relationships.

Furthermore, AFLP were used to estimate molecular diversity among 270 *S. officinarum* genotypes from all the major regions where *S. officinarum* is grown and 151 cultivars produced by the Australian breeding program or important parents used in it. *S. officinarum* clones from New Guinea displayed greater diversity than *S. officinarum* clones from other regions, confirming the hypothesis that New Guinea is the centre of origin for this species. The *S. officinarum* clones from Hawaii and Fiji formed a separate group that may correspond to introgressed clones with other members of the *Saccharum* complex. Besides, greater diversity was found in the cultivars than in the *S. officinarum* clones due to the introgression of *S. spontaneum* chromatin. It must be highlighted that about 90% of markers present in the *S. officinarum* clone collection were also present in the cultivar collection, suggesting that most of the observed genetic diversity in *S. officinarum* has been captured in Australian sugarcane germplasm (Aitken et al. 2006). Also, in a smaller subset of *S. officinarum*, *S. robustum*, *S. spontaneum*, *S. barberi*, *S. sinense,* and the related genus *Erianthus*, AFLP were employed. The phenetic tree obtained was consistent with the known taxonomical relationships. AFLP gave higher resolution of closely related species into discrete groups than that by RFLP markers, reported earlier. For that, AFLP markers would be an appropriate tool in providing better information about the relationships among the species, estimating diversity, and revealing species and genus specific markers that could be directly applied in sugarcane breeding programs (Selvi et al. 2006).

Random amplified polymorphic DNA (RAPD)-PCR markers were used to characterize a local collection of *S. spontaneum* L. clones and cultivars from the United States. As a substantial degree of genetic diversity was found within the local collection, the information provided would help conserve its genetic diversity (Pan et al. 2005). On the other hand, Indian sugarcane cultivars were characterized using RAPD markers, revealing both the limited genetic base and the need to diversify the genetic base by using new sources from the germplasm (Kawar et al. 2009).

The ability of microsatellite or simple sequence repeat (SSR) markers to determine the genetic diversity between members of the genera *Saccharum* (*S. officinarum*, *S. spontaneum*, *S. sinense*), Old World *Erianthus* Michx. sect. *Ripidium*, North American *E. giganteus* (*S. giganteum*), *Sorghum* and *Miscanthus* was assessed. Clustering revealed a genetic structure for *Saccharum* and *Erianthus* sect. *Ripidium* that reflected closely the relationship previously identified using other marker systems (Cordeiro et al. 2003). It must be highlighted that although SSR are codominant markers in the case of the polyploidy sugarcane, they are scored as present or absent (dominant manner).

Besides, to ensure the genetic identity during shipping and handling, 116 Louisiana commercial sugarcane clones were also genotyped with 21 SSR markers by Pan (2007). After that, recognizing the SSR technique as a valuable tool in characterizing and identifying sugarcane genotypes, Pan (2010) developed the world's first SSR marker-based sugarcane database. For that, 1,025 sugarcane clones were genotyped on a fluorescence-capillary electrophoresis detection platform

with 21 highly polymorphic SSR markers. This database, annually updated, provides molecular descriptions for new cultivar registration articles, enables sugarcane breeders to identify mislabeled sugarcane clones in crossing programs, determines the paternity of cross progeny, and ensures that desired cultivars are grown in farmers' fields.

In another study, SSRs were used to access the genetic diversity and interrelationships of *S. barberi*, *S. spontaneum*, *S. officinarum,* and commercial cultivars. Clusters revealed groups of individuals having close genetic relationships based on geographical origin, available pedigree information, adaptation zone, and morphological characters (Singh et al. 2010).

Qi et al. (2012) genotyped with SSR 92 elite parents involved in China Mainland sugarcane breeding programs and four wild relatives. Model-based genetic structure, clustering, and principal components analyses consistently revealed five groups. One of the groups was characterized as the introduction group with 46 cultivars predominantly introduced from Australia, Taiwan, India, and the U.S.A.; another group only contained wild germplasm, whereas the three other groups mostly originated from China Mainland.

Furthermore, You et al. (2013) characterized 115 sugarcane genotypes obtained by different Chinese breeding programs and used for parental crossing by SSR. Results revealed that genotypes released by each breeding program showed closer genetic relationships suggesting that breeding programs should increase their genetic diversity of sugarcane varieties to meet the demand of sugarcane cultivation for sugar and bioenergy use. After that, You et al. (2016) obtained similar results when characterizing 181 sugarcane clones by using fluorescence-labeled genomic with SSR and eight expressed sequence tag (EST)-SSR primer pairs.

In China, 150 of the most popular parental lines used in sugarcane breeding programs were analyzed with 21 fluorescence-labeled SSR markers and high-performance capillary electrophoresis (Wu et al. 2019). Based on the population structure analysis, the accessions were clustered into two distinct sub-populations in PCoA: one that contained the majority of clones introduced to China, and the other that grouped accessions native to China. Alternatively AMOVA fixation index value and gene flow value all indicated the very low genetic differentiation between the two groups. This study illustrated that SSR markers are still a very useful tool for genotyping sugarcane.

The genetic structure of Brazilian cultivars contrasted with a group of basic germplasm mainly composed of *S. officinarum* and *S. spontaneum* accessions was assessed through SSR, revealing that this technique is a fast and cost-effective way for routine cultivar identification and management of sugarcane germplasm banks. Authors detected that 97.6% of the cultivar alleles were found in the basic germplasm, while 42% of the basic germplasm alleles were absent in cultivars. Genetic divergence found between cultivars and *S. spontaneum* accessions has practical applications for energy cane-breeding programs because the choice of more divergent parents will maximize the frequency of transgressive individuals in the progeny (Manechini et al. 2018). In another study, Crystian et al. (2018) analyze the genetic base of the Serra do Ouro germplasm bank in Brazil over the past decades using SSR. Results revealed an increase in genetic similarity in the 1970s due to interbreeding with few parents; however, after the 1990s, with the introduction of new parents, there was a decrease in genetic similarity levels.

Even more, SSRs were recently used to characterize the world collection of sugarcane and related grasses maintained at Miami, Florida, U.S.A. that contains nonredundant clones of different species and genera within the *Saccharum* complex. Results revealed average gene diversity of 0.208 among all species groups, where *Erianthus*-like *Saccharum* species, *Miscanthus* spp., and *S. spontaneum* each formed a distinct cluster, *S. robustum*, *S. officinarum*, hybrid cultivars, and *S. edule* grouped together in a major cluster, and *S. sinense* and *S. barberi* formed a distinct grouping (Fickett et al. 2020).

The aforementioned information reflected that among molecular markers, SSRs are advantageous for a variety of applications in sugarcane genetics research and breeding (genetic fingerprinting, kinship analysis, and population genetic studies). Their benefits are being multi-allele, highly informative, having high relative abundance, providing good coverage across the genome, and being

experimentally reproducible. In addition, more recent evidence suggests they may play an important role in genome evolution and provide hotspots of recombination. SSRs also have a few advantages even over markers obtained using modern genotyping platforms, though they are low throughput and might cost less per unit (Ahmad et al. 2018).

Besides, inter simple sequence repeats (ISSR) primers designed from SSR regions were useful in clarifying the molecular classification of related *Saccharum* species and characterizing the genuine hybrid from *Saccharum* and *Erianthus* (Zhang et al. 2004).

Also related to SSR markers, 18 sugarcane specific sequence-tagged microsatellite site (STMS) markers developed from Genbank databases were used to analyze the genetic divergence of Indian genotypes. Results revealed that the genetic base of sugarcane germplasm used in varietal improvement in India is reasonably large (Hemaprabha et al. 2005). However, Sindhu et al. (2011) when characterizing cultivated sugarcane hybrids representing all agro-eco climatic regions of tropical India also by using STMS markers found low level of genetic diversity among the commercial hybrids under cultivation.

To determine the utility of single nucleotide polymorphisms (SNPs) as a way to improve the understanding of the complex sugarcane genome, Cordeiro et al. (2006) developed markers to a suite of SNPs identified in a list of sugarcane ESTs. This was the first instance of a technology that has provided an insight into the copy number of a specific gene locus in hybrid sugarcane. The identification of specific and numerous haplotypes/alleles present in a genotype provided the basis for identifying associations between specific alleles and phenotype and between allele dosage and phenotype in sugarcane.

Some authors have suggested that, for breeding purposes, the estimation of genetic diversity should be based on functionally characterized genes or genes of interest since they reflect functional polymorphisms (Alwala et al. 2006). Variability of expressed or regulatory sequences might reflect a selection pressure for these regions, which would be different for each gene. Therefore, the variation detected by functional markers randomly distributed across the genome may detect variability more useful for breeders, since they reflect genetic variations of direct interest. Even more, they can be used to accelerate the breeding process and select important agronomic traits (Marconi et al. 2011). In that sense, target region amplified polimorphism (TRAP) markers amplify functional regions of the genome, providing valuable information related to the variations within coding/functional regions of the plant genome (Li and Quiros, 2001; Singh et al. 2017). TRAP with primers designed from sucrose and cold-tolerance-related EST were used to determine the genetic diversity and interrelationships in sugarcane germplasm collections from the genera *Saccharum* (*S. officinarum*, *S. barberi*, *S. sinense* R, *S. spontaneum*, *S. robustum,* and interspecific hybrids between *S. officinarum* and *S. spontaneum*), *Miscanthus,* and *Erianthus*. Cluster and principal coordinate (PCoA) analyses placed the *Erianthus* spp. and *Miscanthus* spp. genotypes distinctly from each other and from the *Saccharum* species, supporting their taxonomic classification as separate genera. Genotypes of the low sucrose and cold tolerant species, *S. spontaneum*, formed one distinct group, while the rest of the *Saccharum* species formed one interrelated cluster with no distinct subgroups. Results ratify TRAP as a potentially useful marker technique for genetic diversity studies in sugarcane (Alwala et al. 2006). TRAP markers were also used to estimate the genetic similarity among sugarcane varieties and five species of the *Saccharum* complex. Fixed primers designed from candidate genes involved in sucrose metabolism and drought-response metabolism were used. Results suggested that genetic variability in the evaluated genes was lower in the sucrose metabolism genes than in the drought-response metabolism ones (Creste et al. 2009).

Besides, TRAP revealed a higher polymorphism than AFLP markers, indicating that TRAP markers are more effective than AFLP when analyzing the sugarcane genome because they are easier, faster, and more economic than applying AFLP markers (Racedo 2014). In addition, Singh et al. (2017) used TRAP markers to identify promising genotypes bearing enviable agronomic traits (sucrose content and multiple disease resistance) to recommend as proven parents. They

found through poly component analysis that 52.48% of the cumulative variation was explained by all the genotypes with respect to the sucrose metabolism-related genes. Moreover, authors reconfirmed that functional TRAP markers would be efficiently useful in genetic studies for sugarcane genetic improvement.

In 2009, Que et al. selected genes related to sucrose metabolism and cold stress to study genetic diversity of important sugarcane clones, including main cultivars, recently released varieties, and germplasms introduced abroad recently through TRAP markers. Although cluster analysis could not contribute completely to parentage grouping of the clones, it could really reflect the variation degree of these clones. Recently, Medeiros et al. (2020) employed TRAP markers to evaluate several accessions of the *Saccharum* complex; they showed that even when considering genes under the selection process, the introgression of new favorable alleles is still possible. Also, TRAP markers revealed that both parental species contributed significantly to the variability in genes involved in sugar metabolism, and considerable variability in the lignin metabolism genes remains underexplored by local Brazilian breeding programs (Junior et al. 2020).

In addition, the potential of the sequence-related amplified polymorphism (SRAP) technique, which preferentially amplifies gene-rich regions, was evaluated to assess the genetic relationships among members of the *Saccharum* species (*S. officinarum*, *S. spontaneum*, *S. robustum*, *S. sinense*, *S. barberi,* and sugarcane hybrids); *Miscanthus* and *Erianthus* were included as outgroups. Dendrogram grouped the genotypes according to their phylogenetic relationships. *Erianthus* and *Miscanthus* were separated as two outgroups, where one clustered *S. robustum*, *S. sinense*, *S. barberi,* and hybrids as different subgroups with each one, including some *S. officinarum* clones, while the second cluster included the *S. spontaneum* clones, exclusively. The high discriminatory power coupled with the possibility that most of the amplicons amplified gene-rich regions of the genome makes SRAP a potentially robust tool for genetic mapping aimed at marker-assisted introgression in sugarcane (Suman et al. 2008).

In another study, the genetic relationships of sugarcane clones were determined based on three types of molecular markers; AFLP, TRAP, and SSR. Although the average genetic similarity estimates based on AFLP (0,675) and TRAP (0,655) was closer to each other than to SSRs (0,522), the correlation between TRAP and SSRs was higher (r = 0.24). Authors suggested that the choice of molecular markers should be considered carefully, based on the purpose of the application in the breeding program, since it is not possible to select a marker system that fits all the requirements for germplasm characterization (Creste et al. 2010). Perera et al. (2012a) also compared AFLP, SSR, and morphological traits to evaluate genetic diversity among the most useful parents of an Argentinean sugarcane breeding program. When morphological data was combined with molecular markers, the dendrogram obtained revealed genetic relationships, and the genetic diversity was better estimated. Authors also suggested that both methods appear to be useful, complementing each other, and should be used together to assist sugarcane breeders in estimating genetic diversity, electing parents for crossings, identifying superior lines, and protecting intellectual property rights. It must be highlighted that, in most cases, a combination of different marker techniques is recommendable for genetic diversity studies since they could reflect different aspects of the genome studied and provide more accurate information (Perera et al. 2016).

EST-SSR markers are preferred to genomic SSR (gSSR) since they represent the functional part of the genome, which can be easily associated with the desired trait. The database of sugarcane EST (SUCEST) offers a great opportunity for developing molecular markers that are directly associated with important agronomic traits. Marconi et al. (2011) employed this database and developed 287 EST-SSRs polymorphic markers that were tested in 18 sugarcane genotypes. Among the markers characterized in this study, some have particular interest, such us those that are related to bacterial-defense responses, generation of precursor metabolites and energy, and involved in carbohydrate metabolic process. Besides, Singh et al. (2013) employed EST-SSR for

studying the population structure of sugarcane accessions. The authors found a high level of polymorphism; this not only suggested their utility in sugarcane genetics and genomics, but also enriched the microsatellite marker resources in sugarcane.

Collard and Mackill (2009) developed another DNA marker in rice, the start codon targeted (SCoT) marker. It is based on the short conserved nucleotide sequence that flanks the start codon ATG. SCoT marker involves a single oligonucleotide primer, is PCR based, and, due to the simultaneous binding of the primer on both DNA strands, the sequence between the two binding sites is amplified. SCoT marker presents the following advantages: simple, low-cost, highly polymorphic, gene-targeted, and abundant in the genome (Que et al. 2014). This marker technology was applied in sugarcane to study genetic diversity among 107 sugarcane accessions within a chinese sugarcane germplasm collection. PCA partitioned the 107 sugarcane accessions into two major groups, the domestic group and the foreign introduction group. Each group was further divided based on institutions, where the sugarcane accessions were originally developed. Results provided foundation data for managing sugarcane germplasm, including construction of a core collection and regional variety distribution (Que et al. 2014).

Besides, new technologies are now available. They include high-throughput next-generation sequencing (NGS), currently used to detect SNPs, that have permitted the accumulation of large quantities of molecular data through the analysis of marker regions to make estimates of genetic diversity extremely robust (Goodwin et al. 2016). Although there is a long history of research and development in sugarcane, it is only comparatively recently that marker technology has caught up with the need for discovery of large numbers of single-dose SNP markers; fast, efficient, and cost-effective methods for genotyping can be used for selection in a breeding program (Aitken 2021). In that sense, diversity arrays technology (DArT) provides whole-genome profiling for hundreds to thousands of polymorphic markers in a single assay using a high-throughput microarray platform. However, due to the genome complexity of this sugar-producing crop of high economic importance, an application of DArT genotyping to this species required extensive testing and optimization. The method employs genome complexity reduction using methylation-sensitive restriction enzymes and high-throughput next-generation sequencing (NGS). The sugarcane DArT markers generated with this method identified high genetic differentiation of sugarcane ancestral species from modern cultivars, in agreement with the data available for other types of molecular markers for this crop. Even more, the majority of sugarcane DArT markers segregated in a Mendelian. When 384 clones were sequenced, one-third of sequenced markers came from the transcribed portion of the sugarcane genome. Results obtained validated the potential of DArT technology in providing cost-effective genetic profiles for plants, irrespective of their genome complexity, for effective applications in molecular-assisted breeding, diversity analysis, or genetic identity testing (Heller-Uszynska et al. 2011). In a recent study, the genetic diversity of local and foreign sugarcane varieties selected from the local germplasm in the Philippines was evaluated through SNPs and silico-DArT dominant markers. Cluster analysis and principal coordinate analysis showed low-sucrose varieties grouping separately from commercial hybrids developed in the Philippines. However, genetic distances among varieties genotyped indicate moderate to high genetic relatedness within the local germplasm, especially among commercially-available varieties in the country. Authors suggested that DArT-Seq genotyping was successfully used in the analysis of genetic diversity among current commercial varieties and can be a useful tool in evaluating new breeding materials for the development of more improved varieties (Bello et al. 2019).

To summarize the importance of genetic diversity studies, it must be mentioned that conserving crop biodiversity is an urgent undertaking. Modern cultivars have limited genetic variation (Fickett et al. 2020) since almost 20 *S. officinarum* clones are involved in the genealogy of sugarcane cultivars with only a few being used extensively (Raboin et al. 2008). According to the Intergovernmental Panel on Climate Change, 25–30% of plant species, including crops, will be extinct or endangered in the next century. Hence, safeguarding the existing diversity and producing new

diversity in cash crops are of utmost importance to keeping the environment intact (Priyadarshan and Mohan Jain 2022).

In conclusion, different molecular marker systems were useful in studying the genetic diversity of populations of diverse origin, whether wild, domesticated, or modern varieties. Unlike the estimates of diversity made from geographic, pedigree, or morphological data, the genetic parameters of diversity calculated from molecular data that reflect variations at DNA level should better reveal the real relationships. However, it is important to note that a sufficient number of molecular tags distributed throughout the genome should be used, and ideally, the information from different molecular marker systems should be combined to more accurately estimate the diversity present in each population under study.

1.3 MOLECULAR TECHNIQUES TO DETECT SYSTEMIC DISEASES DURING QUARANTINE AND TISSUE CULTURE TO SANITIZE DISEASED PLANTS

Probably due to the genetic bottleneck effect that occurred during the first interspecific hybridizations that gave rise to modern cultivars, the genetic diversity of sugarcane is low. This situation forces the breeding programs to have a germplasm bank that allows them to fully exploit the existing allelic diversity and create new combinations from it. For this process, the exchange of materials between different regions and programs is essential. To effectively use the new genetic resources incorporated, either to introduce them to the germplasm bank or to test them as varieties to be released, it is imperative to subject them to a period of quarantine. During this stage, the plants are inspected and screened for diseases to determine their sanitary status and avoid the introduction of pathogens. The available sugarcane phytopathogen-detection techniques, described below, are varied and also differ in their sensitivity. If the presence of diseases is detected in these imported materials that might have valuable traits for the program, the meristem culture technique offers the possibility of sanitizing them effectively.

Actually, more than 100 pathogens (including bacteria, fungi, viruses, phytoplasmas, and nematodes) have been recognized as causal agents of diseases in sugarcane (Govindaraju et al. 2019). Propagation of sugarcane through vegetative cuttings favors the spread of diseases through planting materials. Most of the breeding programs focus on obtaining varieties resistant to diseases since severe epidemics have threatened sugarcane production throughout its history, and many sugarcane varieties were replaced due to their breakdown to a new disease or to a new pathogenic strain (Viswanathan and Rao 2011).

In general, in modern sugarcane production systems, diseases are predominantly controlled by an integrated approach involving the combination of disease-free planting material obtained by tissue culture, disease-resistant cultivars, applicable farm-management practices, and strict quarantine measures for exchanging foreign materials (Babu et al. 2021). However, to perform any management strategy, a correct diagnosis of the pathogen involved is required. In that sense, regarding available diagnosis techniques, serological and molecular methods are commonly employed.

Among serological methods, ELISA and tissue-blot (or tissue-printing) techniques have become very popular and largely used for routine detection of bacterial, phytoplasma, and viral pathogens in sugarcane. These tools are only specific and sensitive enough to detect moderate to high pathogen titters within infected plants (Amata et al. 2016). However, molecular techniques such as polymerase chain reaction (PCR) and reverse-transcriptase (RT)-PCR assays are more sensitive than the serological techniques, specially when detecting sugarcane viruses and phytoplasmas occurring at low concentrations (Viswanathan and Rao 2011). Besides, PCR offers several advantages over other methods, including versatility, relative simplicity, and specificity, which can be increased by a two-step PCR (nested PCR). Due to the aforementioned reasons, PCR and RT-PCT-based diagnostics are recommended for evaluating quarantine materials and plants obtained through tissue culture. It must be mentioned that most of the available diagnosis protocols for

bacteria detection rely on internal transcribed sequences (ITS) regions, while for viruses, the capsid protein is the most commonly used.

In addition, multiplex PCR or RT-PCR allowed, in a single reaction with different pairs of specific primers, the simultaneous detection of various pathogens. This kind of reaction was found to be equally efficient to uniplex-PCR or RT-PCR to amplify the target pathogens from sugarcane (Iglesia et al. 2003) and is very useful in quarantine programs since it saves time and resources.

As PCR and RT-PCR are gel-based systems, they could be laborious and time-demanding for routine diagnosis of a large number of samples. In that sense, quantitative real-time PCR and RT-PCR are gel-free approaches in molecular diagnostics, enabling a sensitive, specific, and fast detection of sugarcane pathogens at very low titters. These techniques could be 100-fold more sensitive than conventional PCR or RT-PCR (Korimbocus et al. 2002).

Quantitative real-time PCR requires using costly instruments that may be unaffordable for some laboratories. To solve this problem, Notomi et al. (2000) developed novel nucleic acid amplification assays termed loop-mediated isothermal amplification (LAMP). This method utilizes auto-cyclic strand displacement at a single, constant temperature in a rapid, simple, and low-cost process. Generally, four to six primers are designed to recognize several regions of the target sequence, which gives LAMP its high specificity; even more, several authors found that LAMP is as sensitive as, or more sensitive than, quantitative real-time PCR assays (Keizerweerd et al. 2015). This diagnosis method was used for detecting bacteria (Liu et al. 2013) and viruses (Keizerweerd et al. 2015) in sugarcane.

A highly sensitive and suitable method for the robust detection of viruses in complex plant tissues, such as sugarcane, is the nucleic acid sequence-based amplification (NASBA) of the target RNA with specific primers and simultaneous real-time detection of the amplification products with molecular beacons (MBs) in a closed tube, called AmpliDet RNA (Gonçalves et al. 2001). MBs are short singled-stranded nucleic acid sequences (30–50 bases) that are designed to have a unique sequence flanked by indirect repeats so that a stem-loop structure is formed.

On the other hand, the detection of pathogens by nucleic acid hybridization tests is not as useful since, many times, it revealed background in healthy plants due to pigment interference. Also, the use of radioactive probes is not always feasible because of the lack of facilities and other inherent limitations in several laboratories (Hema et al. 2003).

Other molecular-diagnostic techniques to detect viruses from leaf extracts and sugarcane juice are immunocapture-reverse transcription-PCR (IC-RT-PCR) and direct-binding RT-PCR (DB-RT-PCR) (Hema et al. 2003). In the first one, viruses are captured with its specific antiserum in antibody-coated tubes and then the RT-PCR is performed, whereas DB-RT-PCR does not require the use of antiserum since it is a nonspecific immobilization system where viruses bound to a solid support can be directly detected by RT-PCR. The second diagnosis method is at least 10-fold less sensitive than IC-RT-PCR.

To provide examples of available diagnostic molecular protocols for detecting the most common bacteria, fungi, and viruses in sugarcane, Table 1.1 is included.

In brief, quarantine is a valuable resource for obtaining foreign material to broaden the genetic diversity of the breeding program. A correct detection and identification of the pathogens present in disease plants avoid introducing and spreading the pathogen in the region. Additionally, due to the difficulties in managing diseases, especially in the case of vegetatively propagated crops like sugarcane, where the planting material is the primary source of the disease in the field, the use of disease-free plants obtained trough tissue culture is recommended to the farmers. In most if not all countries where disease-free sugarcane plants are obtained, molecular diagnosis are routinely applied to ensure the health of the seedlings obtained. The choice of the most appropriate diagnostic method will depend on several factors including sensitivity, number of samples, cost, speed, available equipment, among others. It must be mentioned, that further works are required to develop sensitive diagnostics for field level detection of pathogens to address problems in the field (Viswanathan et al. 2018).

TABLE 1.1
Examples of Available Molecular Diagnosis Protocols for Causal Agents of Sugarcane Diseases

Method	Causal Agent	Disease	Reference
PCR	*Leifsonia xyli susbp. Xyly*	Ratoon stunting disease (RSD)	Pan et al. (1998); Sawazaki et al. (2013)
	Xanthomonas albilineans	Leaf scald (LS)	Fegan et al. (1998); Pan et al. (1997; 1999) Wang et al. (1999)
	Acidovorax avenae subsp. avenae	Red stripe	Song et al. (2003)
	Puccinia melanochepala	Brown rust	Virtudazo et al. (2001) Glynn et al. (2010)
	Puccinia kuehni	Orange rust	Virtudazo et al. (2001) Glynn et al. (2010) Sawazaki et al. (2010a,b, 2013);
	Macruropyxis fulva sp. nov.	Tawny rust	Martin et al. (2015)
	Sporisorium scitaminea	Smut	Sawazaki et al. (2013)
Multiplex PCR	*L. xyli* susbp. X*yly* and *X. albilineans*	RSD and LS	Iglesia et al. (2003)
Nested PCR	*L. xyli* susbp. X*yly*	RSD	Lozano et al. (2003)
Multiplex nested PCR	*L. xyli* susbp. *Xyly* and *X. albilineans*	RSD and LS	Davis et al. (1998)
RT-PCR	SCMV	Sugarcane mosaic virus	Yang and Mirkov (1997); Alegria et al. (2003); Viswanathan et al. (2007); Singh et al. (2009)
	SrMV	Sorghum mosaic virus	Yang and Mirkov (1997)
	Fijivirus	Fiji disease virus	Smith and van de Velde (1994)
	SCSMV	Sugarcane streak mosaic virus	Chatenet et al. (2005); Viswanathan et al. (2007; 2008a); Singh et al. (2009)
	SCYLV	Sugarcane yellow leaf virus	Pan (1997); Angel et al. (2001); Chatenet et al. (2001); Gonçalves et al. (2001); Viswanathan et al. (2008c); Singh et al. (2009); Girard et al. (2012)
	SCBV	Sugarcane bacilliform virus	Braithwaite et al. (1995); Singh et al. (2009); Balan et al. (2020)
Multiplex RT-PCR	Fijivirus, SCMV and SCYLV	Fiji disease virus, Sugarcane mosaic virus and Sugarcane yellow leaf virus	Cadavid et al. (2003)
	SCMV, SrMV, SCSMV and SCYLV	Sugarcane mosaic virus, Sorghum mosaic virus, Sugarcane streak mosaic virus and Sugarcane yellow leaf virus	Xie et al. (2009)
	SCMV, SCSMV and SCYLS	Sugarcane mosaic virus, Sugarcane streak mosaic virus and Sugarcane yellow leaf virus	Viswanathan et al. (2010)

TABLE 1.1 (Continued)
Examples of Available Molecular Diagnosis Protocols for Causal Agents of Sugarcane Diseases

Method	Causal Agent	Disease	Reference
DB-RT-PCR	SCSMV	Sugarcane streak mosaic virus	Hema et al. (2003)
IC-RT-PCR	SCSMV	Sugarcane streak mosaic virus	Hema et al. (2003)
	SCBV	Sugarcane bacilliform virus	Matsuoka et al. (2014)
	SrMV	Sorghum mosaic virus	Chen et al. (2020)
Multiplex IC-RT-PCR	SCMV and SCSMV	Sugarcane mosaic virus and Sugarcane streak mosaic virus	Viswanathan et al. (2008b; 2013)
qPCR	*L. xyli* susbp. X*yly*	RSD	Grisham et al. (2007); Pelosi et al. (2013); Carvalho et al. (2016); Fu et al. (2016); Johnson (2019)
	X. albilineans	LS	Garces et al. (2014); Gutierrez et al. (2016); Wang et al. (2020); Shi et al. (2021)
	SCWL	Sugarcane white leaf phytoplasma	Roddee et al. (2018)
qRT-PCR	SCYLV	Sugarcane yellow leaf virus	Korimbocus et al. (2002); Amata et al. (2016)
	SCSMV	Sugarcane streak mosaic virus	Fu et al. (2015); Wei-Lin et al. (2015)
NASBA	SCYLV	Sugarcane yellow leaf virus	Gonçalves et al. (2001)
LAMP	*L. xyli* susbp. X*yly*	RSD	Liu et al. (2013)
RT-LAMP	SCMV and SrMV	Sugarcane mosaic virus and Sorghum mosaic virus	Keizerweerd et al. (2015); Viswanathan et al. (2018)

1.4 GENETIC VARIABILITY GENERATION THROUGH HYBRIDIZATION

1.4.1 Selection of Parents

A key factor that determines genetic gain and the success of any breeding program is the selection of parents to carry out the crosses that will generate the clones subjected to rigorous field evaluations during the following 10–12 years (Sanghera and Jamwal 2019, Atkin et al. 2009). However, there is evidence that the selection of parents has received little attention in terms of assistance with molecular markers compared with other stages of the sugarcane improvement scheme (Jackson 2016).

The potential of an individual as a parent can be estimated through its breeding value (BV), which is a measure of the individual's ability to transmit its characteristics to its progeny. BV is affected directly by the estimation of variance components, especially the estimate of additive variance (Falconer and Mackay 1996). At present, the restricted maximum likelihood (REML) and best linear unbiased prediction (BLUP) (Piepho et al. 2008) statistical methods are the most used to obtain BV estimates (Barbosa et al. 2005, Neto et al. 2013, Mbuma et al. 2019). In addition to the data from field trials, there is enough evidence that the incorporation of pedigree information from the coefficient-of-coancestry (pedigree, kinship, consanguinity) matrix has improved the precision of BV estimates in sugarcane (Alwala 2007, Piepho et al. 2008, Atkin et al. 2009, Mendes de Paula et al. 2020). To incorporate genetic correlations between individuals improves the statistical model for estimation of the variance components and BVs. Although pedigree provides extremely valuable information, sometimes it does not necessarily represent the real correlation between the individuals under study

since the pedigree records are not always complete and / or may have errors. In addition, the biology of crop reproduction, which presents asymmetric distribution and preferential chromosome pairing in meiosis, added to structural chromosomal differences and frequent aneuploidy (Jannoo et al., 2004), means that not all the assumptions are fulfilled to estimate the kinship coefficients (Lima et al. 2002), which can result in biased estimates of additive variance and BVs (Atkin et al. 2009, Viana et al., 2012). The use of sufficient genotypic data obtained from molecular markers can provide accurate information on shared alleles, better reflecting genetic correlations, and improving estimates of additive variance and genetic merit of individuals (VanRaden 2008, Costa 2015). Few studies in sugarcane incorporated molecular data to improve the estimates of BVs (Jackson 2016, Wei and Jackson, 2017), until the recent development of genomic selection (detailed below) in which high throughput genomic information is incorporated to phenotypic data and appropriate statistical models to accurately estimate BV.

The choice of parents is not an easy task since numerous traits of economic importance are desired to be improved at the same time, and they do not have the same genetic architecture. For example, tons of cane per hectare (TCH) and tons of sugar per hectare (TSH) have been reported as low heritability traits, while Brix presents intermediate heritability (Hoarau et al. 2021, Jackson 2016, Neto et al. 2013; Pisaroglo de Carvalho et al. 2014; Jackson 2018; Wei et al. 2017). This implies that for TCH and TSH, the nonadditive variance components are of great importance (Zhou 2021). In this case, the BVs of the parents will not be very informative about the success of the crosses and the superiority of the clones in the progeny since little of the genetic variance is due to the additive components. This outcome would explain to some extent the low rate of genetic gain obtained in sugarcane breeding compared to other cross-pollinated crops such as maize (Burnquist 2013); traditionally, breeding programs have mostly exploited the additive variance considering BV or general combining aptitude (GCA) of the parents (Zhou 2021). Some recent studies have proposed the exploitation of specific combining aptitude (SCA) and heterosis to exploit the nonadditive genetic effects of important traits (Mendes de Paula 2020, Zhou 2021). This breeding strategy could be assisted by the application of genomics to detect polymorphic molecular markers, which allow the identification of progenitors with different allelic complements to maximize heterosis or hybrid vigour. Some studies have focused on determining genetic distances from different molecular markers systems, suggesting that performing crosses between distant parents would lead to obtaining superior progeny, increasing the genetic gain (Alwala 2007, Lavanya and Hemaprabha 2010, Singh et al. 2017, Wang et al. 2020, Zhou 2021).

In summary, the parent's collection to generate the variability that will be selected for more than a decade is dynamic. New parents must be incorporated while others are discarded so that the available genetic variability does not decrease due to selection. The crosses must be designed to increase the rate of genetic gain; for this, it is essential to know the genetic architecture of the characteristics under selection and the genetic constitution of parents. The precision of estimating BVs can be improved by incorporating genetic information from high-throughput molecular markers. In addition, parents must be studied in their SCA and to exploit the heterosis. For the latter, the aim is to maximize the genetic distance of the parents in a cross, for which the molecular markers distributed throughout the genome provide precise information on the identity between the alleles present in each parent.

1.4.2 Hybridity Determination and Progeny Test Assisted by Molecular Markers

Once progenitors have been chosen, biparental or polycrosses are carried out with genotypes that have enough flowers. At hybridization season, molecular markers are useful tools for the following:

- To verify the hybrid condition of the seeds obtained in a cross since it is very useful to estimate the percentage of intentional paternal hybridizations to better understand and possibly improve the crossing process. As sugarcane inflorescences are hermaphrodite panicles, male or female parental genotypes are classified based on the relative amount of