Chapter 2
Genomic Designing for Biotic Stress Resistant Cocoa Tree

Fredson dos Santos Menezes, Jonathan Javier Mucherino-Muñoz, Cláusio Antônio Ferreira, Saulo Fabrício da Silva Chaves, Ceslaine Barbosa, Livia Santos Lima Lemos, Natalia Jordana, Mariana Carvalho, José Luis Pires, Raner José Santana Silva, Karina Peres Gramacho, Rafael Moysés Alves, Ronan Xavier Corrêa, and Fabienne Micheli

Abstract Cocoa tree (*Theobroma cacao* L.) is cultivated mainly in tropical regions and produces beans that are used for chocolate manufacture. Worldwide, cocoa bean production is threatened by biotic stresses, mainly fungus, oomycetes, virus and other pests. The understanding of the determinism of the plant-pathogen interactions as well as the different and integrated ways to manage the cocoa diseases at field level began the focus of several research groups. Here, we did an overview of the several

F. dos Santos Menezes · J. J. Mucherino-Muñoz · C. A. Ferreira · R. J. S. Silva · R. X. Corrêa · F. Micheli (✉)
Centro de Biotecnologia e Genética, Universidade Estadual de Santa Cruz, Rodovia Ilhéus-Itabuna Km 16, Ilhéus, Bahia 45662-900, Brazil
e-mail: fabienne.micheli@cirad.fr

R. X. Corrêa
e-mail: ronanxc@uesc.br

J. J. Mucherino-Muñoz
Department of Forest Management, Facultad de Ciencias Forestales y Ambientales, Universidad de Los Andes, Mérida 5101, Venezuela

S. F. da Silva Chaves
Universidade Federal de Viçosa, Av. Peter Henry Rolfs, s/n, Campus Universitário, Viçosa, Minas Gerais 36570-900, Brazil

C. Barbosa · N. Jordana · M. Carvalho · J. L. Pires · K. P. Gramacho
Cocoa Research Center (CEPEC), CEPLAC, Km 22 Rod. Ilhéus-Itabuna, Ilhéus, Bahia 45600-970, Brazil

L. S. L. Lemos
Universidade Federal Do Sul da Bahia, Campus Paulo Freire, Praça Joana Angélica, 250, Bairro São José, Teixeiras de Freitas, Bahia 45988-058, Brazil
e-mail: livia.lemos@cpf.ufsb.edu.br

R. M. Alves
Embrapa Amazônia Oriental, Trav. Dr. Enéas Pinheiro, s/n, Bairro Marco, Caixa Postal 48, Belém, Pará 66095-903, Brazil
e-mail: rafael@cpatu.embrapa.br

F. Micheli
CIRAD, UMR AGAP, 34398 Montpellier, France

C. Kole (ed.), *Genomic Designing for Biotic Stress Resistant Technical Crops*,
https://doi.org/10.1007/978-3-031-09293-0_2

cocoa diseases, of the traditional breeding methods as well as the molecular assisted ones recently developed, of the molecular and omics resources currently available, and of the new biotechnology approaches—including genome edition and nanotechnologies—that are used at basic and applied research levels. We also described the main germplasm and collections worldwide as well as the use of the cocoa diversity as main source of disease resistance.

Keywords Biotechnology · Cocoa genes and genomes · Diversity · Plant diseases · Traditional and assisted cocoa breeding

2.1 Introduction

Cocoa tree (*Theobroma cacao* L.), the main source of raw material for the manufacture of chocolate, is a plant originated from the Upper Amazon Basin and currently cultivated mainly in tropical regions of the world, where it finds favourable conditions for its development (Gardea et al. 2017). It is a culture of high economic importance, especially for the African and American continents, which are responsible for about 75% and 17% of world production, respectively (https://www.icco.org/). At a social level, this agricultural activity generates jobs, directly and indirectly, throughout its entire production chain, thus ensuring a greater source of income, especially for smallholder farming (https://www.worldcocoafoundation.org/). According to the world region, cocoa tree also has relevant environmental importance, as observed in Latin America where it is cultivated under the Atlantic Forest covering in a system called *Cabruca*, which contributes to the maintenance of biodiversity and carbon sequestration (Franzen and Borgerhoff Mulder 2007; https://www.tnc.org.br/o-que-fazemos/nossas-iniciativas/cacau-floresta/).

Even its socioeconomical and environmental importance, the cocoa production has been highly threatened by biotic stresses, especially by phytopathogenic attacks such as *Phytophthora* ssp. (Kellam and Zentmyer 1986), *Moniliophthora perniciosa* (Aime and Phillips-Mora 2005), *M. roreri* (Phillips-Mora and Wilkinson 2007), *Ceratocystis cacaofunesta* (Engelbrecht et al. 2007a) and *Cacao swollen shoot virus* (CSSV) (Abrokwah et al. 2016), which, finding susceptible hosts and favourable environments, are able to establish compatible interactions, to trigger the characteristic symptoms of the corresponding diseases, and finally to drastically affect the productivity levels of this crop.

Aiming to contain or mitigate the damage caused by these pathogens, research has been intensively invested in supporting *T. cacao* breeding programs around the world (Bekele and Phillips-Mora 2019), with the main intentions of providing new knowledge to help in adoption of control measures based on genetic resistance. For this, a multidisciplinary knowledge is needed, ranging from classical techniques to modern and sophisticated molecular approaches, which were enhanced thanks to the sequencing of the genome of this species (Argout et al. 2011, 2017; Motamayor et al. 2013). The application of such methodologies, individually or together, mainly

provides knowledge of the genetic diversity and characterization of this species. This serves as a basis to assist in the selection of potential genotypes for resistance through classical breeding and genome-wide association study (GWAS), that can be recommended as crop varieties or as parents for further establishment of promising hybrids (Osorio-Guarín et al. 2020; McElroy et al. 2018; Marita et al. 2001; Lanaud et al. 2004; Romero Navarro et al. 2017), as well as, through genetic engineering-based applications, to select genes potentially involved in triggering defence responses, which can be later used to increase the degree of resistance in a given genotype that already has other favourable characteristics for cultivation (Scotton et al. 2017; Maximova et al. 2006; Fister et al. 2018; Helliwell et al. 2016; Shi et al. 2010, 2013).

As mentioned above, to reach this final result, it is necessary to form a solid and diversified base of knowledge about the genetics and molecular biology of *T. cacao*, and this foundation was divided into topics that will be explained throughout this chapter.

2.2 Description on Different Biotic Stresses

The cocoa diseases with the greatest threat to world production are caused by the basidiomycete fungi of the genus *Moniliophthora.* These are *Moniliophthora roreri* (frosty pod rot) and *Moniliophthora perniciosa* (witches' broom). The third most important disease is black pod of cacao caused by *Phytophthora* ssp., which is practically ubiquitous to all cacao producing countries. Virus disease such as the *Cacao swollen shoot virus* (CSSV) disease, has reduced cacao production and is of particular concern in Africa. In Indonesia, *Ceratobasidium theobromae* is of economic importance. The vascular streak dieback or the *Sahlbergella singularis* (mirids) are also responsible for cocoa production damages, but with less intensity and/or in a geographically localized way. These diseases will be highlighted in this chapter (Table 2.1).

2.2.1 Frosty Pod Rot of Cocoa

The frosty pod rot of cocoa caused by *Moniliophthora roreri* is restricted to the western continent, and it is present in 11 countries of Tropical America. Brazil was free of the disease, but the report of the disease in July 2021 in a small city of Acre, western of north Brazil, is a call for the Brazilian cacao chain. The first official report of *Moniliophthora roreri* was in 1917 in Ecuador, but its centre of origin and distribution is Colombia (Phillips-Mora and Wilkinson 2007). Losses by the disease vary from 25 to 100% depending on climate conditions and plant genetic background, culminating in the field abandonment, having already caused serious socioeconomic problems (Bailey et al. 2018). In natural conditions, only the pods of the genera *Theobroma* and *Herrania* are affected by the disease.

Table 2.1 Information on the distribution of the most prevalent cacao diseases

Disease	Pathogen	Geographical spread
Virus diseases	*Cacao necrosis virus (CNV)*; *Genus Nepovirus*	Ghana, Nigeria
	Cacao swollen shoot virus; *Genus Badnavirus* (CSSV)	Benin, Côte d'Ivoire, Ghana, Liberia, Nigeria, Sierra Leone, Togo Reports also in Sri Lanka
	Cacao yellow mosaic virus; *Genus Badnavirus* (CYMV)	Sierra Leone
	Trinidad cocoa virus; Genus Badnavirus	Trinidad; isolated occurrences
Witches' broom disease	*Moniliophthora perniciosa*	Brazil (Bahia, Espirito Santo, Amazonian regions), Bolivia, Colombia, Dominican Republic, Ecuador, French Guiana, Grenada, Guyana, Panama, Peru, St. Lucia, St. Vincent, Suriname, Trinidad and Tobago, Venezuela
Moniliophthora pod rot (frosty pod rot or moniliasis disease)	*Moniliophthora roreri*	Belize, Bolivia, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Jamaica, Mexico, Nicaragua, Panama, Peru, and western Venezuela
Black pod rot	*Phytophthora* spp.	Most cocoa-producing countries worldwide
	P. citrophthora	Brazil, Cameroon, Costa Rica, Côte d'Ivoire, Dominican Republic, El Salvador, French Guiana, Guatemala, India, Indonesia, Jamaica, Mexico, Panama, Peru, Trinidad, Venezuela
	P. heveae	Brazil, Cameroon, Cuba, India, Malaysia, Mexico, Philippines
	P. megasperma	Brazil, Cuba, India, Malaysia, Venezuela, Philippines
	P. nicotianae var. *parasitica*	Brazil, Cuba, India, Malaysia, Philippines
	P. megakarya	Bioko (Fernando Po), Cameroon, Côte d'Ivoire, Gabon, Ghana, Nigeria, São Tomé and Principe, Togo

(continued)

Table 2.1 (continued)

Disease	Pathogen	Geographical spread
Ceratocystis wilt	*C. cacaofunesta*	Brazil, Cameroon, Colombia, Costa Rica, Ecuador, French Guiana, Trinidad, Venezuela
Vascular streak die-back	*Ceratobasidium theobromae*	Most cacao-growing areas in South and South East Asia: China (Hainan Island), India, Indonesia, West Malaysia and Sabah, Myanmar, PNG, (islands of New Guinea, New Britain, New Ireland), southern Philippines, Thailand, and Vietnam
Mealybug	Several species	All cacao-growing regions

M. roreri has a hemibiotrophic lifestyle with well-defined prolonged biotrophic phase. The spores, the only infective propagule, can infect fruits at any stage of development, but fruits are more susceptible during the first stages (45–60 days old); the older the pods lower the susceptibility. Thus, the pods are the primary sources of dissemination. Spore germination requires high humidity and an average temperature of 22 °C. Infection occurs mainly through the cuticle or stomata, first colonizing the apoplasm of the cortical parenchyma cells. After a prolonged incubation period that can last 45–90 days after the infection (Griffith et al. 2003), brown lesions are developed. The biotroph/necrotroph shift is coordinated with the shift from green to necrotic pods (Bailey et al. 2018). A snow-white pseudostroma and powdery spores of *M. roreri* develop rapidly on the pod surface following necrosis (about 3–10 days later). From then on, the pod becomes sporulating and necrotic (Phillips-Mora and Wilkinson 2007). The potential of inocula is very high, and one pod can produce between 44 million and 7 billion spores/cm^2 (Ram et al. 2004). The average temperature between 22 and 30 °C, air moisture higher than 80%, and available water, favour the germination and penetration of the spores (Evans 1981; Orea et al. 2017). Spores are long-lived, abundant during the dry season, and survive on crop residues. Thus, the dry season is a critical factor in determining *M. roreri* survival between harvests. It is a guarantee of inoculum sources available at the beginning of the wet season (Evans 1981). The spores are dispersed by wind and rain. But, human-mediated dispersal is the main form of long-distance dissemination.

M. roreri propagates mainly clonally (Bailey et al. 2018; Díaz-Valderrama and Aime 2016) and remains in a primarily haploid stage throughout its life cycle. Through the amplified fragment length polymorphism (AFLP) and inter-simple sequence repeat (ISSR) methodology, the genetic groups of *M. roreri* were determined and four main genetic groups were identified: Co-Central, Co-West, Bolivar, and Gileri. Corroborating to the *M. roreri* Colombian origin idea, four of the genetic groups were found in Colombia, with Co-East and Co-Central being apparently

endemic in this country. The Gileri group was the only group not found in Colombia and probably is exclusively found in Ecuador (Phillips-Mora 2003).

Control of frosty pod rot is done mainly by genetic control associated with good agronomic practices. Cultural control consists of the removal of potentially diseased pods right at the onset of necrotic lesion. In addition, cultural treatments such as mowing, threshing, pruning, fertilization of cacao trees, drainage, and treatment with anti-sporulates are efficient (Ram et al. 2004). Chemical control with protective fungicides, such as copper fungicides and oxychloride-based fungicides (Suarez Contreras and Rangel Riaño 2013), and systemic fungicides such as trizols and strobirulines is recommended to avoid build-up of epidemics. Biological control is carried out using antagonistic microorganisms such as bacteria of the genus *Pseudomonas* and *Bacillus* and fungi of the genus *Trichoderma* and *Clonostachys* (Krauss et al. 2003). The use of resistant varieties is also a promising alternative, but clones descending from Scavina (SCA), a reference for resistance to witches' broom disease, do not show resistance to frosty pod rot attack. However, some clones from Peru, Costa Rica, Colombia, and Ecuador show levels of resistance to the disease (Phillips-Mora et al. 2005).

2.2.2 Witches' Broom Disease

Witches' broom disease is the second most important threat of a cocoa plantation in South America. It's causal agent *Moniliophthora perniciosa* (Stahel) Aime Phillips-Mora (Aime and Phillips-Mora 2005), belongs to the phylum Basidiomycota, class Agaricomycetes, order Agaricales, family Marasmiaceae. Although the witches' broom disease was first identified in the Brazilian Amazon Forest in 1785, the disease was only officially mentioned in 1904, reporting the occurrence of the disease in Suriname in 1895 (Went 1904). In the following years, the disease spread rapidly and is currently distributed in the main cocoa-producing countries in tropical America such as Brazil, Bolivia, Ecuador, Colombia, Peru, Panama, Guyana, Grenada, Caribbean Islands, Trinidad and Tobago (Mondego et al. 2008; http://www.cabi.org/isc/datasheet/16054).

Moniliophthora perniciosa has co-evolved with cacao in the western Amazon basin in an area that includes eastern Ecuador (the Orient region), north eastern Peru and south central Colombia (Pound 1938; Clement et al. 2010). Thus, the *Malvaceae, Theobroma,* and *Herrania* species are its primary hosts; the economical ones being cocoa (*Theobroma cacao* L.) and cupuaçu (*T. grandiflorum).* However, after the witches' broom disease introduction in Bahia, in 1989, several alternative hosts belonging to *Solanaceae, Malpighiaceae*, *Bignoneaceae*, *Bixaceae* have been reported as *M. perniciosa* hosts (Griffith and Hedger 1994a, b). However, this is still debatable if these species are indeed *M. perniciosa* hosts as there is a lack of pathogenicity tests to prove these as host's range of *M. perniciosa.* To our knowledge, Brazil harbors the greatest diversity of hosts ever recorded, with 29 wild and cultivated species (Lisboa et al. 2020; Patrocínio et al. 2017). If these species are proven as

primary hosts of *M. perniciosa,* the natural distribution of this fungus will be probably one of the largest ever described for an obligate plant pathogen. A phylogenetic analysis and pathogenicity studies support the hypothesis of specificity of the fungus to the host of origin.

Studies on fungal populations showed that *M. perniciosa* has high genetic variability, with populations varying according to country of origin, and host-specificity. Thus, there are differences in pathogenicity and genetic diversity among Peruvian, Ecuadorian, and Brazilian isolates. Further studies in Brazil showed that *M. perniciosa* strains harvested from cocoa resistant selected varieties recommended for planting are genetically different from the strains collected from unselected local genotypes (Pires 2003). Also, temporal studies have shown a significant increase in disease severity in progenies of the primary source of resistance (Scavina) and varieties recommended to farmers and derived from that source (Pires 2003; Gramacho 2003).Samples of *M. perniciosa* isolated from resistant and susceptible cacao genotypes, differed genetically, thus indicating an evolutionary process towards the Scavinas resistant sources.

The *M. perniciosa* life cycle is characterized by the basidiocarp formation, in which basidiospore, the only infective propagule is produced. Basidiospores are very sensitive and short-lived. Germination occurs under high humidity and average temperatures of 22 °C. Spores are disseminated to the infection courts; any meristematic plant organs such as fruits, flower, and apical shoots, adhering to and penetrating the natural opening, through stomata, the base of trichomes, or directly through the cuticle (through enzymatic digestion) (Purdy and Schmidt 1996; Sena et al. 2014). At this phase, colonization is mainly in the apoplastic tissues, and 45–60 days later *M. perniciosa* shifts to intracellular colonization. In this stage, the mycelia are characterized by thin, dikaryotic hyphae with clamp connection and have a saprophytic life-style (Sena et al. 2014).

The symptoms of witches' broom disease depend on the type and stage of the tissue that is infected by *M. perniciosa.* Infected branches increased in thickness, and compared with uninfected branches, infected branches appear distinctly swollen. Large numbers of smaller branches or witches' brooms are often formed eventually on the infected branches. Symptoms appear 30–60 days after the infection. Afterward, they begin dry and about 90 days later until they reach complete dryness. Infected flower buds can produce several flowers and fruits stuck by the peduncle, forming the floral cushion brooms that eventually dry and produce fruits. Young fruits there may be from abortive parthenocarpic fruits (in the shape of strawberries and carrots) to black, hard, and irregular lesions (Silva et al. 2002). These symptoms can be so severe that they can lead to a loss of 90% of cocoa production in a region.

2.2.3 Black Pod Rot

The black pod is one of the main diseases of the cacao tree, being responsible for losses of 20–30% of the annual cocoa production, which can lead to the loss of 10%

of the plants (Bailey and Meinhardt 2016; de Oliveira and Luz 2005). It is caused by many species of *Phytophthora*—Kingdom Straminipyla, Phylum Oomycota, Class Oomycetes, Order Peronosporales, Family Phytiaceae—considered a pseudo-fungus with varied geographical distribution (Acebo-Guerrero et al. 2011; Ho 2018).

There are about 300 species of *Phytophthora* reported in the world, with only 7 causing black rot in cacao trees (Bailey and Meinhardt 2016; Ho 2018) and only four species have a commercial impact—*Phytophthora palmivora, Phytophthora megakarya, Phytophthora citrophthora*, and *Phytophthora tropicalis (P. capsici)* (Ho 2018; Bailey and Meinhardt 2016; End et al. 2017). All these species affect the plants with the same types of symptoms, differing in aggressiveness and geographic distribution. *Phytophthora megakarya* and *Phytophthora palmivora* are considered the most important for cacao cultivation in Central and West Africa and most studies are related to them (Guest 2007; Adomako 2007), however *P. megakarya* (considered the most aggressive) has been reported only in Africa and it is still in an invasive phase. Since *P. megakarya* is more aggressive and causes higher yield losses than *P. palmivora*, special care should be given when moving plant/soil materials and cocoa beans to production areas and countries that are not yet affected by *P. megakarya* (End et al. 2017). On the other hand, *P. palmivora* is the widest world distribution species in Africa, Asia, and America (Guest 2007).

High incidence levels of black pod disease have been reported in Brazil, being induced by *Phytophthora capsici*, *Phytophthora citrophthora*, *Phytophthora heveae*, and *Phytophthora palmivora* (Luz et al. 2001). Recently, the first description of the species *P. theobromicola* was reported, which seems to be prevalent in plantations in the state of Bahia, Brazil, inducing more extensive lesions in various genetic materials (Decloquement et al. 2021).

The pathogen's life cycle is divided into a sexual and asexual phase, the first being more commonly found in nature and the second marked by the formation of chlamydospores (vegetative structures) that germinate giving rise to sporangia that in turn release zoospores that can be dispersed over long distances by water and indirectly infect other plants, as opposed to the direct infection that occurs through the mycelium (Ho 2018; Luz et al. 2001; Oliveira et al. 2014; Kudjordjie 2015).

The most evident symptom of the disease in cacao is the formation of black pods. Initially, small dark spots appear on the pods approximately 30 h after infection. These symptoms appear in pods of any age, but in more developed pods, the beans can be totally or partially used, while in younger pods, spots, wrinkling, and darkening are common and may be confused with physiological withered. Although the pod is the most affected site, the disease can also affect leaves and roots that can serve as a source of inoculum, resulting in dark sunken lesions in the stem that often develops as a result of mycelium spread from pods into flower cushions and further along the stem or directly through wounds (End et al. 2017). In addition, extensive necrosis of leaves and shoots of seedlings, flower cushions, and root infections can occur, and these lesions can even cause the death of the plant (de Oliveira and Luz 2005; Luz et al. 2001; Silva Neto et al. 2001). The trunk can also be a target of infection due to the dispersion of pathogen structures, which can even cause the death of the plant (de Oliveira and Luz 2005; Luz et al. 2001; Silva Neto et al. 2001).

In this scenario, to prevent new pathogens from spreading or prevent them from multiplying further if they have already gained entry and have been established in new restricted areas are used legal enforcement of the quarantine measures. The plant parts that are likely to carry pathogen in trade and transport are: pods, roots, budwood, trunk/branches, leaves, and soil or growth media accompanying plants, while seeds originating from healthy pods are unlikely to carry the disease (End et al. 2017).

2.2.4 Ceratocystis Wilt of Cacao

The Ceratocystis wilt of the cocoa tree, also known as machete disease, is characterized by drying out and completely killing the cocoa tree. This disease is caused by Ascomycota *Ceratocystis cacaofunesta* Engelbrecht & T.C. Harr (Engelbrecht and Harrington 2005), a species within the Latin American clade of the *Ceratocystis fimbriata* species complex. *Ceratocystis cacaofunesta* belongs to the class Sordariomycetes, order Microascales, and family Ceratocystidaceae. The first report of *Ceratocystis* infecting cacao trees was in 1918 in Ecuador, where the disease was confined until the 1950s. In that same decade, the disease was reported in other countries in South and Central America (Silva et al. 2004; Rorer 1918).

When Ceratocystis wilt of cacao emerged in Ecuador there was not much damage to cacao production, as the country's cacao trees were resistant to this disease. However, in 1957 a more virulent form of the fungus was identified and since then it has affected many cacao trees in South American countries such as Brazil, Colombia, Guyana, Peru, Venezuela and, Trinidad and Tobago; in Central America such as Guatemala, Costa Rica, and Mexico and the Caribbean island of Haiti (Cabrera et al. 2016). Ceratocystis wilt is considered one of the most important emerging diseases of cacao because the agricultural practices of crop management (such as pruning and harvesting the fruits) favor the penetration of the fungus into the host tissue when the tools used are contaminated with the fungus (Engelbrecht et al. 2007a).

Species of the genus Ceratocystis have relatively low genetic diversity. This fact may be related to the type of reproduction of the fungus (homothalic), the limited flight range of the dispersing beetles and the introduction of the disease via human action, generally having few genetic variants in long-term dispersion (Engelbrecht et al. 2007b) report that populations in Colombia, Costa Rica and Bahia have limited genetic diversity, characteristic of introduced populations that have suffered recent genetic bottlenecks. While the populations of Rondônia (Brazil) and western Ecuador have similar diversity to populations of other species of the genus, indicating that they are natural populations.

The main route of colonization of *C. cacaofunesta* is through natural wounds on the surface of the host, which can be caused by beetles (carrying the conidia adhered to their paws) or contaminated tools. The conidia of this pathogen infect the plant's xylem parenchyma cells and move in a radial direction, invading the xylem vessels (Araujo et al. 2014). Conidia of *C. cacaofunesta* are very small in the early stages of colonization; this characteristic may be related to the rapid distribution of this

fungus in the plant and allows the fungus to more easily penetrate the host cell wall (Santos et al. 2013). The plant defence system, in turn, recognizes the pathogen, and a containment barrier against the fungus is formed, causing vascular occlusion. This barrier can favour both host resistance or pathogenesis, depending on the rate of occlusion formation. That is, if this obstruction prevents the spread of the pathogen, it will generate resistance to the host, but if this obstruction makes it difficult to transport water in many vessels, it may cause the plant to collapse (Talboys et al. 1972).

Ceratocystis cacaofunesta is a necrotrophic fungus, as such, it causes necrosis of the vascular parenchyma cells, which contributes to triggering the main symptoms of the disease, which is the wilt and dryness of the cocoa tree. Even after death, the leaves can still stick to the tree for weeks. In addition to these, there may be symptoms on the stem of the plant, presenting as brownish-brown lesions with lighter regions, seen in advanced stages of the disease. And in the roots, similar symptoms can manifest, which indicates that there is the transmission of the pathogen through the soil and contact between roots (Cabrera et al. 2016; Santos et al. 2013; Silva et al. 2004).

2.2.5 Cocoa Swollen Shoot Virus

Virus diseases are the most critical disease limiting cacao production in West Africa (Table 2.1), a region responsible for ~70% of the world's cocoa production. The *Cocoa swollen shoot virus* (CSSV) is one of the most destructive phytopathogens for the cocoa crop, especially in West Africa, where it causes significant economic damage (Padi et al. 2013). Marelli et al. (2019) accounted for an annual loss of 96,000 metric tons in production due to the virus. CSSV is considered endemic to West Africa and its importance in Malaysia and Sri Lanka was highlighted (Steven 1936; Geering and Hull 2012). Other viral diseases have been reported in Trinidad, but they were not associated with swellings. An attenuated and localized form of CSSV is reported in Indonesia, Sabah, and Sri Lanka (Muller 2016).

CSSV was reported in West Africa, Ghana, in 1936, but it was probably present since 1922 (Muller 2016). CSSV dissemination followed the cacao cultivation in the Eastern Region, and although few scattered outbreaks were observed in the Western region, in the 20th decade, this region became the most severely affected area in Ghana. After that, the disease has spread throughout the entire widely cacao-growing areas in West Africa, and nowadays, the disease is considered endemic to the Eastern Hemisphere.

Cacao swollen shoot virus disease is caused by a complex of badnaviral species (family, Caulimoviridae) referred to as pararetroviruses. Several different strains of the virus exist and can cause defoliation, dieback of the plant, and severe yield losses. Following the International Committee on Taxonomy of Viruses (ICTV), CSSV taxonomy is based on the nucleotide diversity in the RT/RNaseH region (Kouakou et al. 2012; Oro et al. 2012; Chingandu et al. 2017), and five different species, A, B–C,

D, E, and G have been described to cause CSSV disease. The virus is naturally transmitted to cocoa by at least 16 species of mealybugs (Hemiptera: Pseudococcidae) (Ameyaw et al. 2014; Hughes and Ollennu 1994), the vector of most badnaviruses. Particles are bacilliform and measure 121–130 × 28 nm. The virus does not multiply in the vector and is not transmitted to its progeny. CSSV can infect cacao at any stage of plant growth, and so far, there is no evidence of seed or pollen transmission. CSSV natural hosts are species from the Malvaceae families: *Adansonia digitata, Bombax* spp., *Ceiba pentandra, Cola chlamydantha, Cola gigantea, Theobroma cacao* and other tree species of the Malvaceae.

The mealybugs are the vectors responsible for spreading the disease over a short distance (radially) by crawling the canopy from tree to tree. New outbreaks are associated with jump spread over greater distances by wind-borne viruliferous mealybugs or by the very active small first instar nymphs (Strickland 1950; Thresh et al. 1988). Once a plant is infected, it cannot be cured. Like most plant viral diseases, the disease can be contained or prevented if healthy plants are isolated within barriers of CSSV-immune crops. The disease is not seed-borne, but it may be introduced in clones imported as plants or budwood.

Additionally, experimental pathogenicity tests with indigenous plants of West Africa also revealed the families Bombaceae, Sterculiaceae, and Malvaceae as potential alternative hosts of CSSV (Posnette et al. 1950). Several plant species have been reported as CSSV hosts, among those, 30 plant families are used as shade for cacao and other crops (Abrokwah et al. 2016; Friscina et al. 2017). Their geographical origin influenced phylogenetic relationships between Ghanaian and Togolese sequences rather than whether they originate from mild or severe isolates. CSSV populations have now been analysed molecularly in the major West African countries (Ivory Coast, Ghana, Nigeria, and Togo) (Oro et al. 2012; Kouakou et al. 2012), and six structural groups were proposed according to the diversity in the first part of open reading frames (*ORF3*) with high genetic variability within them.

It is admitted that no virus disease has been found in cocoa in South America. The high variability within CSSV populations and the combined knowledge of CSSV disease-badnavirus on molecular and pathogenicity studies and the historical data of the disease emergence have led researchers to hypothesize that CSSV in cocoa emerged from host jumps indigenous plants. Likely its introduction in Africa is due to several host shifts from indigenous hosts. Selective pressures in alternate host plants may induce a differential evolution of the virus compared to its evolution in cacao. This diversity is reflected in the differential preference of each species, as well as in symbiotic interactions with other organisms, making it difficult to control the insect and, therefore, the virus (Muller 2016; Ofori et al. 2015; Padi et al. 2013; Roivainen 1976).

Control measurements include preventive measures through the use of resistant material eradication of the infected tree and trees around it. Upper Amazon hybrids with good agronomic characteristics have been demonstrated more resistant to infection than other genotypes being grown (Amelonado type) (Amon-Armah et al. 2021).

2.2.6 Other Diseases and Pests

2.2.6.1 Vascular Streak Dieback

The vascular streak dieback, caused by *Ceratobasidium* (Oncobasidium) *theobromae*, is a disease of paramount importance to cacao orchards in Southeast Asia and Melanesia regions where the pathogen is endemic (Bekele and Phillips-Mora 2019; McMahon and Purwantara 2016). As the common name suggests, the symptoms of *C. theobromae* in cocoa progressively evolve, culminating in the plant's death. It can infect both adult and seedling individuals (Samuels et al. 2012). This basidiomycete causes considerable losses in the countries where it occurs. Marelli et al. (2019) estimated an annual loss due to disease of 61 thousand metric tons in almond production by 2016. Vascular streak dieback has been reported in a lot of cacao-growing areas as South and Southeast Asia, Melanesia, Kerala (India), Myanmar, Thailand, Hainan Island (China), Vietnam, Malaysia and Indonesia (Ploetz 2007).

The symptoms of vascular streak dieback initially begin with a chlorosis in the leaf, flushes behind the shoot apex and scattered islets of green tissue. The symptomatics leaves become fully chlorotic in a few days and the symptoms progressively develop in all leaves. Internally, the steams present infected xylem with dark streaks within the vascular tissue. During the wet weather, when an infected leaf falls, the hyphae can emerge from the leaf scar and develop into a tullasneloid basidiocarp. After that, it results in the formation of a white, flat, velvety covering of the leaf scar and bark. When it occurs on a dry day, the scar heals over and the fungus fails doesn't develop. Basidiospores keep viable for a few hours and require free water for the germ tube growth (Prior 1979).

The *C. theobromae* basidiospores are dispersed by wind and the effective spore dispersal is probably limited to a few hours in the early morning with high humidity and wet leaves (Keane 1981). In this way, just a few infections occur beyond 100 m from the diseased cacao.

The management of vascular streak dieback includes the use of resistant varieties. Crosses using Trinitarios clones have shown some resistance to the disease. Despite biocontrol of *C. theobromae*, treatment with *Trichoderma harzianum* T-22 suppressed vascular streak dieback development (Vanhove et al. 2016). Also, fungicides including flutriafol, hexaconazole, propiconazole, tebuconazole and triadimenol had been used to control vascular streak dieback but none of them showed to be commercially viable in cacao plantations (Holderness 1990).

2.2.6.2 Pests

Some pests of greater economic importance can be cited, such as *Distantiella theobroma*, *Monalonion bondaris*, *Selenothrips rubrocinctus*, *Conotrachelus humeropictus,* as well as some species of the genus *Helopeltis*. However, the bug *Sahlbergella singularis*, the cocoa borer, *Conopomorpha cramerella*, as well as the species of the

family Pseudococcidae, vectors of the CSSV, have been reported as causing the greatest losses in production (Bekele and Phillips-Mora 2019; Muller 2016).

2.3 Genetic Resources of Resistance Genes

Resistance genes to the main cocoa diseases have been sought almost exclusively in the species itself (primary gene pool), in wild or cultivated materials. The first technical reports of works aimed at this search are from the beginning of the 1930s, in Latin America and the Caribbean; and it is from collections conducted in Peru (Pound 1938) that the first genotypes that proved to be highly resistant to witches' broom disease were obtained (Bartley 1994). Since then, several collection expeditions were conducted in areas of natural dispersal of the species in Brazil, French Guyana, Ecuador and Peru (http://www.icgd.rdg.ac.uk), with a strong focus on resistance, especially to witches' broom disease. And, seeking resistance, countless collections and selection procedures in populations of traditional varieties and improved seminal varieties, and selection procedures in breeding cycles, have been conducted in many of the producing areas around the world—the International Cocoa Germplasm Database presents, in detail, the genotypes identified as resistant or promising for the main cocoa diseases and some pests. Currently, 49 national and international ex situ collections of germplasm participate in the preservation of the genetic resources of resistance genes of cocoa (http://www.icgd.rdg.ac.uk).

2.4 Glimpses on Classical Genetics and Traditional Breeding

2.4.1 Breeding Objectives

The breeding of cocoa, carried out by classical methods or by modern biotechnological techniques, plays a fundamental role in promoting the development of this species. The continuous technological advances obtained to promote the orchards' sustainability, earning a differentiated financial return, favours economic viability; and guarantees food security for both the producer and the consumer. Furthermore, it provides ecologically friendly farming, either by reducing the use of pesticides or reducing the use of chemical fertilizers (Dennis et al. 2008). Therefore, there are several goals to be achieved, which are dynamic, with synergistic or antagonistic interactions between them. They concern the productive, vegetative and health aspects of cocoa, and breeding will be challenged to develop genotypes that bring together favourable genes in these three aspects. Thus, the producer will be able to count on cultivars that will provide security, stability, productivity and longevity for field cultivation.

2.4.1.1 Productive Aspects

It aims to improve productivity and the qualitative characteristics of almonds, cacao's main product. Genotypes with a high number of fruits per plant, a high amount of almonds per fruit, and a high weight of almonds are sought (Lopes et al. 2011). These characters are related and can be expressed in a series of indexes such as fruit index, seed index, fruit value, among others (Bekele and Phillips-Mora 2019; Dias 2001), or even indexes that group other characters, such as production and resistance (Jaimez et al. 2020). The almond weights component, which makes up the seed index, tends to have less environmental influence, and consequently, greater heritability. However, it suffers interference from other factors, such as the length of the production cycle, the quantity and location of the almonds in the fruit (Doaré et al. 2020). Market demands exert pressure on breeding programs to also pay attention to the qualitative characteristics of almonds, which influence the quality of chocolate. Among them are the content and consistency of the butter, from the almond, as well as the nutritional and organoleptic characteristics themselves (Adeigbe et al. 2021; Araújo et al. 2009; Pinheiro et al. 2012).

2.4.1.2 Vegetative Aspects

It considers the plant's vigour, as well as the characteristics that facilitate the management, such as size, uniformity, precocity and genotypes' adaptability to different growing conditions, to which the cocoa tree can be subjected (Bekele and Phillips-Mora 2019; Mustiga et al. 2018). Particularly, for this last character, economic factors should be considered, such as the increase in planting density aiming at higher productivity, and/or plant in full sun (Olufemi et al. 2020); ecological, such as cultivation in agroforestry systems with different compositions and, consequently, levels of competition (Schneider et al. 2016); and climatic, due to the increasingly frequent changes as a result of global warming (Farrell et al. 2018).

2.4.1.3 Health Aspects

These refer to biotic stresses and will have a special focus in this chapter. Cocoa tree cultivation brings with it phytosanitary problems, among pests and diseases that, depending on the level of infestation, can drastically reduce the viability of the orchards. Therefore, breeding programs primarily aim at the development of genotypes with long-lasting resistance to these pathogens, preferably of the horizontal type, promoting adaptability and geographic and temporal stability against a range of phytopathogens simultaneously (Dias 2001; Bekele and Phillips-Mora 2019).

2.4.2 Classical Mapping Efforts

Markers are important tools for optimizing classical breeding. They provide fundamental information for the adequacy of the approach, as well as the selection methods used. In addition, they are essential in the characterization of genetic materials, as well as in the detection of genetic variability, raw material for breeding (Kordrostami and Rahimi 2015). In cocoa, several markers have already been developed and used as auxiliary to classical breeding. In chronological order of use, morphoagronomic, biochemical and DNA markers can be mentioned. The latter will be covered in the following sections of this chapter.

2.4.2.1 Morphoagronomic Markers

They constitute the basic tool of the cocoa breeder, being historically the most used in the crop breeding. In Active Germplasm Banks—AGB's, for example, the use of morphoagronomic markers is fundamental for the characterization and phenotypic distinction of accessions. These markers provide relevant information about the morphological characteristics of the vegetative and reproductive structures of each accession, to differentiate between them. They also refer to agronomic characteristics, such as production components and resistance to pests and diseases, allowing the efficient pre-selection of accesses to form working collections (Bekele et al. 2006; Dias 2001). Through simple morphological markers, it is possible, for example, to distinguish between the main cocoa tree ecotypes: Forastero, Criollo and Trinitário (Bekele et al. 2006, 2020; Bidot Martínez et al. 2017). As they are phenotypic, and, in most cases, related to polygenic characteristics, morphoagronomic markers suffer an expressive environmental influence, which interferes with heritability. However, its use is not unnecessary, having value as a complementary tool to genetic markers, increasing the reliability of information (Bidot Martínez et al. 2017; Gopaulchan et al. 2019). Robust information on morphological markers applied to germplasms can be found in the International Cocoa Germplasm Database (http://www.icgd.reading.ac.uk/icgd/; ICGD).

2.4.2.2 Biochemical Markers

Biochemical markers are less used in cocoa than morphological and DNA markers. Even so, this methodology has already been used for the assessment of genetic diversity and accession conservation strategy at the International Cocoa Genbank, Trinidad (ICGT), proving to be as effective as random amplified polymorphic DNA (RAPD)-type DNA markers, for example (Sounigo et al. 2005; Warren et al. 1995). Furthermore, these markers were used to study the pattern of self-compatibility and incompatibility in the species (Warren et al. 1995), as well as for making a gene linkage map (Lanaud 1986; Lanaud et al. 1995), fundamental for subsequent molecular studies.

2.4.3 Classical Breeding Achievements

Cocoa farming would not be at the technological level that it is today, were it not for the advances provided by research on genetic improvement. Only from the turn of the 20th to the twenty-first century, biotechnology became a reality in the species' improvement. Therefore, most of the results that were and are still being obtained, especially the development of cultivars, are the result of the classical approach, showing its importance (Bekele and Phillips-Mora 2019; Dias 2001). Below, the main advances provided by classical breeding are highlighted, with emphasis on resistance to biotic stresses.

2.4.3.1 Resistance to *Moniliophthora perniciosa*

The concern with *M. perniciosa* dates back to the initial stages of cocoa breeding, in the first half of the twentieth century, when pioneering prospecting was carried out in search of sources of resistance to the fungus (Pound 1943). The supposedly resistant SCA 6 and SCA 12 clones were material collected in these first expeditions (Dias 2001). These clones became, for a long period, the main source of resistance to the fungus, with clone SCA 6 still being classified in this way (Rodrigues et al. 2020). However, this resistance was overcome with the diversification of the pathogen in some countries such as Brazil, Peru and Ecuador (de Albuquerque et al. 2010; Gramacho et al. 2012; Lopes et al. 2011). This revealed the need to diversify the sources of resistance to the fungus, as a strategy to maintain the durability of resistance in the field.

Several studies were carried out to find clones and hybrid combinations with high resistance indexes, without, however, disregard the durability and geographic range. Currently, genetic materials that have resistance alleles to the pathogen, in addition to Scavina clones, are ICS (Imperial College Selection) 1, 60, 98, 45, 85 and 10; CAB (Cocoa from the Brazilian Amazon) 0371, 0388, 0392, 0410, 0169, 0352, 0214, 0208 and 0270; IMC (Iquitos Mixed Calabacillo) 67 and 47, PA (Parinari) 121, EET 272, POUND 18 and CC 41, as well as their hybrid combinations, such as the TSH (Trinidad Selected Hybrids) and TSA series (de Albuquerque et al. 2010; Bekele and Phillips-Mora 2019; Dias 2001; Benjamin et al. 2016).

2.4.3.2 Resistance to *Moniliophthora roreri*

Genetic resistance to the pathogen is the main strategy for controlling the disease. In line with the polygenic nature of the trait in question, the species' breeding programs seek to pyramid resistance genes, that is, to concentrate genes from different sources in a few genotypes (Phillips-Mora et al. 2018). In this sense, there is a need to constantly search for genetic materials that contain different sources of resistance to moniliasis, as well as hybrid combinations that manifest such resistance, preferably

horizontally. Thus, a range of genotypes have already been identified and are currently being used in breeding programs as clones or parents for controlled crosses, such as, for example, resistant clones UF 273 and 712, ICS 95 and 10, and PA 169, which, crossed with each other and with other clones with varied resistance indices, constituted the hybrids of the CATIE series, such as, for example, CATIE-R1, CATIE-R4 and CATIE-R6, promising materials for both resistance and almond production (Bekele and Phillips-Mora 2019; Jaimes et al. 2011; Phillips-Mora et al. 2009, 2018). Osorio-Guarín et al. (2020) report that several genetic materials with evidence of resistance to moniliasis and witches' broom have been identified, such as FCM 19, SUI 72, SCC 85, EBC 09, among others. The authors also indicate that there is a low correlation between the symptoms of the two diseases, explained by the competition for the same site of infection. The search for materials that can add resistance alleles for both pathogens of the *Moniliophthora* genus is currently the motto of the main Latin American breeding programs (Bekele and Phillips-Mora 2019; Lopes et al. 2011; Osorio-Guarín et al. 2020).

2.4.3.3 Resistance to *Phytophthora* spp.

It is noticed that the breeding for resistance to *Phytophthora* has a complicating factor, concerning the aforementioned pathogens: the great variety of etiological agents, which demands a proportional diversity of resistance sources capable of withstanding the pressure of most species, that is, efficient use of the available cocoa germplasm (Nyadanu et al. 2012). As with moniliasis, the nature of resistance to pod rot is quantitative and additive (Fister et al. 2019; Nyadanu et al. 2012), with relatively high heritability (Bekele and Phillips-Mora 2019; Nyassé et al. 2007) which allows the pyramiding of resistance genes from different origins. Several tolerant genotypes to different species have been developed in the main breeding programs in the world, such as, for example, POUND 7, SPA 9, ICS 1, IMC 47, SCA 6, IFC 5, PA 150, UF 12, as well as their hybrids, from crosses with each other and with genotypes with similar or slightly lower resistance (Bekele and Phillips-Mora 2019; Dias 2001; Fister et al. 2019; Nyadanu et al. 2012; Nyassé et al. 2007; Thevenin et al. 2012).

2.4.3.4 Resistance to CSSV

Besides, the recommended control actions, such as the elimination of diseased plants and the use of chemical pesticides, have not been efficient in reducing the viral spread. For this reason, the most sustainable and least costly option for disease control is the breeding route, with the development of genotypes that are not attractive to scale insects, as well as resistant to viral symptoms (Trebissou et al. 2020). As a result of these researches, several genotypes with evidence of genetic resistance to the virus were identified. It was observed that Amelonado-type genetic-based materials, which were widely used in the African continent, are highly susceptible to the disease, unlike genotypes originating from the Lower and Upper Amazon (Padi et al. 2013). With

these, there was even a successful attempt to obtain tolerant clones via mutagenesis, using gamma rays (Adu-Ampomah et al. 1996). The additive nature of resistance to CSSV was found, indicating the possibility of accumulation of resistance alleles via hybridization (Lockwood 1981). Furthermore, the influence of shading management to mitigate the effects of CSSV was defined, with the possibility of integrating the shade tolerance character with virus resistance in breeding programs (Andres et al. 2017, 2018). Thus, plantations in full sun tend to be more favourable to the rapid evolution of this virus (Andres et al. 2018).

To date, genotypes that are completely resistant to CSSV have not been developed, and there were low genetic gains for the trait (Padi et al. 2013). However, some materials with evidence of genetic resistance have already been identified, which are being used in breeding programs, such as Upper Amazon clones IMC 67 and 47, NA 33; the Lower Amazon clones Catongo, RB 49 and C-Sul 7, some Guyanese materials such as GU 239/H, 225/V and 290/H, some hybrid combinations between the genotypes of these origins, such as TC65 (PA 7 × IMC 35), in addition to clones from mutation-induced mvP30 and mvT85 (Bekele and Phillips-Mora 2019; Trebissou et al. 2020; Padi et al. 2013; Muller 2016; Ofori et al. 2015).

2.4.3.5 Resistance to *Ceratocystis cacaofunesta*

Ceratocystis cacaofunesta is difficult to control due to the rapid progression of the visible symptoms of the disease until the plant's death (Silva et al. 2007). Added to this fact is the way the fungus penetrates the plant, through the action of some Coleoptera of the Scolytidae family, as well as injuries caused by the management of cacao (pruning, crowning, etc.) (Santos et al. 2012b). After the plant is infected, tissues dry out and plant death, but leaves and fruits remain attached to the plant for a long time. For these reasons, development of resistant varieties via genetic improvement of the species is the most viable option to mitigate the damage caused by this pathogen (Yamada et al. 2015).

Analysing the pattern of plant mortality in the field in di-allele analyses, Gardella et al. (1982) found the influence of several genes in the manifestation of resistance to the fungus, which was later confirmed by Sanches et al. (2008) and Santos et al. (2012b). This polygeny attaches great importance to the additivity of the genes involved, attesting to the need to direct crosses between genotypes with different sources and resistance, for the pyramiding of favourable genes. Dominance also plays an important role, making it possible to exploit heterosis from crosses through cloning.

The search for sources of genetic resistance generated significant results, having identified genotypes that were later used in crosses. These include BMI 67, ICS 6, EET 400, PA 7, POUND 18, EET 272, TSH 1188, VB1151, CEPEC 2008, among others (Bekele and Phillips-Mora 2019; Santos et al. 2012b; Sanches et al. 2008; Silva et al. 2013; Yamada et al. 2015).

2.4.3.6 Resistance to *Ceratobasidium theobromae*

Cultural and chemical control methods, even when integrated, are not efficient in mitigating the harmful effects of the pathogen on crop production. However, genetic breeding, associated with such methods, constitutes an effective tool to control the fungus (Asman et al. 2021; McMahon and Purwantara 2016). As observed for resistance to the aforementioned pathogens, the additivity of the genetic resistance trait to *C. theobromae* was also observed, evidencing the importance of selecting parents with favorable genes for the production of hybrids and clones. For this character, intra-allelic dominance relationships are less important (Tan and Tan 1988). Despite being quantitative, studies show that few genes govern the trait (McMahon and Purwantara 2016).

Some clones with evidence of genetic resistance to *C. theobromae* that are currently being used in the main breeding programs are of the Trinitarian and Upper Amazon type, as well as their hybrids. As an example, we have KA2-106, KA2-101, PBC 123, BR 25, PA 191, TSH 858, SCA 9, ICS 95, UF 667, among others (Bekele and Phillips-Mora 2019; McMahon et al. 2015; McMahon and Purwantara 2016; Guest and Keane 2018).

2.4.3.7 Resistance to Major Pests

Although a range of pests coexists with the cocoa crop, reducing production, most have local economic importance, unlike the aforementioned diseases, which have a global scope and, therefore, greater appeal. For this reason, and due to the difficulty in identifying efficient sources of resistance, the search for genotypes that manifest a genetic resistance mechanism is still little explored by the classical breeding of cocoa (Bekele and Phillips-Mora 2019).

For *S. singularis*, an important pest for the African continent, Upper Amazon and Guyanese genotypes have more concrete signs of resistance, with a complex defence structure that combines antibiosis, antixenosis and tolerance itself. Some clones can be highlighted, such as ICS 1, UF 676, PA 102, EET59, in addition to the hybrids T65/7 × T57/22, T65/7 × T9/15, among others (Anikwe et al. 2009; N'Guessan et al. 2008).

Together with the pathogens *C. theobromae* and *P. palmivora*, the *Conopomorpha cramerella* borer is one of the main responsible for promoting significant losses in cocoa crops on the Asian continent (McMahon et al. 2009; Niogret et al. 2020). Estimates made by Marelli et al. (2019) attributed to the action of the pest an annual loss of 81,000 metric tons in production. For this reason, the most advanced pest-related studies aimed at achieving genetic resistance have been dedicated to *C. cramerella.* Characteristics related to insect tolerance were identified, consisting mainly of physical barriers in the genotypes, such as the thickness of almonds' sclerotic layer (Soesilo et al. 2015). There is a differential manifestation of these characteristics between genotypes, denoting genetic control and, therefore, amenable to selection

(Soesilo et al. 2015; Teh et al. 2006). In this context, some clones have greater potential for resistance to the pest, such as ARDACIAR 10, Na 33, Paba/V/81L/1, Aryadi 2, SCA12 KKM22, BR25, among others. The last three, despite having low infestation rates, do not have thick fruits and/or almonds, suggesting an alternative resistance mechanism, such as, for example, antixenosis (Soesilo et al. 2015; McMahon et al. 2009).

The resistant clones mentioned in each topic are just examples of the wide range of genotypes developed and used by classical cocoa breeding programs around the world. In the consulted literature, other genotypes available in these programs can be found.

2.4.4 Limitations of Traditional Breeding and Rationale for Molecular Breeding

As a perennial species, the cocoa tree has particularities that make it difficult to generate short-term results and, therefore, demand that selection is as accurate and efficient as possible. Among these specificities are the long reproductive period, the modification of characters over the years of cultivation, the need for large areas for genetic evaluation of a substantial number of genotypes, among others. It is noteworthy that, during the development of a cultivar, each breeding cycle takes around 10–12 years (Bekele and Phillips-Mora 2019; Dias 2001; Resende 2002).

On the other hand, the current market dynamic requires more speed and precision in generating results. Therefore, with this growing demand for research, which has intensified in recent years, the cocoa genetic breeding, using only a classical approach, will not be enough to deliver the assets required by the production chain promptly, given the particularities of the species mentioned in the previous paragraph. In this sense, biotechnological tools such as molecular markers, genomic maps, high-throughput phenotyping, among others, are strategic to increase the efficiency of the species improvement (Wickramasuriya and Dunwell 2018).

The use of molecular markers, for example, makes it possible to shorten the breeding cycle, by prematurely identifying the presence of favourable alleles in an individual. In terms of biotic stresses, the identification of marks attributed to resistance genes is essential to save months or even years of field trials, with inoculations and/or phenotyping (Osorio-Guarín et al. 2020). Once the genotypes have been sequenced, it is possible to predict the performance of their progenies or clones, enabling the targeting of crosses and thus contributing to the pyramiding of resistance genes to a particular pathogen, or tolerance to some category of abiotic stress (McElroy et al. 2018).

With the advancement of computer resources, the use of large databases is already a reality in most economically relevant crop breeding programs. For cocoa, it should be no different. Therefore, combining classical breeding with sophisticated biotechnology and statistical genetic tools can—and must—significantly increase selection

gains, and make cocoa breeding more efficient and responsive to the new challenges and opportunities that currently present themselves.

2.5 Brief on Diversity Analysis of Cocoa Germplasm

2.5.1 Phenotype-Based Diversity Analysis

Diversity analysis is supported by two actions: (i) the genetic improvement of culture aiming to overcome the environmental adversities that the culture faces throughout its domestication and use, such as biotic stress factors; and (ii) the genetic study that supports the main practices involving plant genetic resources (Henry 2005; Acquaah 2020). The estimation of a given crop that has phenotypic diversity can be understood as the differences in the expression of phenotypic characteristics between individuals or populations (Fuccillo et al. 2007). *Theobroma cacao* has a long history of domestication (Motamayor et al. 2002, 2003), which started in Mesoamerica despite the primary centre of origin and diversity being regions of the upper Amazon. Archeological findings indicate that the use of cocoa as food, in rituals and in medicine by Mesoamerican populations dates back to around 5300 years ago, indicating the pre-Columbian use of this genetic resource (Zarrillo et al. 2018).

In the last 250–300 years, cocoa cultivation spread to several tropical regions, and cultivars or varieties were used to establish productive systems in several regions of the Americas and Africa. This distribution combined with local and regional breeding has led to a wide diversity in qualitative and quantitative terms in the phenotypic descriptors of the plant and productivity mainly with the production of new cultivars and hybrids (Mustiga et al. 2018). However, the crossing between plants does not increase the genetic diversity of a crop, but evolutionary events such as the artificial selection of plants that are increasingly promising and more resistant to biotic and abiotic factors are important factors for the increase in genetic diversity (Henry 2005; Coates et al. 2018).

In cocoa, the different characteristics associated with the fruits have been the main differentiating descriptors of variety and cultivars, because it is the most desired plant organ to improve itself (Acquaah 2020; Bartley 2005). This greater emphasis on the cocoa fruit as a differentiating factor, led to estimate and divide the crop into distinct genetic groups and has also been used as a way to measure genetic diversity.

The characteristics of the fruits linked to geographic origin have also been two important classification factors in the subdivision of the crop into genetic groups, such as the varieties of the Criollo and Forastero groups. Varieties of the Criollo group are native and their origin is well known. The Forastero group, on the other hand, encompasses the varieties cultivated in various regions, and their origin is widely discussed. Both groups have a range of varieties and cultivars and even hybrids that are part of a third Trinitarian genetic group, expanding the classification system of cocoa (Bartley 2005). However, the revision of the classification of the species has been carried

out, proposing a total of ten different genetic groups, also taking the geographic distribution as a basis (Motamayor et al. 2008).

The genus *Theobroma* comprise about 22 species of origin Neotropics, with low species number diversity compared to other genera of angiosperm (Cuatrecasas 1964). However, the genus *Theobroma* is phylogenetically close to the genus *Herrania* (Schultes 1958). The intra- and intergeneric variations are based on morphological descriptors. The genus *Theobroma* has the species *T. cacao* as one of the most diversified in morphoagronomic descriptor variability, as a result of domestication and breeding. The Amazon region is an important centre of species diversity and contains genetic groups and germplasm with endemic characteristics to this region and also species at risk of extinction (Hammer and Khoshbakht 2005; González-Orozco et al. 2020).

The characteristics of fruits and seeds are the most important from an economic point of view, although wild species usually present smaller quantitative data for these descriptors, which are used in yield estimates. The species *T. grandiflorum* revealed higher fruit weight compared to *T. cacao* (Santos et al. 2012a). However, the descriptors for fruit and seeds are extremely variable in *T. cacao*, depending on the variety, clone or cultivar (Bartley 2005). The knowledge between the phenotypic and genetic relationships between wild species of the genus *Theobroma* and their wild relatives is important not only to understand the phylogeny and evolution of the group, but also to support genetic improvement with the use of biotechnological tools aimed at gene introgression and agronomically important characteristics (Henry 2005; González-Orozco et al. 2020).

The cocoa phenotypic study in locally or regionally used varieties is important, as the environment and the agronomic management associated with the crop considerably influence the descriptors related to fruits and seeds (Bartley 2005). In the coastal region of the state of Espírito Santo in Brazil, the evaluation of phenotypic and chemical descriptors associated with the fruit in seven promising cultivars for cultivation and regional production, distributed the germplasm into three statistically distinct groups, revealing that the clones CCN10 and CCN51 are the most suitable for the production of almonds aimed at the chocolate industry (Alexandre et al. 2015). The phenotypic and chemical characterization of cocoa was also carried out in regions with little availability of water (Brazilian semi-arid), in cultivars that are already widely used in the region; this demonstrate the use of the Brazilian genetic resource in several regions of the country. CCN51 and CEPEC2005 clones were the most promising for almond fermentation, while other clones (PS1319 and CEPEC2004) are also indicated for the production of other food items, such as jellies, nibs and sweets with low fat content (Reges et al. 2021). In Indonesia, several clones and hybrids resulting from the crossing between clones were evaluated for genetic diversity using qualitative and quantitative phenotypic descriptors, indicating that the generated cluster was more influenced by qualitative descriptors, generating two main groups and three subgroups in the larger group “A”, with only two clones in group “B” (AD-04 and M-01) (Lembang et al. 2019). However, it has already been observed that the length and width of the fruits have a greater impact on the multivariate distribution of cultivars (Alexandre et al. 2015). In recent years, the

characterization of cocoa has considered parameters directly related to its commercial use, with organoleptic characteristics among the products generated by different cultivars or clones. The sensory profile is important to direct the agricultural product to its final destination and consequently to the consumer market. In Peru, characteristics aimed at the production of fine-flavour cocoa have been evaluated in hundreds of plants, indicating the variability of 64 unique characteristics in combinations, expanding the possibility of using the cocoa genetic resource for the production of various items of great commercial value (Eskes et al. 2018).

In the Southern region of Bahia in Brazil, the characterization of local cocoa varieties has revealed a broad genetic base, due to the long history of introduction and breeding, fostered by local research and production institutions. In this region the local varieties Comum, Pará and Maranhão have been widely exploited in the formation of production systems, representing a good portion of the genetic resource available in the region, along with several clones, such as CCN51, TSH1188, SCA 6 and many others, whose characteristics are not only linked to the commercial value of the fruits and seeds, but also the resistance to biotic factors. In this sense, the Bahia's genetic resource has been considered a rich system in terms of genetics and phenotypic expression, revealing germplasm for the most diverse actions of improvement and selection of plants with agronomically desirable characteristics (Santos 2019) (Fig. 2.1).

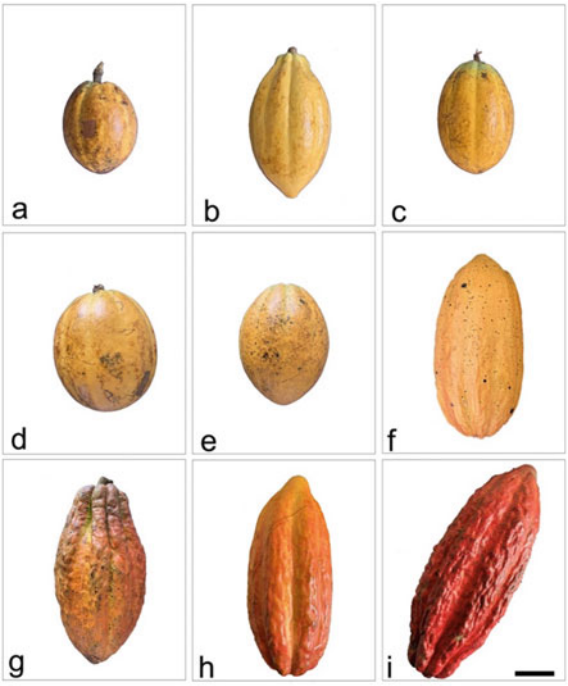

Fig. 2.1 Some shapes and sizes of ripe cocoa fruits found in plantations in southern Bahia. Local ancient varieties (**a**—Pará, **b**—Maranhão, **c**—Comum). Local mutant varieties (**d**—Redondo; **e**—Catongo). Commercial cultivars available from seedling producers (**f**—PS1319, **g**—TSH1188, **h**—SJ02, **i**—CCN51) (bar = 5 cm)

2.5.2 Genotype-Based Diversity Analysis

Until the 1980s, in general, the evaluation of genetic variation was performed based on phenotypic descriptors (Engels 1986; Nevo 1988). Additionally, the evaluation by enzymatic profile was also widely performed as a way to assess the genotypic variation using isoenzymes (Elliot and Kennedy 1988). From the 1990s onwards, molecular markers were developed, expanding the analytical perspectives of the cocoa genome. The different types of markers proposed and applied in the study of cocoa vary depending on the genetic and evolutionary principle of the analysed sequences and also the technical variation (Henry 2013).

In cocoa, molecular markers have been used to evaluate population and geographic distribution in relation to classification parameters (Motamayor et al. 2008), detecting and relating genomic and phenotypic characteristics (Barreto et al. 2018), assisting breeding programs and assessing reproductive characteristics such as self-incompatibility (Royaert et al. 2011).

Polymorphism among populations from different genetic groups was an important target for initial investigations with molecular markers-linked to the study of genetic diversity - genome size estimates, and polymorphic and specific sequence hybridizations (Figueira et al. 1992; N'Goran et al. 1994). Molecular markers based on sequence hybridization, such as restriction fragment length polymorphisms (RFLPs)—used for the evaluation of polymorphisms together with random markers based on polymerase chain reaction (PCR) random amplification of polymorphic DNAs (RAPDs), enabling the evaluation of diversity genetics, but with little certainty in the relationship between the germplasm and the proposed genetic groups (N'Goran et al. 1994). RFLP probes with low-copy sequences enabled the detection of variability between different genetic groups, with indications of introgression by hybridization between these genetic groups of different origin in germplasm. However, there is a certain genomic purity with low heterozygosity and other genetic and diversity specifications of each group already described, such as Criollo, Forastero and Trinitário (Lerceteau et al. 1997).

In the last three decades, around 50 different diversity studies based on molecular markers have resulted in highlighting the genetic diversity of cocoa in the different regions where it is cultivated (Fig. 2.2). The genome sequencing, and the use of nuclear and chloroplastic microsatellite sequences (SSR—simple sequence repeat) have been an effective strategy for the analysis of genetic polymorphism in cocoa (Lanaud et al. 2004, 2009; Motamayor et al. 2008; Santos et al. 2015; Lachenaud and Zhang 2008). SSR markers in cocoa have been efficient in the methodological standardization for the evaluation of the genetic diversity (Saunders et al. 2004).

The diversity of cocoa has brought about great discussions about the number and classification of genetic groups, a factor that often reflects on the phenotypic variability. The differentiation of Amazonian cocoa populations revealed, based on geographic distribution and molecular markers, the formation of ten distinct genetic groups: Amelonado, Contamana, Criollo, Curaray, Guyana, Iquitos,

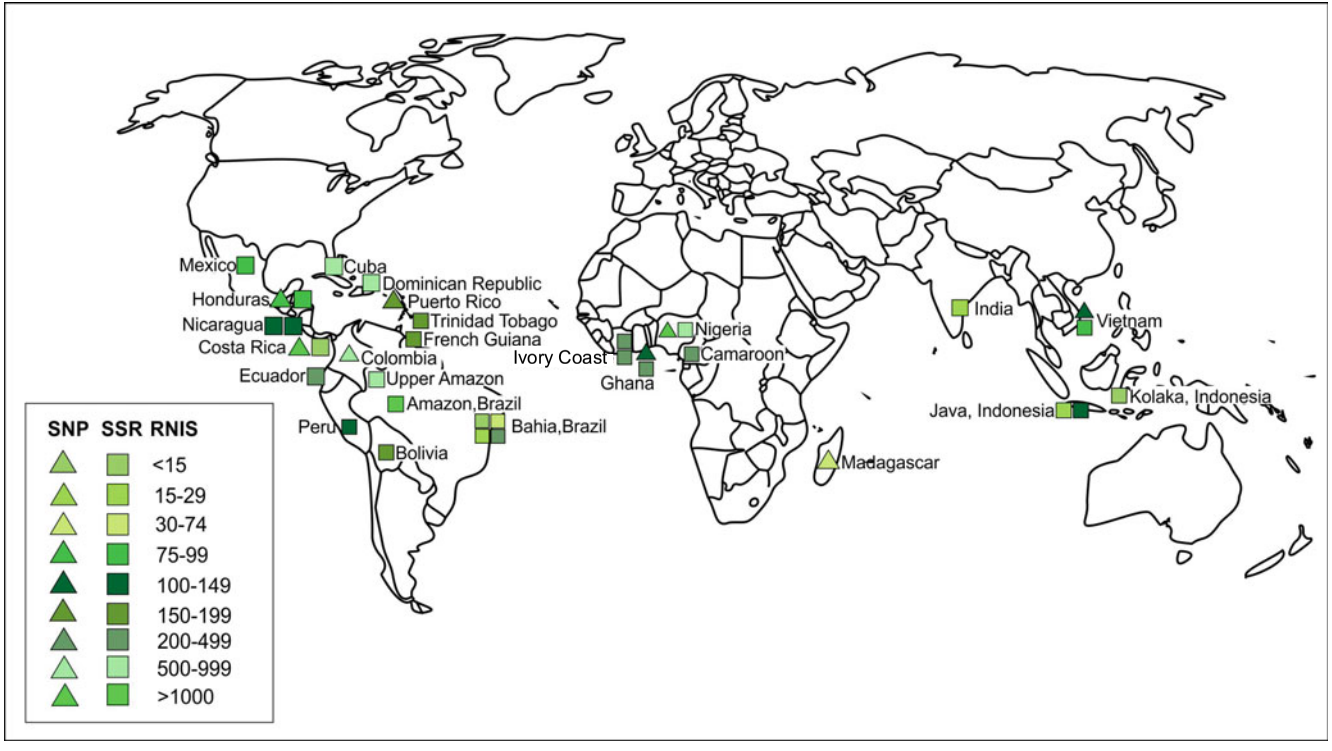

Fig. 2.2 Studies on the genetic diversity of cacao around the world, based on SNP and SSR markers. RNIS, range of the number of individuals per each study in different locations

Marañon, Nacional, Nanay and Purús expanding the possibilities of use of the cocoa genetic resource in breeding and conservation programs (Motamayor et al. 2008).

In Brazil, both the diversity and the genetic structure of plants cultivated in Bahia, revealed that the genetic resource of this location has particular characteristics, presenting a specific identity in terms of variability, differing from plant populations from other regions. In this case, cocoa produced in the Bahia State showed great genetic divergence between plants from other cultivation locations (Santos et al. 2015). Another approach to study the diversity of local cocoa in Bahia was to assess the diversity of materials that showed resistance to witches' broom after its occurrence in this cocoa-producing region. Then, RAPD markers were used, enabling access to genetic diversity and subsequent selection of potentially resistant plants (Faleiro et al. 2004a; Leal et al. 2008). Later, SSR markers were used for the same purpose of finding genetic novelties as alternative sources of plants resistant to witches' broom (Yamada et al. 2009; Faleiro et al. 2004b; Lima et al. 2013).

A population of Ecuador was evaluated for polymorphic SSR loci; it has been shown the trend of genetic erosion caused by the breeding and introgression of germplasm, reducing the number and diversity of native allele or wild plants (Loor et al. 2009). In Peru, high diversity has been revealed, depending on the region, and also due to the spatial structure of the collected accessions. However, the introduction of germplasm from other regions has altered the region's native germplasm, forming other populations in natural and semi-natural farming systems (Zhang et al. 2009b). In West Africa, genetic analysis of cocoa trees from different collection sites revealed great allelic richness with only 12 loci evaluated, indicating a greater diversity in germplasm from the upper Amazon compared to the genetic resource introduced from other regions (Aikpokpodion et al. 2009).

Expressed sequence tag (EST)-SSR markers have been used for functional genomic analysis, serving as a basis for the study of genome-phenotype relationship (Argout et al. 2011). Additionally, SSR markers have been related to genetic regions responsible for disease resistance in cocoa, one of the biggest problems associated with the crop in terms of reduced productivity, crop dropout by producers and plant death (Pugh et al. 2004; Brown et al. 2007; Lanaud et al. 2009; Akaza et al. 2016).

In the 2010 decade, the genome sequencing of cocoa was reported, expanding the perspectives in the evolutionary and genetic study of the species (Argout et al. 2011, 2017). The information generated by the cocoa genome sequencing allowed the identification and development of single nucleotide polymorphisms (SNPs). SNPs markers were also obtained to analyse the origin of traditional varieties cultivated in Madagascar, indicating that the genetic resource of this region has descent from the Criollo, Amelonado and Trinitário groups (Li et al. 2021).

2.6 Brief Account of Molecular Mapping of Resistance Genes and QTLs

2.6.1 Genetic Maps of Cocoa, Marker Evolution and Segregating Populations

The genetic mapping of cocoa based on molecular markers allowed the identification of genes and QTLs for resistance to different diseases that affect this crop, as well as other characteristics, mainly those that affect production, consequently with an economic effect. The genetic mapping of the first cocoa populations segregating for witches' broom and black pod resistance predominantly used the RFLP, AFLP and RAPD markers. They were carried out with F1 population and backcross population involving Catongo and Pound 12, identifying QTLs for resistance to black pod disease (Crouzillat et al. 2000), as well as with the population derived from the artificial self-pollination of the hybrid TSH 516, identifying QTLs with large effect for resistance to witches' broom (Queiroz et al. 2003). This map containing QTLs for resistance to witches' broom was expanded from 193 to 342 markers and each individual was multiplied by grafting for repetition of phenotypic evaluations, enabling greater accuracy in the phenotyping step (Faleiro et al. 2006). In this case, the focus was to identify genomic regions related to the resistance of cocoa to witches' broom, in order to assist in genetic improvement steps. Subsequently, different populations and marker types were used to identify QTLs and candidate genes for resistance to these two diseases (Tables 2.2 and 2.3).

The first genetic map from segregating populations for frosty pod rot resistance involved the F1 population derived of Pound 7 × UF 273, using SSR markers or single strand conformation polymorphism (SSCP) of resistance gene homolog (RGH) and WRKY markers (Brown et al. 2007). The two genetic maps of cocoa population segregating for Ceratocystis wilt resistance used SSR and SNP markers. These maps involved F2 population derived from SCA6 × ICS1 (Santos et al. 2012b), and population F1 population derived from TSH1188 × CCN51 (Fernandes et al. 2018).

Different strategies were used for the development and mapping of SNP markers in cocoa. For example, almost a hundred conserved ortholog set II (COS-II) have been identified and mapped in cocoa, with SNPs being identified for 83 genes, among which 19 cosegregated with QTLs (Kuhn et al. 2012). The strategies used in this study consisted of (1) evaluating polymorphisms in the mapping population, (2) identifying SNPs in a diversity panel through DNA amplification from 15 different cocoa accessions based on conserved genes, (3) analysing the transcriptome from leaf RNA from these 15 cocoa accessions. The 15 different genotypes used by Kuhn et al. (2012) to identify SNPs in conserved sequences were previously allocated to 10 different genetic groups based on SSR markers (Motamayor et al. 2008).

Another SNP discovery and mapping initiative was carried out based on a germplasm bank of 249 accessions and two linkage genetic mapping populations.

Table 2.2 Generation, size and genealogy of different populations of cocoa genetic mapping

Generation	Genealogy[a]	Size	QTL or sequence[b]	References
F1 BC1	Catongo × Pound 12 → F1 (Catongo × Pound 12) × Catongo → BC1	55 131	BP	Crouzillat et al. (2000)
F2	SCA6 × ICS1 → TSH516 ⊗ → F2	82	WB	Queiroz et al. (2003)
F2	SCA6 × ICS1 → TSH516 ⊗ → F2	82	WB	Faleiro et al. (2006)
F2	SCA6 × ICS1 → TSH516 ⊗ → F2	146	COS-II	Kuhn et al. (2012)
F1	UPA402 × UF676 → F1	125	Candidate gene	Fouet et al. (2011)
F2	SCA6 × ICS1 → TSH516 ⊗ → F2	143	CW	Santos et al. (2012b)
F1	TSH1188 × CCN51 → F1	459	WB	Royaert et al. (2016)
F1	TSH1188 × CCN51 → F1	265	BP	Barreto et al. (2018)
F1	TSH1188 × CCN51 → F1	266	CW	Fernandes et al. (2018)
F1	Pound 7 × UF 273 → F1	256	BP, FP	Brown et al. (2007)
F1	Pound 7 × UF 273 → F1	179	BP, FP	Gutiérrez et al. (2021)
F1	EET 95 × Silecia 1 → F1	733	WB, FP	Livingstone et al. (2017)
F1	SCA 12 × unknown genotype → F1	251	WB, FP	Livingstone et al. (2017)
Trihybrid	(SCA6 × H) × C1 → F1 (P7 × ICS100) × C1 → F1 (P7 × ICS95) × C1 → F1	179 173 183	BP	Akaza et al. (2016)
F1	DR1 × Catongo → F1 S52 × Catongo → F1 IMC78 × Catongo → F1	96 94 125	BP	Clement et al. (2003)
Trihybrid	(Na34 × IMC60) × IFC2 → F1	59	BP	Flament et al. (2001)
F1	T60/887 × IFC5 → F1	56	BP	Flament et al. (2001)
Trihybrid	(SCA6 × H) × C1 → F1	151	BP	Risterucci et al. (2003)

[a]For more details on the genetics of each genotype denoted here by the acronyms, see the original articles. ⊗ , selfcross. BC, backcross. F1, first generation. F2, second generation

[b]QTL for different diseases (BP, black pod; WB, witches' broom; FP, frost pod; CW, Ceratocystis wilt) and sequence or candidate genes (COS-II, conserved ortholog set of resistance genes; candidate genes; EST sequence)

Table 2.3 QTLs associated with resistance to pathogens in cocoa

Disease	Species	Population	Marker	Phenotyping	Software linkage/QTL	Statistical methodology	QTL number	LOD	LG	%VE (R^2)	References
BP	*Phytophthora palmivora*	Catongo × Pound 12(F1), F1 × Catongo (BC)	RFLP, RAPD, AFLP	Artificial inoculation in fruits	MAPMAKER/Q-gene	One-way ANOVA, simple interval mapping	6	>2	1, 2, 4, 5, 9(× 2)	7.4 to 47.9	Crouzillat et al. (2000)
BP	*Phytophthora palmivora*	T60/887 × IFC2, T60/887 × IFC5	AFLP, RFLP	PRR, Artificial inoculation in fruits and leaf-disc	JOINMAP/MAPQTL	Nonparametric marker by marker	7	2.2 to 4.2	2(×2), 3(×2), 6(×2), 10	9 to 17	Flament et al. (2001)
BP	*Phytophthora palmivora (×2), P. megakarya(× 2), P. capcisi(× 2)*	SCA6_H × IFC1	AFLP, SSR	Artificial leaf-disc inoculation	JOINMAP/MAPQTL	Composite interval mapping	13	2.93 to 4.06	1(×3), 3, 5(× 5), 6(×3), 7	7.5 to 12.4	Risterucci et al. (2003)
BP	*Phytophthora palmivora*	IMC78 × Catongo, DR1 × Catongo, S52 × Catongo	AFLP, RFLP, SSR	PRR	MAPMAKER-EXP/QTLCartographer	Composite interval mapping	2	2.5, 7.4	4(×2)	10.1, 22.6	Clement et al. (2003)

(continued)

Table 2.3 (continued)

Disease	Species	Population	Marker	Phenotyping	Software linkage/QTL	Statistical methodology	QTL number	LOD	LG	%VE (R^2)	References
BP	*Phytophthora palmivora*	Pound7 × UF271 (F1)	SSR	Artificial inoculation in fruits	JOINMAP/MAPQTL	Restricted multiple QTL mapping	3	6.1 to 14.6	4, 8, 10	7.3 to 23	Brown et al. (2007)
BP	*Phytophthora palmivora*	(SCA6 × H) × C1, (P7 × ICS100) × C1 and (P7 × ICS95) × C1	SSR	PRR, artificial inoculation in leaf-disc	JOINMAP/MAPQTL	Single marker locus analysis, simple interval mapping	11	2 to 4.35	1(×2), 2, 3, 4(×2), 6(×3), 8, 10	13.2 to 27.6	Akaza et al. (2016)
BP	*Phytophthora palmivora, P. citrophthora, P. capsici*	TSH1188 × CCN51 (F1)	SRR	Artificial inoculation in leaf-disc	OneMap/OneMap	Composite interval mapping	6	3.224 to 3.873	1, 2, 3, 4, 6(× 2)	1.776 to 3.299	Barreto et al. (2018)
BP	*Phytophthora palmivora*	Pound7 × UF273 (F1)	SNP	Artificial inoculation in fruits	JOINMAP/GenStat	Marker regression, simple interval mapping, composite interval mapping, multiple QTL mapping	4	3.82 to 10.88	2, 4, 8, 10	8.23 to 23.07	Gutiérrez et al. (2021)

(continued)

Table 2.3 (continued)

Disease	Species	Population	Marker	Phenotyping	Software linkage/QTL	Statistical methodology	QTL number	LOD	LG	%VE (R^2)	References
WB	*Moniliophthora perniciosa*	SCA6 × ICS1 (F2)	RAPD, AFLP	Average number of vegetative brooms per canopy area	MAPMAKER-EXP/Q-gene	Single factor, composite interval mapping	1	>4	11	34.8	Queiroz et al. (2003)
WB	*Moniliophthora perniciosa*	SCA6 × ICS1 (F2)	SSR	Number of brooms produced by infection	JOINMAP/MAPQTL	Simple interval mapping, multiple QTL mapping, restricted multiple QTL mapping	3	3.38 to 10.55	1, 9(× 2)	6.7 to 51.1	Brown et al. (2005)
WB	*Moniliophthora perniciosa*	TSH1188 × CCN51 (F1)	SSR	Number of brooms produced by infection (natural and artificial)	-/SAS	Simple regression	3	–	3, 9(× 2)	74.6 to 82.7	Santos et al. (2007)
WB	*Moniliophthora perniciosa*	TSH1188 × CCN51	SNP	Vegetative brooms, cushion brooms	JOINMAP/GenStat	Restricted maximum likelihood	7	2.318 to 11.24	3, 4, 6(×2), 7, 9(× 2)	0.6 to 13.5	Royaert et al. (2016)

(continued)

Table 2.3 (continued)

Disease	Species	Population	Marker	Phenotyping	Software linkage/QTL	Statistical methodology	QTL number	LOD	LG	%VE (R^2)	References
FP	*Moniliophthora roreri*	Pound7 × UF273 (F1)	SSR	Artificial internal and external inoculation of pods	JOINMAP/MAPQTL	Restricted multiple QTL mapping	5	3.9 to 7.2	2(×2), 7, 8(×2)	4.5 to 9.8	Brown et al. (2007)
FP	*Moniliophthora roreri*	Pound7 × UF273 (F1)	SNP	Artificial inoculation of pods	JOINMAP/GenStat	Marker regression, simple interval mapping, composite interval mapping, multiple QTL mapping	11	3.87 to 9.19	2, 4(×2), 7(×2), 8(×2), 9(×2), 10(×2)	4.96 to 11.19	Gutiérrez et al. (2021)
CW	*Ceratocystis cacaofunesta*	SCA6 × ICS1 (F2)	SSR, EST-SRR	Stem inoculation	JOINMAP/MAPQTL	Simple interval mapping, composite interval mapping, multiple QTL mapping	2	2.57 to 3.27	3, 9	7.7 to 9.6	Santos et al. (2012b)
CW	*Ceratocystis cacaofunesta*	TSH1188 × CCN51 (F1)	SNP	Stem inoculation	JOINMAP/MAPQTL	Interval mapping, multiple QTL mapping	2	4 to 48.85	4, 6	3.8 to 62.6	Fernandes et al. (2018)

BP black pod; *CW* Ceratocystis wilt; *FP* frosty pod; *LG* linkage group; *WB* witches' broom

Among these different cocoa genotypes, 15 were used to identify EST-SNPs, among which the four genotypes used by Allegre et al. (2012) for the linkage genetic mapping of the discovered EST-SNPs. Two-population maps and a saturated consensus genetic map were developed for cocoa; they included SNPs and multiple markers, some of them located in coding sequences of candidate genes (Fouet et al. 2011; Royaert et al. 2011; Fernandes et al. 2018).

The three genetic maps developed from the TSH1188 × CCN51 population represent an effort to identify genomic regions involved in the resistance response to witches' broom, Ceratocystis wilt and black pod, which are the three main diseases that affect cocoa in several regions where it is cultivated. SNP and SSR markers were co-localized with genome sequences to correlate gene function to observed traits such as plant resistance (Royaert et al. 2016; Fernandes et al. 2018; Barreto et al. 2018).

DNA ships containing species-representative SNPs were used for the genetic mapping of cocoa. A chip with 6 k cocoa SNPs was used to map and identify QTLs for resistance to witches' broom and Ceratocystis wilt (Royaert et al. 2016; Fernandes et al. 2018). Additionally, the 15 k strategy consisted of identifying thousands of SNPs, many of them within genes annotated in the cocoa genome, and transforming them into an Illumina Infinity II array. These SNPs were identified within a diversity panel of 11 different cocoa accessions. The mapping of these SNPs was performed in two F1 populations of full-sib: one was derived from the cross between EET 95 × Silecia 1; the other was from SCA12 × unknown genotype (Livingstone et al. 2017). In this last work, genomic information and plant phenotyping were intensively used for resistance to frosty pod rot, the incidence of witches' broom disease and other characteristics.

Cocoa genetic mapping populations include different generations, genealogies and sizes (Table 2.2). These segregating populations range from 82 to 733 individuals, with a predominance of medium-sized populations (150–250 individuals), which are suitable for genetic mapping. They were generated from controlled crosses between wild varieties (SCA 6), mutants (Catongo), simple hybrids (self-pollination of TSH516) and complex hybrids (TSH1188, CCN51, etc.). Thus, these populations are full-sib progenies.

2.6.2 QTL Regions Disease Resistance in Cocoa

Efforts to identify QTL regions associated with biotic stress in cocoa have focused on the four main pathogens that affect culture: black pod, witches' broom, frosty pod and Ceratocystis wilt.

2.6.2.1 Black Pod Resistance

Since the early 2000s, most of the efforts have focused on the elucidation of the QTL regions involved in resistance to the different species of the genus *Phytophthora* that cause black pod, due to the devastating consequences of the pathogens in culture; these efforts have used various molecular markers such as AFLP, RAPD, RFLP, EST-SSR, SSR and more recently SNP. A QTL region located in the linkage group (LG) 4, initially identified by Crouzillat et al. (2000), represents one of the most promising regions in the interaction with the different species of *Phytophthora*. This region has been validated in different genotypes and in different countries (Table 2.3). Other regions located in the LGs 1, 2, 4, 8 and 10 have been recurrent in various studies. The polygenic character of black pod resistance was initially proposed by Spence & Bartley (Spence and Bartley 1966), which has been validated by the identification of various QTLs, from QTLs of major effect to minor effect by different studies (Table 2.3). It is important to highlight that the values of % phenotypic variation explained (%VE) presented by the various authors have been influenced by the different phenotyping methodologies and statistical methodologies implemented in each study, ranging from point-to-point analysis to multipoint analyses potentiated with machine learning algorithms like HMM.

2.6.2.2 Witches' Broom Resistance

The main phenotyping methodology used in the studies that have addressed to the witches' broom resistance is the quantification of infected vegetative brooms in the field and the studies have focused on two crosses, SCA6 × ICS1 and TSH1188 × CCN51. The markers used were restricted to RAPD, AFLP, SSR and more recently SNP. A first study identified a QTL in LG 11 with a %VE = 34.8 (Queiroz et al. 2003). Although it has been refined in terms of phenotyping and of use of a greater number of markers (REF), this results most likely was due to the use of a linkage map wih insufficient coverage for the date. Subsequently, Brown et al. (2005) identified a QTL with a major effect with %VE = 51.1 in LG 9, among others with a minor effect (Table 2.3). In 2007, Santos et al. also identified QTLs in the same LG and LG 3 with %VE between 74.6 to 82.7, and this study used a simple regression statistical approach without the use of a linkage map. More recently, Royaert et al. (2016) using SNP also identified two QTLs in LG 9, but with substantially lower %VE and also identified several QTLs of minor effect varying between 0.6 and 13.5 of %VE, due to a different genetic background used to generate the segregating population (Table 2.3).

2.6.2.3 Frosty Pod Resistance

Efforts to elucidate QTL regions associated with resistance to frosty pod caused by the *M. roreri* fungus, are less compared to black pod and witches' broom. This may

be due to the less cosmopolitan distribution of the disease, present only in northern South America and have reports in Central America to southern Mexico. To date, only two studies have been carried out on detection of QTLs associated with resistance to frosty pod (FP) (Table 2.3), both using the same genotypes, Pound7 × UF273 (F1). Initially Brown et al. (2008) identified five QTLs in LG 2, 7, and 8 with a %VE varying between 4.5 and 9.8 using SSR markers. Later, Gutiérrez et al. (2021) using SNP markers identified 11 QTLs confirming those previously identified by Brown et al. (2008), and identified 6 more regions with a %VE varying between 4.96 and 11.19.

2.6.2.4 Ceratocystis Wilt Resistance

The vascular disease Ceratocystis wilt, caused by the fungus *Ceratocystis cacaofunesta*, is of utmost importance since it causes the death of infected plants and it has also been proven that the vast majority of witches' broom resistant genotypes show susceptibility to Ceratocystis wilt. The first study carried out by Santos et al. (2012b) used a linkage maps previously developed using SSR and EST-SSR markers, and identified two QTLs in LG 3 and 9, with a %VE of 7.7 to 9.6 on the cross SCA6 × ICS1 (F2). The first and only effort in the identification of QTLs associated with Ceratocystis wilt resistance based on a linkage map constructed with the use of SNP markers using the TSH1188 × CCN51 (F1) cross, was carried out by Fernandes et al. (2018), identifying two QTLs in LG 4 and 6, with a %VE between 3.8 and 62.6, the last in LG 6, being the first QTL of major effect reported as associated with resistance to Ceratocystis wilt in cocoa.

2.6.2.5 Brief Summary of QTL Regions Associated with Resistance to Biotic Stresses in Cocoa

The first decade of QTL mapping of disease resistance in cacao was examined in detail using meta-analysis (Lanaud et al. 2009). Nowadays, the number of QTLs and types of diseases have been increased (Table 2.3). A summary of QTL regions is shown from the main QTLs associated with resistance to biotic stress of the four diseases addressed in QTLs studies (Fig. 2.3) (Akaza et al. 2016; Gutiérrez et al. 2021; Brown et al. 2005, 2007; Royaert et al. 2016; Santos et al. 2012a, b; Fernandes et al. 2018). In this summary, only the QTLs identified by interval mapping methodologies in studies with $n \geq 150$ individuals, and showing the highest %VE (R^2) by disease are shown.

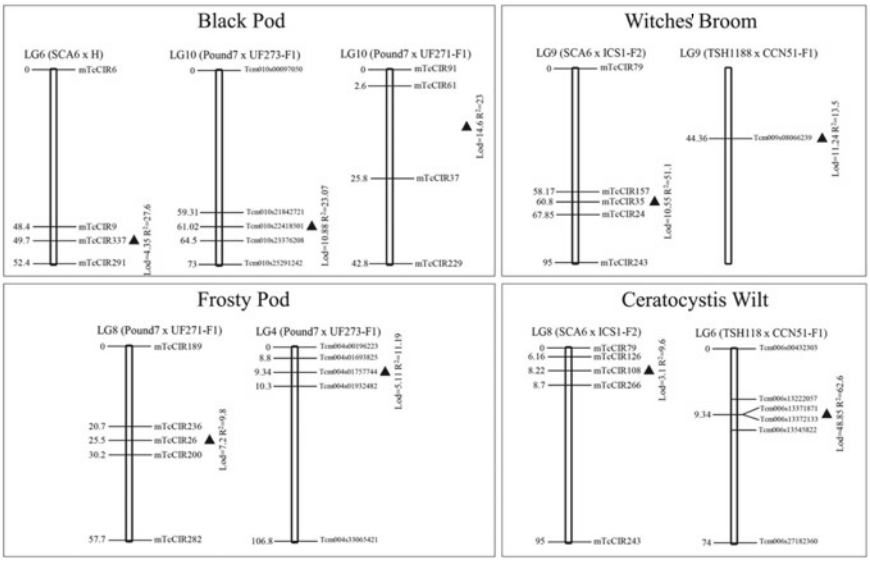

Fig. 2.3 Main linkage groups, crosses and QTLs associated with resistance to biotic stresses in cocoa. In each linkage group the relative position (cM) is shown on the left and the markers on the right. ▲: QTL Region, LG: Linkage group, Lod: LOD peak value of the logarithm of the likelihood ratio observed, R^2: proportion of the phenotypic variation explained by the QTL. This map was built with data from previous studies (Akaza et al. 2016; Gutiérrez et al. 2021; Brown et al. 2005, 2007; Royaert et al. 2016; Santos et al. 2012b; Fernandes et al. 2018)

2.7 Cocoa Germplasm Characterization

The characterization of a plant germplasm gives crucial information to understand the genetic bases of the given species, allowing the comprehension of the identity of the individuals that make it up, the genealogical and geographic relationships, the genetic variability present in populations, and the geographic origin of the accessions present in the germplasm bank, which together serve as a basis for both the conservation of genetic resources and for the improvement of certain characteristics of agronomic interest (Lindo et al. 2018; Zhang et al. 2009a). In the case of *T. cacao*, morphological and/or molecular characterizations have already been performed in germplasm banks and/or farms present in several countries, such as Costa Rica (Zhang et al. 2009b), Peru (Zhang et al. 2009a), Colombia (Osorio-Guarín et al. 2017), Jamaica (Lindo et al. 2018), China (Wang et al. 2020), Trinidad (Bekele et al. 2020), Brazil (Santos et al. 2015), Cameroon (Efombagn et al. 2009), Dominican Republic (Boza et al. 2013), Honduras and Nicaragua (Ji et al. 2013), among others.

In Costa Rica, 688 accessions of *T. cacao* from the International Cocoa Collection of CATIE (IC3) maintained in the country were used, which were genotyped using 15 SSR molecular markers. The level redundancy analysis revealed that, in a random sample, 113 accessions were representing 90% of the allelic diversity

present in the germplasm collection. Two hundred and thirty-one alleles distributed in 548 genotypes were identified. These 548 genotypes, based on geographic origin, were classified into 12 groups. The groups from Brazil and Ecuador accounted for about 56% of the total number of exclusive alleles, noting that the collection was composed of only 18.4% from these two groups. Through genetic structure analyses, it was observed that within-country/within-region was responsible for 84.6% of the total molecular variation, while between-the-countries/between-regions represented 15.4%. The results found in this work mainly indicated that the IC3 contains a high level of redundancy (Zhang et al. 2009b).

In Peru, the evaluation was made using a total of 612 cocoa accessions from the Pound collection, the first to be established in South America and which has important sources of genetic variation that can be exploited in breeding programs, especially for resistance to diseases. These accessions were genotyped using 15 microsatellite markers. As main results, 180 cases of mislabelling and a total of 116 duplicates were detected, and 316 accessions were then selected for the diversity analyses. The half-sib and full-sib families were rebuilt for the five access groups in the collection, with 78 half-sib and 48 full-sib families being rebuilt. Through analysis of probability simulations, eight parents were identified as probably responsible for 117 pairs of mother–child relationships present in the collection. The principal coordinate analysis (PCA) together with the Bayesian method of clustering indicated a marked structure of genetic diversity stratified by the fluvial systems of the Peruvian Amazon (Zhang et al. 2009a).

In Colombia, a set of 565 accessions was used for evaluation, of which 450 came from the Colombian Corporation for Agricultural Research (Corpoica) germplasm bank and 115 from breeding collections. For genotyping, these 565 accessions plus 252 accessions from reference populations were characterized using 87 SNP markers. For phenotyping, 141 accessions were analysed using 18 morphological characters and 94 accessions were analysed with four biochemical characters, both from the UPOV descriptor list. PCA analysis of the morphological characters showed that 60.6% of the total variation was represented in seven descriptors, while for the biochemical characters 100% of the variation was explained in the four characters evaluated, noting that for both sets of characters the genotypes analysed were grouped into four clusters. Genotyping analysis showed that this collection from Corpoica has a high genetic diversity (Osorio-Guarín et al. 2017).

In Jamaica, 160 accessions from the germplasm banks of the Orange River Agricultural Research Station in the parish of St. Mary and Montpelier Agricultural Research Station in St. James, plus 150 reference accessions from farms have been characterized, both sets of accessions being genotyped using 94 SNP markers. The results indicated that most of the genotypes present in Jamaica collection are hybrids originating mainly from the genetic groups Parinari, Iquitos Mixed Calabacillo, Scavina, Amelonado and Criollo, with the greatest contributions being from the Parinari and Amelonado groups. Through the construction of the Neighbour-joining dendrogram, the formation of two large clusters was observed, in cluster 1 the accessions from the Nacional and Scavina groups were inserted and in cluster 2 the accessions from the Amelonado, Criollo, Trinitario, Nanay, Parinari and Iquitos Mixed

groups Calabacillo. Through the analysis of molecular variance, it was also found that the highest level of differentiation occurred among individuals within the population (97%) (Lindo et al. 2018).

In China, where a collection of 170 accessions of *T. cacao* is maintained in Yunnan, a sample of 88 accessions was selected for characterization and was genotyped with 91 SNP markers. It is also worth mentioning the use of over 140 reference accessions. Through PCA, it was observed that most accessions belonging to the Yunnan collection were dispersed in the Amelonado group and other reference groups. Through the UPGMA dendrogram, the formation of large clusters was observed, cluster 1 with all reference populations, except Amelonado, and cluster 2 with all accessions of the Yunnan collection together with the reference population Amelonado. These results showed that the introductions of cocoa accessions in the Yunnan collection consisted mainly of hybrids derived from the Amelonado group, thus demonstrating a low genetic diversity and suggesting the need for new introductions of genotypes from the other groups, to obtain a greater representation of the genetic diversity of this species (Wang et al. 2020).

In Cameroon, 300 accessions from farms of the South and West regions, and 77 accessions from the Institute of Agricultural Research for Development (IRAD) germplasm bank were selected for characterization, and both sets were evaluated using eight quantitative and nine qualitative descriptors. Morphological variations were observed between accessions from farms and those from the germplasm bank. When evaluating the quantitative characters related to the pods, it was possible to find differences between the sets of accessions from farms in the South and West regions. By PCA, it was found that 86% of the total phenotypic variance was contained in four principal components. Comparing morphological and molecular results, it was observed that the differences found between farms in the South and West region were mainly influenced by non-genetic factors. While accessions from the germplasm bank and those from farms were considered genetically distant, which suggested a low introduction of genotypes from breeding programs in farm crops (Efombagn et al. 2009).

In the Dominican Republic, 803 accessions from the germplasm bank and 55 accessions cultivated in local farms were selected for characterization by genotyping with 14 SSR markers. It was detected in the set of accessions from the germplasm bank the presence of 15 synonymous groups containing 48 accessions, while for the set coming from local farms, the occurrence of 13 synonymous groups containing 30 accessions was identified. Analysing the genetic diversity, the presence of 117 and 113 alleles for the germplasm and farm accessions, respectively was observed. When evaluating the population structure, a predominance of accessions in the Amelonado group was detected, which represented 43.9% for germplasm accessions and 72.1% for the farm accessions (Boza et al. 2013).

In Trinidad, 1900 accessions from the International Cocoa Genbank Trinidad (ICGT) germplasm, of which 260 belong to the Trinitario group were characterized using 25 descriptors from the International Board for Plant Genetic Resources. Upon observation of the Pod Index, which is used as an indicator of potential yield, a variation between 13.9 and 92.8 was found, with values below 21.0 being considered

favourable. It was also observed that 23% of the accessions considered superior belonged to the Trinitario group. Through PCA, it was identified that 69% of the total variation was contained in nine principal components. Observing the characters 'number of seeds' and 'cotyledon mass' for the 260 accessions of the Trinitario group, six were considered as elite for potential yield (Bekele et al. 2020).

In Honduras and Nicaragua, 84 accessions collected from traditional cocoa farms and 31 reference clones from the IC3 were selected and genotyped with 70 SNP markers. It was found, in the group of 84 accessions collected on farms, the presence of six synonymous groups, in which 16 accessions were included, representing 11.9% of the varieties included in this group. Through multivariate analysis, the formation of five clusters was identified, which represented the genetic groups Antigo Criollo, Amelonado, Trinitario from Nicaragua, Trinitario from Honduras and Forastero from Superior Amazon. It was found that the classification of accessions by farmers had high coincidence with the groups formed through genotyping by SNPs, highlighting that from the 28 accessions named as Criollo, 22 were in fact inserted in the Antigo Criollo group; two out of three accessions named as the Trinitario of Nicaragua were grouped into the Trinitario reference group; and 12 of the 13 accessions named as Trinitario de Honduras actually belonged to the Trinitario de Honduras group (Ji et al. 2013).

In Brazil, 279 accessions of *T. cacao* were selected for characterization, from which 179 were from the IFBAIANO and CEPLAC institutions, located in Bahia State in the cities of Canavieiras, Camacan, Gandu and Uruçuca, and another 103 genotypes from the CEPEC/CEPLAC institution, in the cities of Ilhéus and Igrapiúna in Bahia State. Genotyping was carried out with 30 SSR markers. Genetic structure analysis allowed the formation of two groups, one named as Bahia cocoa and the other as Non Bahia cocoa. In the Bahia cocoa group, two subgroups were observed, one where the genotypes cultivated on farms were inserted and the other composed of SIAL/SIC clones. Based on genetic analysis, the 30 SSR markers identified 209 alleles. Analyzing the groups separately, the genetic diversity present in the Bahia cocoa group was considered low, when compared to that presented in the Non Bahia cocoa group (Santos et al. 2015).

By analyzing the work already carried out, it is observed that the characterization of germplasm banks provides results that mainly allow to understand the genetic diversity of the species present in a given collection, influencing both genetic conservation measures and breeding programs. These data can support the selection of superior accessions that can be indicated for planting, as well as genotypes that can serve as parents in crosses in order to insert or increase traits of agronomic importance (fruit quality, yield, flavor and disease resistance) or to identify genes that can be used in genetic engineering work, mainly for disease resistance.

2.8 Map-Based Cloning of Resistance Genes

2.8.1 BAC Libraries

The construction of genomic DNA libraries based on bacterial artificial chromosomes (BAC) can be used for genomic mapping, positional cloning, complex genome sequencing and gene annotation (Holmes et al. 2015; Luo and Wing 2003). For *T. cacao*, a BAC library was built and validated in 2004, aiming to develop molecular resources to enable the study of the structure and evolution of the genome of this species (Clement et al. 2004). For the construction of the *T. cacao* BAC library, DNA extraction was performed from collected leaves of the SCA6 genotype from the germplasm collection of the Center National de Recherche Agronomique (CNRA) of Ivory Coast. As a result, this library has 36,864 clones with an average size of 120 kb each, with about 80% of the clones having a size greater than 100 kb and 13% greater than 150 kb, thus representing a coverage around 10 times greater in relation to the size of the haploid genome. Additionally, from the establishment of this library, the authors characterized and refined regions associated with disease responses. Relationships between physical and genetic distances were performed, and a screening was carried out in the BAC library using nine resistance gene analog (RGA) and defence gene analog (DGA) probes previously obtained (Lanaud et al. 2004). Two resistance clusters were identified: a region of 3.4 cM located on chromosome 4 and another of 1.1 cM on chromosome 7, which, respectively, were used to construct two contigs, one of 1004 kb and another of 404 kb, thus providing the first clues about the relationships between physical and genetic distances in the genome of *T. cacao* (Clement et al. 2004).

2.8.2 Cytogenetic Studies

Cytogenetic studies provided a great contribution to the understanding of the genomic structure and the evolutionary biology of cultivated plants, mainly because they provide important integrative tools allowing genetic and genomic analyses of plant chromosomes and genomes (Figueroa and Bass 2010). Such studies may facilitate the understanding of behaviour, as well as genetic regulations and mechanisms that, at the gene level, control chromosomal dynamics, and may also contribute to the identification and transfer of resistance genes from exotic to native species (Younis et al. 2015).

For *T. cacao*, some cytogenetic studies have already been carried out, proving that this species has a symmetrical karyotype consisting of $2n = 20$ chromosomes (da Silva et al. 2017; Dantas and Guerra 2010; Figueiredo et al. 2013), with size haploid around 0.45 pg (da Silva et al. 2017) and length of each chromosome ranging from 1.19 to 2.00 μm, with arm ratios from 1.12 to 1.32 μm (Dantas and Guerra 2010). Specifically, the varieties Cacau Comum (native to the Lower Amazon Region)

and Cacau Rui (originated in the Cocoa Region of Bahia State in Brazil) contain exclusively metacentric chromosomes, while the varieties Cacau Jaca (originated in the Cocoa Region of Bahia State in, Brazil) and Cacau Papala (originated in Peru) presented one pair of submetacentric chromosomes each, pairs 5 and 7, respectively (Figueiredo et al. 2013). Analysing commercial cocoa species, by conventional staining with DAPI, an interphasic reticulate type of nucleus was observed, with 19–20 chromocenters with regular shape and size. In prophase, through conventional staining and C-Banding, it was observed that the chromosomes presented high condensation in the proximal region and decondensation in the terminal region (Dantas and Guerra 2010). Using the CMA+/DAPI− double staining technique, it was observed the presence of this band in the terminal region of the long arm in only one pair of chromosomes, the CMA+ band being often heteromorphic in size and distended in one or both homologues (da Silva et al. 2017; Dantas and Guerra 2010). Using the fluorescence in situ hybridization (FISH) technique, analyses of rDNA sites were performed, observing the occurrence of a 5S rDNA site in the proximal region of one of the three largest chromosomes and a 45S rDNA site located together with the CMA+ band. By using C-banding followed by staining with DAPI or Giemsa, it was observed in the interphase nuclei the presence of 20 well-defined chromocenters, all presenting centromeric or proximal heterochromatic bands of similar sizes (Dantas and Guerra 2010).

From cytogenetic studies, 20 EST-SSR type markers were also developed, and informative and satisfactory results on allelic variations between wild and cultivated species were obtained. Thus, 60% of these markers were polymorphic and 40% monomorphic in *T. cacao* (da Silva et al. 2017), strengthening the assertion that these studies can bring essential contributions both to structural and functional genomics as well as to the evolutionary biology of species (Figueroa and Bass 2010).

2.9 Genomics-Aided Breeding for Resistance Traits

2.9.1 Large Scale Transcriptomic Resources

A transcriptome is a snapshot of the gene expression profile of a cell and/or tissue at a given time and/or physiological situation, which can be analysed through the establishment of transcript libraries and subsequent application of specific tools (Ward et al. 2012). In *T. cacao*, functional genomics analyses have mainly sought to understand the roles of certain genes, from redundant and shared functions to unique functions (Gesteira et al. 2007), which has led to the construction of EST libraries from cocoa submitted to various factors (Argout et al. 2008; Gesteira et al. 2007; Jones et al. 2002) or disease resistance inducers (Verica et al. 2004) and more recently to high throughput analysis such as RNAseq (Teixeira et al. 2014; Fister et al. 2015).

The first effort to establish large scale transcriptomics data from *T. cacao* was made in 2002 using complementary DNA (cDNAs) from seeds and leaves of five

cocoa varieties (Amelonado, P7B, R10, Spec54, UF221 and Sic 5) (Jones et al. 2002). Two libraries from Amelonado seeds and leaves, respectively, were obtained, and an unigene of 1380 sequences was identified and annotated. In the seeds, transcripts associated mainly with storage and defence were found (e.g. vicilin, 21-kDa seed protein), while leaves mainly contained sequences related to photosynthetic process (Jones et al. 2002). In 2004, suppression subtraction hybridization (SSH) libraries resulting from *T. cacao* leaves (Forastero and Comum varieties) treated with the defence response inducers benzothiadiazole (BTH), Nep1 and methyl jasmonate/ethylene were built. A dataset of 2114 ESTs, which resulted in the assembly of 1256 unigenes was obtained. From them, 865 had corresponding genes already annotated in the database, and 330 corresponded to genes induced during the defence response (Verica et al. 2004). In 2007, an EST library was developed specifically for the *T. cacao-Moniliophthora perniciosa* interaction, using Catongo (susceptible) and TSH1188 (resistant) cocoa varieties as plant material. In this study, the authors established two EST libraries, one originating from the interaction of *M. perniciosa* with resistant genotype (RT) and another with susceptible genotype (SP), generating a dataset of 1719 unigenes and 1207 for RT and SP, respectively. From them 1371 and 859 unigenes were specific for RT and SP, respectively, and 16 functional classes were established (Gesteira et al. 2007). In 2008, 56 libraries were built, of which 25 corresponded to cocoa tissues subjected to different biotic stresses, highlighting pods infected by the *P. palmivora*, *P. megakarya*, *M. perniciosa* and *M. roreri*; leaves infected by *P. palmivora* and *P. megakarya*; stems inoculated with *M. perniciosa* and *C. fimbriata*; and stems attacked by *Sahlbergella singularis*. From the 56 libraries, 146,650 ESTs were generated, corresponding to 48,594 unigenes, 12,692 contigs and 35,902 singletons. The authors identified 1001 genes associated with stress responses, mainly resistance proteins or proteins involved in plant defence mechanisms, such as chitinases, 1-3 beta glucanases and pathogenesis-related proteins (PRs) (Argout et al. 2008). In 2014, a transcriptomic analysis of the *T. cacao-M. perniciosa* interaction was obtained by RNA-seq (Teixeira et al. 2014). The cocoa cultivar Comum was inoculated with the BP10 *M. perniciosa* isolate and material from infected seedlings was harvested 30 days after inoculation. The authors identified 562 million and 436 million of reads for infected seedlings and healthy seedlings (control), respectively. The alignment of reads against the Cacao Genome Database (Motamayor et al. 2013) (see also Sect. 2.9.2) and the differential expression analysis between infected and control samples allowed the identification of 1967 differentially expressed genes, 1269 upregulated and 698 downregulated (Teixeira et al. 2014). In 2015, through microarray analysis, genes expressed in the SCA6 and ICS1 cocoa genotypes inoculated with *P. tropicalis* and treated with salicylic acid (SA) were identified. As main results, the authors found that the treatment with SA reduced the size of the lesions and limited the growth of the pathogen in the two cocoa genotypes, especially in SCA6. Among other results, it was observed that the ICS1 genotype treated with SA showed a greater number of upregulated PR genes compared to the SCA6 genotype, also treated with SA (Fister et al. 2015) (Table 2.4).

Table 2.4 Main large-scale omics data from cocoa related to biotic stress

Resource type	Technical data	Cocoa genotype	Pathogen or inducer	Data amount	Database or accession ID	Publication year	References
Transcriptomics	ESTs, microarrays	Amelonado, P7B, R10, Spec54, UF221, Sic5	Non-inoculated leaves and seeds	1380 unigenes	np	2002	Jones et al. (2002)
	SSH, microarrays	Forastero, Comum	Benzothiadiazole, Nep1, methyl jasmonate/ethylene	1256 unigenes	CF972636–CF974749	2004	Verica et al. (2004)
	ESTs	TSH1188, Catongo	*M. perniciosa*	2926 unigenes	ES439783-ES440989 ES440990-ES442709	2007	Gesteira et al. (2007)
	ESTs, SSH	ICS1, SCA6	*P. palmivora*, *P. megakarya*, *M. perniciosa, M. roreri*, *C. fimbriata*, *Sahlbergellasingularis*	48,594 unigenes	https://esttik.cirad.fr/	2008	Argout et al. (2008)
	RNA sequencing (RNA-seq)	Comum	*M. perniciosa*	562,491,362 paired reads (infected)	SRA066232	2014	Teixeira et al. (2014)
	Microarrays	SCA6, ICS1	*P. tropicalis*, salicylic acid	1974 differentially expressed genes (salicylic acid)	GPL18260	2015	Fister et al. (2015)
Genomics	Genome-wide shotgun strategy, Roche/454, Illumina and Sanger sequencing	Belizean Criollo	na	76% genome, 82% genes anchored (version 1)	https://cocoa-genome-hub.southgreen.fr/	2011	Argout et al. (2011)

(continued)

Table 2.4 (continued)

Resource type	Technical data	Cocoa genotype	Pathogen or inducer	Data amount	Database or accession ID	Publication year	References
	Illumina large insert size mate paired libraries, 52× long reads, Genotyping-by-sequencing	Belizean Criollo	na	75.5% genome, 99% genes anchored (version 2)	https://cocoa-genome-hub.southgreen.fr/	2017	Argout et al. (2017)
	Genome-wide shotgun strategy, Roche/454 sequencing, fosmid and BAC libraries	Matina 1,6	na	29,408 loci, 98.9% of the scaffolds anchored	https://www.cacaogenomedb.org/ PRJNA51633, ALXC00000000 Phytozome 12	2013	Motamayor et al. (2013), Goodstein et al. (2012)
Proteomics	2D gel, mass spectrometry	TSH1188, Catongo	*M. perniciosa*	554 proteins	np	2020	dos Santos et al. (2020)

np non-provided; *na* non-applicable

The overall data obtained through transcriptomic approaches allowed the identification of genes specifically involved in resistant *vs* susceptible cocoa genotypes (from different resistant sources, e.g. SCA6, TSH), in different metabolic pathways responsible for important agronomic traits (e.g. resistance, defence) or in specific plant organs (e.g. pod, seeds, leaves) (Table 2.4). Such data provided solid bases for an increase of the global knowledge of the plant-pathogen interaction, but also as source of information useful for molecular strategies (e.g. marker development) associated to cocoa breeding programs.

2.9.2 Genome Sequencing

The genome is the set of all the genetic information of an organism or a cell, which can be determined in an increasingly broad, fast and efficient way thanks to the development of new sequencing platforms, known as next generation sequencing (NGS), enabling the whole genome sequencing (WGS) (Giani et al. 2020). The genomic sequencing of species of agronomic interest has allowed breeders more consistent approaches, especially when they are dealing with multigene inheritance traits, which consequently suffer a greater environmental influence and are more difficult to be evaluated and selected based only on classical breeding techniques (Ray and Satya 2014).

For the *T. cacao* species, the first genome sequencing was accomplished in 2011, from the Belizean Criollo (B97-61/B2) genotype, which is appreciated because its almonds result in the production of fine chocolate. For this sequencing, the authors used the genome-wide shotgun strategy incorporating Roche/454, Illumina and Sanger technologies, generating a total of 26 Gb of data. The raw data were used for assembly using the Newbler tool, generating 25,912 contigs and 4792 scaffolds, which represented a total assembly length of 326.9 Mb, covering approximately 76% the genome size of this species. Through the use of 1259 molecular markers, 94% located in the assembly, it was possible to anchor 67% of the sequenced genome of *T. cacao* along the 10 linkage groups. By linking genetic distance to physical distance, relationships of 1 cM per 444 kb and 1 cM per 146 kb were observed for centromeric and chromosomal regions, respectively. It was identified, through a search for homology and functional annotation, that the genome of *T. cacao* has 28,798 genes that encode proteins, each gene having an average size of 3346 bp and 5.03 exons (Argout et al. 2011). In 2017, a second version of the Belizean Criollo genome was established (Argout et al. 2017) using some data from the first version (Argout et al. 2011) and adding new results generated by sequencing Illumina HiSeq 2000 and 52× long reads sequencing. In order to fill gaps existing in the first annotation system, these new data were integrated, and a consensus annotation was established, containing a total of 29,071 contigs encoding for proteins. In summary, this new version of the *T. cacao* genome added data compared to the previous version, improving assembly, correcting some inconsistencies, reducing the

number of scaffolds and unanchored regions, and updating structural and functional information (Argout et al. 2017).

Independently, another *T. cacao* genome was sequenced: the cultivar Matina 1-6 was used, standing out for belonging to the most cultivated cocoa group (Motamayor et al. 2013). For sequencing, the Sanger and Roche 454 pyro-sequencing methods were used, with subsequent assembly performed via Arachne. Altogether, around 32.4 Gb of sequences were generated, as well as 20,103 contigs, which formed 711 scaffolds, representing a total length of 346.0 Mbp. It was observed that 98.9% of these data were mapped on the 10 chromosomes, where 29,408 loci were evidenced. Aiming to establish a comparison between this genome of Matina 1-6 with that of Criollo, the authors performed a synteny analysis, resulting in the identification of 271 orthologous regions between the two genomes. It was found that most genes responsible for LLR-RLK protein synthesis, flavonoid and lipid biosynthesis, and terpenoid synthesis are inserted outside orthologous regions, which led the authors to suggest that genes located outside these regions or in orthologous regions between non-orthologous chromosomes may be responsible for the existing differences related to disease resistance and fruit quality between the two genotypes (Motamayor et al. 2013).

From the sequencing, annotation and availability in databases of the Belizean Criollo and Matina 1-6 genomes, genomic and post-genomic approaches were facilitated, mainly serving as a basis for projects involving comparative and/or evolutionary studies in *T. cacao* and enhancing genetic analysis to identify genes involved in important characteristics of the cocoa tree, which has been directly contributing to the breeding programs of this species.

2.9.3 Proteomics Resources

Currently, only few studies related to proteomics of cacao-pathogen interaction have been developed. However, in 2020, comparative proteomics analysis of the *T. cacao-M. perniciosa* interaction was performed. TSH1188 (resistant) and Catongo (susceptible) cocoa genotypes were inoculated with *M. perniciosa*, samples were harvested through the time course disease, and their proteins analysed by mass spectrometry. The authors obtained a total of 554 protein spots with a greater amount of proteins observed in TSH1188, mainly in inoculated samples. In the Catongo genotype, uninoculated samples (control) presented more proteins than the inoculated ones, indicating a possible gene repression after inoculation with the pathogen. The identified proteins were distributed into eight functional groups, and it was noted that the TSH1188 genotype presented a greater amount of defence and stress proteins than the Catongo one, both inoculated with *M. perniciosa.* Subsequently, these proteins were identified: PRs, chitinases, proteins related to oxidative stress regulation and trypsin inhibitors. For proteins differentially expressed in the two genotypes, interactomics analysis revealed the occurrence of complex networks of protein–protein interactions specific of each genotype (dos Santos et al. 2020).

2.9.4 Bases for Marker Assisted Selection

With the support of structural and functional *T. cacao* omics data, the cocoa breeding has contributed to identify potential sources of resistance to diseases (Wickramasuriya and Dunwell 2018). By accessing the information present in the genome and how it is structured, it is possible to locate genomic regions that are possibly responsible for triggering resistance responses in the plant. These regions have already been described in *T. cacao* for interaction with several pathogens, especially *T. cacao-M. perniciosa* (McElroy et al. 2018; Osorio-Guarín et al. 2020; Schnell et al. 2007; Motilal et al. 2016), *T. cacao-M. roreri* (Romero Navarro et al. 2017; Gutiérrez et al. 2021; McElroy et al. 2018; Osorio-Guarín et al. 2020; Schnell et al. 2007), *T. cacao-Phytophthora* spp. (Gutiérrez et al. 2021; Motilal et al. 2016; Schnell et al. 2007) and *T. cacao-C. cacaofunesta* (Fernandes et al. 2018; Santos et al. 2012b). These and other works cited throughout this chapter have provided important support to know how certain characters are genetically determined and at the same time allowing their selection assisted by markers, thus contributing to the development of new cocoa cultivars with certain characteristics of agronomic interest in a shorter time span (see also Sect. 2.6 above).

2.10 Brief on Genetic Engineering for Resistance Traits and Recent Concepts and Strategies Developed

2.10.1 Review on Achievements of Transgenics

In addition to support by works involving molecular marker assisted selection, genomic data provide a crucial basis for selection of potential genes to be used in genetic engineering applications, mainly through recombinant DNA technology. For the application of this technique, *Agrobacterium tumefaciens* has become the most popular tool for plant transformation, by delivering genes of interest in specific plants including cocoa (Hwang et al. 2017; Maximova et al. 2003). In *T. cacao*, some works involving the use of this technique have already been carried out, highlighting functional studies of *T. cacao* genes such as *chitinase 1* (Maximova et al. 2006), *NPR1* (Shi et al. 2010), *TcNPR3* (Shi et al. 2013), *PI3P* (Helliwell et al. 2016) and *TcBI-1* (Scotton et al. 2017).

The *chitinase 1* gene was overexpressed in the PSU-SCA6 cocoa genotype though *Agrobacterium* mediated-plant transformation (Maximova et al. 2006). Cocoa transgenic plants, compared do the wild-type (non-transgenic plants), showed an increase of the chitinase activity in their leaves as well as a reduction of necrotic lesions when the plants were challenged with the pathogen *Colletotrichum gloeosporioides.* Globally, the results showed that the plants overexpressing the *chitinase 1* gene presented a higher resistance response against the pathogen *C. gloeosporioides* than non-transformed plants (Maximova et al. 2006).

The *TcNPR1* cDNA was isolated from of the SCA6 cocoa genotype and used to complement, by genetic transformation, Arabidopsis plants presenting mutation for *NPR1* (npr1-2 mutant) as well as a high susceptibility to the pathogen *Pseudomonas syringae* pv. tomato (Shi et al. 2010). Complementation showed that the *TcNPR1* gene is a functional ortholog of *A. thaliana* NPR1, and that the overexpression of *TcNPR1* in npr1-2 mutant plants conferred an upregulation to the *PR1* gene after treatment with salicylic acid, and increased the resistance to the pathogen *P. syringae* pv. tomato (Shi et al. 2010). Similar methodology was used to elucidate the function of the cocoa *NPR3* gene, considered as a possible repressor of defence responses mediated by *NPR1*. The *TcNPR3* cDNA was isolated from SCA6 cocoa genotype, and then transferred to *A. thaliana* plants mutated for *npr3*, the Arabidopsis endogenous version of this gene. Functional complementation confirmed that *TcNPR3* is a negative regulator of defence responses in floral tissues, and a functional ortholog of the *A. thaliana NPR3* gene (Shi et al. 2013). *Theobroma cacao* leaves were also submitted to transient transformation with *TcNPR3*: the knockout of this gene led to an increase of the resistance responses against the fungus *Phytophthora capsici*, resulting in smaller lesion sizes and reduction of the pathogen replication (Shi et al. 2013).

Functional study of the *PI3P* gene was made by transient and stable SCA6 cocoa genotype transformation (Helliwell et al. 2016). Cocoa leaves transiently overexpressing the *PI3P* gene showed an increase of resistance against the pathogen *P. palmivora*, while the transient overexpression of *PI3P* directed to the apoplast showed an increase to the resistance against the pathogen *P. tropicalis*. Stable transformation of cocoa plants with *PI3P* showed an increased resistance against the pathogen *C. theobromicola*, while addressing the *PI3P* sequence to the apoplast, stable transformation showed an increased resistance against the *P. palmivora* and *P tropicalis* pathogens (Helliwell et al. 2016).

The negative programmed cell death regulator Bax-1isolated from CAB214 cocoa genotype was used to transformed the Micro-Tom tomato cultivar (Scotton et al. 2017), a plant model for cocoa-*M. perniciosa* interaction (Marelli et al. 2009). As main results, the Micro-Tom TcBI-1-transformed plants inoculated with the necrotrophic pathogens *Sclerotinia sclerotiorum*, *Sclerotium rolfsii*, *Botrytis cinerea*, and *M. perniciosa* showed a significant reduction of the severity of the corresponding symptoms. Additionally, the authors suggested that the overexpression of *TcBI-1* affects the penetration of germinating *M. perniciosa* spores into susceptible tissues of *T. cacao*, and may be able to restore part of the non-host resistance against the S-biotype of *M. perniciosa* (Scotton et al. 2017).

These studies showed that genetic transformation of cocoa (or of related plant model) allowed to functionally confirm the role of some cocoa candidate genes in plant resistance to different pathogens. These results also opened up potential opportunities for the development of further work both in conventional genetic breeding and in biotechnological applications, aiming to increase the resistance responses of *T. cacao* (Maximova et al. 2006; Shi et al. 2010, 2013; Helliwell et al. 2016).

2.10.2 Genome Editing

A great advance with the *T. cacao* genome sequencing (see Sect. 2.9.2) is the possibility to efficiently and precisely edit sequences of interest related to important traits such as pathogen response or resistance. Among the gene editing techniques, the one that has been gaining recently more prominence is the clustered regularly interspaced short palindromic repeats-associated protein 9 (CRISPR/Cas9), due to its simplicity, high efficiency, easy use in the production of transgenics (Mali et al. 2013). Additionally, CRISPR/Cas9 has a high versatility and can be used for knockout, insertion, deletion and replacement of genes, as well as in the activation of gene expression, providing approaches that directly or indirectly help researchers in the processes of plant domestication and improvement of traits of agronomic interest (Ahmad et al. 2020). In agriculture, CRISPR/Cas9 stands out above all for the development of works aimed to the resistance of plants to pathogens, with three factors supporting this assertion: (i) there is a high availability of information related to specific pathosystems allowing the identification of potential genes involved in triggering defence responses to be edited; (ii) disease resistance is often achieved after editing a gene; and (iii) targeted mutagenesis is easily applied to disease resistance, as the inactivation of susceptibility genes can lead to the manifestation of a resistant phenotype (Borrelli et al. 2018).

In *T. cacao*, the first work involving the use of the CRISPR/Cas9 technique took place in 2018 (Fister et al. 2018). In this work, the use of CRISPR/Cas9 transient-mediated transformation to introduce the CRISPR/Cas9 components into cocoa leaves and cotyledon cells would limit the expression of *TcNPR3*, a defence response suppressor gene (see also Sect. 2.10.1). It was hypothesized that the knockout of this gene would result in an increase in resistance responses in tissue treated with CRISPR/Cas9. As main results, the CRISPR/Cas9 edition system allowed the deletion of 27% of *TcNPR3* copies in treated tissues (leaves). Moreover, 72 h after inoculation with the pathogen *P. tropicalis*, the tissues treated with the *TcNPR3*-CRISPR/Cas9 construct exhibited a reduction of the lesion size when compared to the control. The transformed leaves inoculated with the pathogen also showed an increase of the expression level of the *PR-2, PR-3, PR-4, PR-5* and *TcNPR1* genes in relation to control. Then, stable transgenics for *TcNPR3* in mutant embryos were obtained, being noted that the expression of single guide RNA (sgRNAs) and Cas9 in developing embryos can increase the proportion of mutated *TcNPR3* in the tissue under study, which would lead to a possible homozygosis for this mutated embryo or plant. The overall results confirmed the role of *TcNPR3* as a repressor of resistance responses in *T. cacao*, and additionally the application of this technique in a pioneering way for this species represents a great advance in functional genomics and allows greater precision in genetic engineering applications in the genetic breeding of cacao (Fister et al. 2018).

2.10.3 Nanotechnology

The advent of nanotechnology has brought new perspectives to face the constant challenges encountered when working with phytopathology, and may in the future be used to assist in the management, diagnosis and genetic transformation of plant diseases (Elmer and White 2018). In *T. cacao*, some studies aiming the development of nanoparticles have already been developed. In 2019, work was carried out to synthesize silver nanoparticles (AgNPs) from three different parts of the cocoa fruit: seed, husk and pulp. As part of this study, the antimicrobial potential of nanoparticles against *Bacillus subtilis* and *Escherichia coli* was tested; the authors found that the nanoparticles affected bacterial growth, probably causing protein leakage and cell death. Additionally, it was observed that AgNPs from the pulp showed better antimicrobial activity than the AgNPs from seed and husk. The authors argued that the these results would open new perspectives, especially for antimicrobial applications in environmental sciences, health and related fields (Thatikayala et al. 2019). In 2021, another work was developed, this one using cocoa husks for the synthesis of zinc oxide nanoparticles (ZnONPs), verifying its antimicrobial potential against foodborne pathogens. As main results, the authors found that ZnONPs showed antimicrobial activity against *E. coli* and *Staphylococcus aureus* when used at concentrations of 6.25 and 12.5 μg /mL, respectively. Additionally, it was noted that ZnONPs had a more potent antimicrobial activity than chloramphenicol. The authors suggested that the results open possibilities to explore underutilized plant materials for the synthesis of ZnONPs in an ecological and economical way, as well as for future applications of ZnONPs in food, cosmetic, textile and therapeutic medicine packaging (Sarillana et al. 2021).

2.11 Brief Account on Tole of Bioinformatics as a Tool

2.11.1 Gene and Genome Databases

Currently available *T. cacao* genomes are from 2 distinct genotypes, the Criollo genotype which can be found in Cocoa Genome Hub (Argout et al. 2017) and National Center for Biotechnology Information (NCBI), Matina genotype has its genome available in Cocoa Genome Database (www.cacaogenomedb.org/) and Phytozome 12 (Goodstein et al. 2012). The genomes of the two genotypes (Criollo and Matina) have an excellent assembly quality and are good reference genomes (see also Sect. 2.9.2 and Table 2.4). Unfortunately, there is no unified database in which to integrate genotype genomes with tools to compare them.

2.11.2 Comparative Genome Databases

Some comparative genomic analyses of *T. cacao* can be done in databases such as NCBI and Phytozome for example. The quality of the *T. cacao* genome is also important for comparative studies of genetic families in other species or plant families such as Brassicaceae (Hofberger et al. 2015), which ends up indirectly providing new information on comparative genomics of *T. cacao.*

2.11.3 Gene Expression Databases

Although transcriptomics studies from *T. cacao* exist, most of the data and results are not available or accessible in public databases. The two main *Theobroma cacao* ESTs data that can be accessed are the ESTtik database (https://esttik.cirad.fr/) (Argout et al. 2008) and the CocoaESTdb bank (http://cocoaestdb.cpcribioinformatics.in/) (Naganeeswaran et al. 2015). A unique ESTs database from ESTtik and dbEST was obtained and is available at http://cocoaestdb.cpcribioinformatics.in/ (Naganeeswaran et al. 2015). Transcriptomics studies of the *T. cacao* and *M. perniciosa* interaction (Teixeira et al. 2014) and more recently gene expression studies related to evolutionary adaptations (Hämälä et al. 2019) were also obtained (see also Sect. 2.9.1 and Table 2.4).

2.11.4 Protein Databases

Although there are no specific proteomic databases for *T. cacao*, there are proteomic studies in somatic embryogenesis (Noah et al. 2013; Alexandra Pila Quinga et al. 2018; Niemenak et al. 2015), seed proteomic studies (Kumari et al. 2018; Scollo et al. 2018), abiotic stress (Monteiro Reis et al. 2020), interaction with pathogens (dos Santos et al. 2020) and fruit husk proteomics (Awang et al. 2010). These studies, among others, reinforce the importance of the proteomics study of *T. cacao* as well as provide information for a better omic knowledge of the species.

2.11.5 Integration of Different Data

In silico tools and *T. cacao* genome analysis provided a great source of knowledge to integrate and correlate data as well as to carry outstudies of genome-wide gene families such as transcription factors (Silva Monteiro de Almeida et al. 2017) or pathogenesis related proteins (Fister et al. 2016). Some studies also focused on understanding plant resistance mechanisms through characterization of a genome-wide

pattern recognition receptors and *R* genes, which make up the system innate immune *T. cacao* (Santana Silva and Micheli 2020; Li et al. 2016). As for the evolutionary aspect, studies of cacao ancestral gene recombination contributed to a better understanding of the relationship between cocoa species (Utro et al. 2012). In general, the literature on *T. cacao* is rich in information, data and omics studies related to evolutionary, agronomic and responses to biotic and abiotic stresses. The biggest challenge would be the integration of omics data from the literature in one or more *T. cacao* databases. Some efforts were made to integrate and modelled data through interactomics or systems biology allowing the drawing of general and integrative schemes related to specific physiological situation (e.g. interaction cocoa-*M. perniciosa*) (da Hora Junior et al. 2012; dos Santos et al. 2020).

References

Abrokwah F, Dzahini-Obiatey H, Galyuon I, Osae-Awuku F, Muller E (2016) Geographical distribution of cacao swollen shoot virus molecular variability in Ghana. Plant Dis 100(10):2011–2017

Acebo-Guerrero Y, Hernández-Rodríguez A, Heydrich-Pérez M, El Jaziri M, Hernández-Lauzardo AN (2011) Management of black pod rot in cacao (*Theobroma cacao* L.): a review. Fruits 67(1):41–48

Acquaah G (2020) Principles of plant genetics and breeding, 3rd ed. Wiley-Blackwell, Chichester, West Sussex, United Kingdom, 848 pp

Adeigbe OO, Adewale BD, Muyiwa AA (2021) Variability and character association of bean biochemical traits of cocoa (*Theobroma cacao*) genotypes in four Nigerian field banks: providing a platform for nutrition-based selection. J Agric Food Res 3:100115

Adomako B (2007) Causes and extent of yield losses in cocoa progenies. Trop Sci 47(1):22–25

Adu-Ampomah Y, Owusu GK, Sackey S, Padi B, Abdul-Karimu A (1996) Use of gamma rays to induce mutants resistant to cocoa swollen shoot disease in *Theobroma cacao* L. Plant Breed 115(1):74–76

Ahmad S, Wei X, Sheng Z, Hu P, Tang S (2020) CRISPR/Cas9 for development of disease resistance in plants: recent progress, limitations and future prospects. Brief Funct Genom 19(1):26–39

Aikpokpodion PO, Motamayor JC, Adetimirin VO, Adu-Ampomah Y, Ingelbrecht I et al (2009) Genetic diversity assessment of sub-samples of cacao, *Theobroma cacao* L. collections in West Africa using simple sequence repeats marker. Tree Genet Genomes 5(4):699–711

Aime MC, Phillips-Mora W (2005) The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. Mycologia 97(5):1012–1022

Akaza MJ, Kouassi AB, Akaffou DS, Fouet O, N'Guetta AS-P et al (2016) Mapping QTLs for Black pod (*Phytophthora palmivora*) resistance in three hybrid progenies of cocoa (*Theobroma cacao* L.) using SSR markers. Int J Sci Res Publ 6(1)

Alexandra Pila Quinga L, Heringer AS, Pacheco de Freitas Fraga H, do Nascimento Vieira L, Silveira V et al (2018) Insights into the conversion potential of *Theobroma cacao* L. somatic embryos using quantitative proteomic analysis. Sci Hortic 229:65–76

Alexandre RS, Chagas K, Marques HIP, Costa PR, Cardoso Filho J (2015) Caracterização de frutos de clones de cacaueiros na região litorânea de São Mateus, ES. Rev Bras Eng Agric Ambient 19(8):785–790

Allegre M, Argout X, Boccara M, Fouet O, Roguet Y et al (2012) Discovery and mapping of a new expressed sequence tag-single nucleotide polymorphism and simple sequence repeat panel for large-scale genetic studies and breeding of *Theobroma cacao* L. DNA Res 19(1):23–35

Ameyaw GA, Dzahini-Obiatey HK, Domfeh O (2014) Perspectives on cocoa swollen shoot virus disease (CSSVD) management in Ghana. Crop Prot 65:64–70

Amon-Armah F, Domfeh O, Baah F, Dzahini-Obiatey HK (2021) Effectiveness of the eradication campaign of cocoa swollen shoot virus disease in Ghana: the extension and implementation problem. Int J Pest Manag:1–14

Andres C, Gattinger A, Dzahini-Obiatey HK, Blaser WJ, Offei SK et al (2017) Combatting cocoa swollen shoot virus disease: what do we know? Crop Prot 98:76–84

Andres C, Blaser WJ, Dzahini-Obiatey HK, Ameyaw GA, Domfeh OK et al (2018) Agroforestry systems can mitigate the severity of cocoa swollen shoot virus disease. Agric Ecosyst Environ 252:83–92

Anikwe JC, Omoloye AA, Aikpokpodion PO, Okelana FA, Eskes AB (2009) Evaluation of resistance in selected cocoa genotypes to the brown cocoa mirid, *Sahlbergella singularis* Haglund in Nigeria. Crop Prot 28(4):350–355

Araújo IS, de Souza Filho GA, Pereira MG, Faleiro FG, de Queiroz VT et al (2009) Mapping of quantitative trait loci for butter content and hardness in cocoa beans (*Theobroma cacao* L.). Plant Mol Biol Rep 27(2):177–183

Araujo L, Silva Bispo WM, Cacique IS, Cruz MFA, Rodrigues FA (2014) Histopathological aspects of mango resistance to the infection process of *Ceratocystis fimbriata*. Plant Pathol 63(6):1282–1295

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T et al (2008) Towards the understanding of the cocoa transcriptome: production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. BMC Genomics 9:512–512

Argout X, Salse J, Aury JM, Guiltinan M, Droc G et al (2011) The genome of *Theobroma cacao*. Nat Genet 43(2):101–108

Argout X, Martin G, Droc G, Fouet O, Labadie K et al (2017) The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. BMC Genomics 18(1):730

Asman A, Purung MHb, Lambert S, Amiruddin A, Rosmana A (2021) Effect of rootstock and scion on resistance of cocoa clones to vascular streak dieback caused by *Ceratobasidium theobromae*. Ann Agric Sci 66(1):25–30

Awang A, Karim R, Mitsui T (2010) Proteomic analysis of *Theobroma cacao* pod husk. J Appl Glycosci 264

Bailey BA, Meinhardt LW (2016) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, New York, USA, XIII, 633 pp

Bailey BA, Evans HC, Phillips-Mora W, Ali SS, Meinhardt LW (2018) *Moniliophthora roreri*, causal agent of cacao frosty pod rot. Mol Plant Pathol 19(7):1580–1594

Barreto MA, Rosa JRBF, Holanda ISA, Cardoso-Silva CB, Vildoso CIA et al (2018) QTL mapping and identification of corresponding genomic regions for black pod disease resistance to three *Phytophthora* species in *Theobroma cacao* L. Euphytica 214(10):188

Bartley BG (1994) A review of cacao improvement. Fundamental, methods and results. Paper presented at the international workshop on cocoa breeding strategies, Kuala Lumpur

Bartley BGD (2005) The genetic diversity of cacao and its utilization. Online Editor: CABI books, 356 pp

Bekele F, Phillips-Mora W (2019) Cacao (*Theobroma cacao* L.) breeding. In: Al-Khayri J, Jain S, Johnson D (eds) Advances in plant breeding strategies: industrial and food crops. Springer, Cham, pp 409–487

Bekele FL, Bekele I, Butler DR, Bidaisee GG (2006) Patterns of morphological variation in a sample of cacao (*Theobroma cacao* L.) germplasm from the International Cocoa Genebank, Trinidad. Genet Resour Crop Evol 53(5):933–948

Bekele FL, Bidaisee GG, Singh H, Saravanakumar D (2020) Morphological characterisation and evaluation of cacao (*Theobroma cacao* L.) in Trinidad to facilitate utilisation of Trinitario cacao globally. Genet Resour Crop Evol 67(3):621–643

Benjamin CS, Luz EDMN, Santos WO, Pires JL (2016) Cacao families and parents selected as resistant to natural infection of *Moniliophthora perniciosa*. Crop Breed Appl Biotechnol 16:141–146

Bidot Martínez I, Valdés de la Cruz M, Riera Nelson M, Bertin P (2017) Morphological characterization of traditional cacao (*Theobroma cacao* L.) plants in Cuba. Genet Resour Crop Evol 64(1):73–99

Borrelli VMG, Brambilla V, Rogowsky P, Marocco A, Lanubile A (2018) The enhancement of plant disease resistance using CRISPR/Cas9 technology (review). Front Plant Sci 9(1245)

Boza EJ, Irish BM, Meerow AW, Tondo CL, Rodríguez OA et al (2013) Genetic diversity, conservation, and utilization of *Theobroma cacao* L.: genetic resources in the Dominican Republic. Genet Resour Crop Evol 60(2):605–619

Brown J, Schnell R, Motamayor J, Lopes U, Kuhn D et al (2005) Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an F2 population using SSR markers and candidate genes. J Amer Soc Hort Sci 130:366–373

Brown JS, Phillips-Mora W, Power EJ, Krol C, Cervantes-Martinez C et al (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. Crop Sci 47(5):1851–1858

Brown JS, Sautter RT, Olano CT, Borrone JW, Kuhn DN et al (2008) A composite linkage map from three crosses between commercial clones of cacao, *Theobroma Cacao* L. Trop Plant Biol 1(2):120

Cabrera OG, Molano EPL, José J, Álvarez JC, Pereira GAG (2016) Ceratocystis wilt pathogens: history and biology—highlighting *C. cacaofunesta*, the causal agent of wilt disease of cacao. In: Bailey B, Meinhardt L (eds) Cacao diseases. Springer, Cham, pp 383–428

Chingandu N, Zia-Ur-Rehman M, Sreenivasan TN, Surujdeo-Maharaj S, Umaharan P et al (2017) Molecular characterization of previously elusive badnaviruses associated with symptomatic cacao in the New World. Arch Virol 162(5):1363–1371

Clement D, Risterucci AM, Motamayor JC, N'Goran J, Lanaud C (2003) Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. Genome 46(2):204–212

Clement D, Lanaud C, Sabau X, Fouet O, Le Cunff L et al (2004) Creation of BAC genomic resources for cocoa (*Theobroma cacao* L.) for physical mapping of RGA containing BAC clones. Theor Appl Genet 108(8):1627–1634

Clement C, De Cristo-Araújo M, Coppens D'Eeckenbrugge G, Alves Pereira A, Picanço-Rodrigues D (2010) Origin and domestication of native amazonian crops. Diversity 2(1):72–106

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics (Review). Front Ecol Evol 6(165)

Crouzillat D, Phillips W, Fritz PJ, Pétiard V (2000) Quantitative trait loci analysis in *Theobroma cacao* using molecular markers. Inheritance of polygenic resistance to *Phytophthora palmivora* in two related cacao populations. Euphytica 114(1):25–36

Cuatrecasas J (1964) Cacao and its allies, a taxonomic revision of the genus *Theobroma*. In: Systematic plant studies contributions from the United States national herbarium. Smithsonian Institution Press, Washington, DC, pp 379–614

da Hora Junior BT, de Faria Poloni J, Lopes MA, Dias CV, Gramacho KP et al (2012) Transcriptomics and systems biology analysis in identification of specific pathways involved in cacao resistance and susceptibility to witches' broom disease. Mol Biosyst 8(5):1507–1519

da Silva RA, Souza G, Lemos LSL, Lopes UV, Patrocínio NGRB et al (2017) Genome size, cytogenetic data and transferability of EST-SSRs markers in wild and cultivated species of the genus *Theobroma* L. (Byttnerioideae, Malvaceae). PLos One 12(2):e0170799

Dantas LG, Guerra M (2010) Chromatin differentiation between *Theobroma cacao* L. and *T. grandiflorum* schum. Genet Mol Biol 33(1):94–98

de Albuquerque PSB, Silva SDVM, Luz EDMN, Pires JL, Vieira AMC et al (2010) Novel sources of witches' broom resistance (causal agent *Moniliophthora perniciosa*) from natural populations of *Theobroma cacao* from the Brazilian Amazon. Euphytica 172(1):125–138

de Oliveira ML, Luz EDMN (2005) Identificação e manejo das principais doenças do cacaueiro no Brasil. CEPLAC/CEPEC/SEFIT, 132 pp

Decloquement J, Ramos-Sobrinho R, Elias SG, Britto DS, Puig AS et al (2021) *Phytophthora theobromicola* sp. nov.: a new species causing black pod disease on cacao in Brazil. Front Microbiol 12:537399

Dennis ES, Ellis J, Green A, Llewellyn D, Morell M et al (2008) Genetic contributions to agricultural sustainability. Philos Trans R Soc Lond B Biol Sci 363(1491):591–609

Dias LAS (2001) Melhoramento genético do cacaueiro. FUNAPE-UFG, Viçosa, Minas Gerais, Brazil, 578 pp

Díaz-Valderrama JR, Aime MC (2016) The cacao pathogen *Moniliophthora roreri* (Marasmiaceae) produces rhexolytic thallic conidia and their size is influenced by nuclear condition. Mycoscience 57(3):208–216

Doaré F, Ribeyre F, Cilas C (2020) Genetic and environmental links between traits of cocoa beans and pods clarify the phenotyping processes to be implemented. Sci Rep 10(1):9888

dos Santos EC, Pirovani CP, Correa SC, Micheli F, Gramacho KP (2020) The pathogen *Moniliophthora perniciosa* promotes differential proteomic modulation of cacao genotypes with contrasting resistance to witches' broom disease. BMC Plant Biol 20(1):1

Efombagn MIB, Sounigo O, Eskes AB, Motamayor JC, Manzanares-Dauleux MJ et al (2009) Parentage analysis and outcrossing patterns in cacao (*Theobroma cacao* L.) farms in Cameroon. Heredity 103(1):46–53

Elliot A, Kennedy AJ (1988) The use of isozyme analysis in the characterization of cocoa germplasm. Paper presented at the annual seminar on agricultural research, St. Augustine (Trinidad y Tobago)

Elmer W, White JC (2018) The future of nanotechnology in plant pathology. Annu Rev Phytopathol 56(1):111–133

End MJ, Daymond AJ, Hadley P (2017) Technical guidelines for the safe movement of cacao germplasm. In: FAO/IPGRI technical guidelines, 89 pp

Engelbrecht CJ, Harrington TC (2005) Intersterility, morphology and taxonomy of Ceratocystis fimbriata on sweet potato, cacao and sycamore. Mycologia 97(1):57–69

Engelbrecht CJ, Harrington TC, Alfenas A (2007a) Ceratocystis wilt of cacao—a disease of increasing importance. Phytopathol 97(12):1648–1649

Engelbrecht CJB, Harrington TC, Alfenas AC, Suarez C (2007b) Genetic variation in populations of the cacao wilt pathogen, *Ceratocystis Cacaofunesta.* Plant Pathol 56(6):923–933

Engels JMM (1986) The identification of cacao cultivars. Paper presented at the international society for horticultural science, Leuven, Belgium

Eskes AB, Rodriguez CAC, Condori DC, Seguine E, Carrion LFG et al (2018) Large genetic diversity for fine-flavor traits unveiled in cacao (*Theobroma cacao* L.) with special attention to the native chuncho variety in cusco, Peru. Agrotropica 30(3)

Evans HCP (1981) OD root of cacao caused by *Moniliophthora roreri*. In: Phytopathological papers. C.M.I KEW, Surrey, England

Faleiro FG, Lopes UV, Yamada MM, Melo GRP, Monteiro WR et al (2004a) Genetic diversity of cacao accessions selected for resistance to witches' broom disease based on RAPD markers. Crop Breed Appl Biotechnol 4:12–17

Faleiro ASG et al (2004b) Variability in cacao selected by producers for resistance to witches' broom based on microsatellite markers. Crop Breed Appl Biotechnol 4:3

Faleiro FG, Queiroz VT, Lopes UV, Guimarães CT, Pires JL et al (2006) Mapping QTLs for witches' broom (*Crinipellis perniciosa*) resistance in cacao (*Theobroma cacao* L.). Euphytica 149(1–2):227–235

Farrell AD, Rhiney K, Eitzinger A, Umaharan P (2018) Climate adaptation in a minor crop species: is the cocoa breeding network prepared for climate change? Agroecol Sustain Food Syst 42(7):812–833

Fernandes LDS, Royaert S, Corrêa FM, Mustiga GM, Marelli J-P et al (2018) Mapping of a major QTL for ceratocystis wilt disease in an F1 population of *Theobroma cacao* (Original Research). Front Plant Sci 9(155)

Figueira A, Janik J, Goldsbrough P (1992) Genome size and DNA polymorphism in *Theobroma cacao*. J Am Soc Hortic Sci 117:673–677

Figueiredo GSF, Melo CAF, Souza MM, Araújo IS, Zaidan HA et al (2013) Karyotype variation in cultivars and spontaneous cocoa mutants (*Theobroma cacao* L.). Genet Mol Res 12(4)

Figueroa DM, Bass HW (2010) A historical and modern perspective on plant cytogenetics. Brief Funct Genom 9(2):95–102

Fister AS, O'Neil ST, Shi Z, Zhang Y, Tyler BM et al (2015) Two *Theobroma cacao* genotypes with contrasting pathogen tolerance show aberrant transcriptional and ROS responses after salicylic acid treatment. J Exp Bot:6245–6258

Fister AS, Mejia LC, Zhang Y, Herre EA, Maximova SN et al (2016) *Theobroma cacao* L. pathogenesis-related gene tandem array members show diverse expression dynamics in response to pathogen colonization. BMC Genomics 17:363–363

Fister AS, Landherr L, Maximova SN, Guiltinan MJ (2018) Transient expression of CRISPR/Cas9 machinery targeting *TcNPR3* enhances defense response in *Theobroma cacao* (Original Research). Front Plant Sci 9(268)

Fister AS, Leandro-Muñoz ME, Zhang D, Marden JH, Tiffin P et al (2019) Widely distributed variation in tolerance to *Phytophthora palmivora* in four genetic groups of cacao. Tree Genet Genomes 16(1):1

Flament MH, Kebe I, Clément D, Pieretti I, Risterucci AM et al (2001) Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. Genome 44(1):79–85

Fouet O, Allegre M, Argout X, Jeanneau M, Lemainque A et al (2011) Structural characterization and mapping of functional EST-SSR markers in *Theobroma cacao*. Tree Genet Genomes 7(4):799–817

Franzen M, Borgerhoff Mulder M (2007) Ecological, economic and social perspectives on cocoa production worldwide. Biodivers Conserv 16(13):3835–3849

Friscina A, Chiappetta L, Jacquemond M, Tepfer M (2017) Infection of non-host model plant species with the narrow-host-range Cacao swollen shoot virus. Mol Plant Pathol 18(2):293–297

Fuccillo D, Sears L, Stapleton P (2007) Biodiversity in trust: conservation and use of plant genetic resources in CGIAR centres. Cambridge University Press, Cambridge, p 372

Gardea AA, García-Bañuelos ML, Orozco-Avitia JA, Sánchez-Chávez E, Sastré-Flores B et al (2017) Cacao (*Theobroma cacao* L.). In: Yahia EM (ed) Fruit and vegetable phytochemicals. Wiley-Blackwell, Chichester, West Sussex, United Kingdom, pp 921–940

Gardella DS, Enriquez GA, Saunders JL (1982) Inheritance of clonal resistanceto *Ceratocystis fimbriata* in cacao hybrids. Paper presented at the 8th international cocoa research conference, Cartagena, Colombia

Geering ADW, Hull R (2012) Family caulimoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy 9th report of the international committee on taxonomy of viruses. Press EA, London, UK, pp 429–443

Gesteira AS, Micheli F, Carels N, Da Silva AC, Gramacho KP et al (2007) Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa*. Ann Bot 100(1):129–140

Giani AM, Gallo GR, Gianfranceschi L, Formenti G (2020) Long walk to genomics: history and current approaches to genome sequencing and assembly. Comput Struct Biotechnol J 18:9–19

González-Orozco CE, Galán AAS, Ramos PE, Yockteng R (2020) Exploring the diversity and distribution of crop wild relatives of cacao (*Theobroma cacao* L.) in Colombia. Genet Resour Crop Evol 67(8):2071–2085

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD et al (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40(Database issue):D1178–1186

Gopaulchan D, Motilal LA, Bekele FL, Clause S, Ariko JO et al (2019) Morphological and genetic diversity of cacao (*Theobroma cacao* L.) in Uganda. Physiol Mol Biol Plants 25(2):361–375

Gramacho KP (2003) Development of microsatellitemarkers for the genetic analysis of *Crinipellis perniciosa.* Paper presented at the 14th international cocoa research conference, Ghana

Gramacho KP, Pires JL, Clément D, Juca FF, Melo GRP et al (2012) Breakdown of scavina resistance in Bahia caused by the evolution of *Moniliophthora perniciosa.* Paper presented at the 17th Conférence Internationale Sur La Recherche Cacaoyère, Yaounde, Cameroun

Griffith GW, Hedger JN (1994a) The breeding biology of biotypes of the witches' broom pathogen of cocoa, *Crinipellis Perniciosa.* Heredity 72(3):278–289

Griffith GW, Hedger JN (1994b) Spatial distribution of mycelia of the liana (L-) biotype of the agaric *Crinipellis perniciosa* (Stahel) Singer in tropical forest. New Phytol 127(2):243–259

Griffith GW, Nicholson J, Nenninger A, Birch RN, Hedger JN (2003) Witches' brooms and frosty pods: two major pathogens of cacao. N Z J Bot 41(3):423–435

Guest D (2007) Black pod: diverse pathogens with a global impact on cocoa yield. Phytopathol 97(12):1650–1653

Guest D, Keane P (2018) Cacao diseases: vascular-streak dieback. In: Umaharan P (ed) Achieving sustainable cultivation of cocoa. Burleigh Dodds Science Publishing, Cambridge, UK, pp 1654–1657

Gutiérrez OA, Puig AS, Phillips-Mora W, Bailey BA, Ali SS et al (2021) SNP markers associated with resistance to frosty pod and black pod rot diseases in an F1 population of *Theobroma cacao* L. Tree Genet Genomes 17(3):28

Hämälä T, Guiltinan MJ, Marden JH, Maximova SN, dePamphilis CW et al (2019) Gene expression modularity reveals footprints of polygenic adaptation in *Theobroma cacao*. Mol Biol Evol 37(1):110–123

Hammer K, Khoshbakht K (2005) Towards a 'red list' for crop plant species. Genet Resour Crop Evol 52(3):249–265

Helliwell EE, Vega-Arreguín J, Shi Z, Bailey B, Xiao S et al (2016) Enhanced resistance in *Theobroma cacao* against oomycete and fungal pathogens by secretion of phosphatidylinositol-3-phosphate-binding proteins. Plant Biotechnol J 14(3):875–886

Henry RJ (2005) Plant diversity and evolution: genotypic and phenotypic variation in higher plants. CABI Publishing, Wallingford, UK, viii + 332 pp

Henry RJ (2013) Molecular markers in plants. Wiley-Blackwell, Chichester, West Sussex, United Kingdom, 216 pp

Ho HH (2018) The taxonomy and biology of phytophthora and pythium. J Bacteriol Mycol 6:40–45

Hofberger JA, Nsibo DL, Govers F, Bouwmeester K, Schranz ME (2015) A complex interplay of tandem- and whole-genome duplication drives expansion of the L-type lectin receptor kinase gene family in the brassicaceae. Genome Biol Evol 7(3):720–734

Holderness M (1990) Control of vascular-streak dieback of cocoa with triazole fungicides and the problem of phytotoxicity. Plant Pathol 39(2):286–293

Holmes S, Lyman S, Hsu JK, Cheng J (2015) Making BAC transgene constructs with lambda-red recombineering system for transgenic animals or cell lines. Methods Mol Biol 1227:71–98

Hughes JA, Ollennu LAA (1994) Mild strain protection of cocoa in Ghana against cocoa swollen shoot virus—a review. Plant Pathol 43(3):442–457

Hwang H-H, Yu M, Lai E-M (2017) Agrobacterium-mediated plant transformation: biology and applications. The Arabidopsis Book 2017(15):e0186

Jaimes Y, Aránzazu F, Rodríguez EESLSC, Martínez N (2011) Behavior of introduced regional clones of *Theobroma cacao* toward the infection *Moniliophthora roreri* in three different regions of Colombia. Agron Colomb 29(2):171–178

Jaimez RE, Vera DI, Mora A, Loor RG, Bailey BA (2020) A disease and production index (DPI) for selection of cacao (*Theobroma cacao*) clones highly productive and tolerant to pod rot diseases. Plant Pathol 69(4):698–712

Ji K, Zhang D, Motilal LA, Boccara M, Lachenaud P et al (2013) Genetic diversity and parentage in farmer varieties of cacao (*Theobroma cacao* L.) from Honduras and Nicaragua as revealed by single nucleotide polymorphism (SNP) markers. Genet Resour Crop Evol 60(2):441–453

Jones P, Allaway D, Gilmour D, Harris C, Rankin D et al (2002) Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta 216(2):255–264

Keane PJ (1981) Epidemiology of vascular-streak dieback of cocoa. Ann Appl Biol 98(2):227–241

Kellam MK, Zentmyer GA (1986) Morphological, physiological, ecological, and pathological comparisons of *Phytophthora* species isolated from *Theobroma cacao*. Phytopathol 76(2)

Kordrostami M, Rahimi M (2015) Molecular markers in plants: concepts and applications. Genet Third Millennium 13:4024–4031

Kouakou K, Kébé BI, Kouassi N, Aké S, Cilas C et al (2012) Geographical distribution of cacao swollen shoot virus molecular variability in Côte d'Ivoire. Plant Dis 96(10):1445–1450

Krauss U, Hoopen MT, Hidalgo E, Martínez A, Arroyo C et al (2003) Manejo integrado de la moniliasis (*Moniliophthora roreri*) del cacao (*Theobroma cacao*) en Talamanca, Costa Rica. Agroforestería En Las Américas 10(37–38):52–58

Kudjordjie EN (2015) *Phytophthora megakarya* and *P. palmivora* on *Theobroma cacao*: aspects of virulence and the effects of temperature on growth and resistance to fungicides. Ph.D., University of Copenhagen

Kuhn DN, Livingstone D, Main D, Zheng P, Saski C et al (2012) Identification and mapping of conserved ortholog set (COS) II sequences of cacao and their conversion to SNP markers for marker-assisted selection in *Theobroma cacao* and comparative genomics studies. Tree Genet Genomes 8(1):97–111

Kumari N, Grimbs A, D'Souza RN, Verma SK, Corno M et al (2018) Origin and varietal based proteomic and peptidomic fingerprinting of *Theobroma cacao* in non-fermented and fermented cocoa beans. Food Res Int 111:137–147

Lachenaud P, Zhang D (2008) Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana. Ann Forest Sci 65(3):310–310

Lanaud C (1986) Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer: *Theobroma cacao* L. I. Contrôle génétique et "linkage" de neuf marqueurs génétiques (Article de revue). Café, Cacao, Thé 30(4):259–270

Lanaud C, Risterucci AM, N'Goran AKJ, Clement D, Flament MH et al (1995) A genetic linkage map of *Theobroma cacao* L. Theor Appl Genet 91(6):987–993

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A et al (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol Ecol 8(12):2141–2143

Lanaud C, Risterucci AM, Pieretti I, N'Goran JAK, Fargeas D (2004) Characterisation and genetic mapping of resistance and defence gene analogs in cocoa (*Theobroma cacao* L.). Mol Breed 13(3):211–227

Lanaud C, Fouet O, Clément D, Boccara M, Risterucci AM et al (2009) A meta–QTL analysis of disease resistance traits of *Theobroma cacao* L. Mol Breed 24(4):361–374

Leal JB, dos Santos LM, dos Santos CAP, Pires JL, Ahnert D et al (2008) Diversidade genética entre acessos de cacau de fazendas e de banco de germoplasma na Bahia. Pesqui Agropecu Bras 43(7):851–858

Lembang RP, Mustari K, Asrul L (2019) Identification and analysis of relationship several local promising clones cocoa (*Theobroma cacao* L.) based on morphological characters in South Sulawesi. IOP Conf Ser: Earth Environ Sci 270:012024

Lerceteau E, Robert T, Pétiard V, Crouzillat D (1997) Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. Theor Appl Genet 95(1):10–19

Li P, Quan X, Jia G, Xiao J, Cloutier S et al (2016) RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. BMC Genomics 17(1):852

Li Y, Zhang D, Motilal LA, Lachenaud P, Mischke S et al (2021) Traditional varieties of cacao (*Theobroma cacao*) in Madagascar: their origin and dispersal revealed by SNP markers. Beverage Plant Res 1:4

Lima EM et al (2013) Genetic molecular diversity, production and resistance to witches' broom in cacao clones. Crop Breed Appl Biotechnol 13:2

Lindo AA, Robinson DE, Tennant PF, Meinhardt LW, Zhang D (2018) Molecular characterization of cacao (*Theobroma cacao*) germplasm from Jamaica using single nucleotide polymorphism (SNP) markers. Trop Plant Biol 11(3):93–106

Lisboa DO, Evans HC, Araújo JPM, Elias SG, Barreto RW (2020) *Moniliophthora perniciosa*, the mushroom causing witches' broom disease of cacao: insights into its taxonomy, ecology and host range in Brazil. Fungal Biol 124(12):983–1003

Livingstone D, Stack C, Mustiga GM, Rodezno DC, Suarez C et al (2017) A larger chocolate chip-development of a 15K *Theobroma cacao* L. SNP array to create high-density linkage maps. Front Plant Sci 8:2008–2008

Lockwood G (1981) Genetic aspects of resistance to cocoa swollen-shoot virus in Ghana. Ann Appl Biol 98(1):131–141

Loor RG, Risterucci AM, Courtois B, Fouet O, Jeanneau M et al (2009) Tracing the native ancestors of the modern *Theobroma cacao* L. population in Ecuador. Tree Genet Genomes 5(3):421–433

Lopes UV, Monteiro WR, Pires JL, Clement D, Yamada MM et al (2011) Cacao breeding in Bahia, Brazil—strategies and results. Crop Breed Appl Biotechnol 1:73–81

Luo M, Wing RA (2003) An improved method for plant BAC library construction. Methods Mol Biol 236:3–20

Luz EDMN, Santos AF, Matsuoka K, Bezerra JL (2001) Doenças causadas por Phytophthora no Brasil. Livraria e Editora Rural Ltda, 754 pp

Mali P, Yang L, Esvelt KM, Aach J, Guell M et al (2013) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826

Marelli J-P, Maximova SN, Gramacho KP, Kang S, Guiltinan MJ (2009) Infection biology of *Moniliophthora perniciosa* on *Theobroma cacao* and alternate solanaceous hosts. Trop Plant Biol 2(3–4):149–160

Marelli J-P, Guest DI, Bailey BA, Evans HC, Brown JK et al (2019) Chocolate under threat from old and new cacao diseases. Phytopathol 109(8):1331–1343

Marita JM, Nienhuis J, Pires JL, Aitken WM (2001) Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches' broom disease resistance. Crop Sci 41(4):1305–1316

Maximova S, Miller C, Antunez de Mayolo G, Pishak S, Young A et al (2003) Stable transformation of *Theobroma cacao* L. and influence of matrix attachment regions on GFP expression. Plant Cell Rep 21(9):872–883

Maximova S, Marelli J-P, Young A, Pishak S, Verica J et al (2006) Over-expression of a cacao class I chitinase gene in *Theobroma cacao* L. enhances resistance against the pathogen *Colletotrichum gloeosporioides*. Planta 224(4):740–749

McElroy MS, Navarro AJR, Mustiga G, Stack C, Gezan S et al (2018) Prediction of cacao (*Theobroma cacao*) resistance to *Moniliophthora* spp. diseases via genome-wide association analysis and genomic selection. Front Plant Sci 9:343–343

McMahon P, Purwantara A (2016) Vascular streak dieback (*Ceratobasidium theobromae*): history and biology. In: Bailey BA, Meinhardt LW (eds) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, Cham, pp 307–335

McMahon P, Iswanto A, Susilo AW, Sulistyowati E, Wahab A et al (2009) On-farm selection for quality and resistance to pest/diseases of cocoa in Sulawesi: (i) performance of selections against cocoa pod borer, *Conopomorpha Cramerella.* Int J Pest Manag 55(4):325–337

McMahon P, Purung HB, Lambert S, Mulia S, Nurlaila et al (2015) Testing local cocoa selections in three provinces in Sulawesi: (i) productivity and resistance to cocoa pod borer and Phytophthora pod rot (black pod). Crop Prot 70:28–39

Mondego JMC, Carazzolle MF, Costa GGL, Formighieri EF, Parizzi LP et al (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into witches' broom disease of cacao. BMC Genomics 9(1):548

Monteiro Reis GS, Furtado de Almeida A-A, de Almeida SN, Santana JO, de Souza M, Da A et al (2020) Proteomic profiles of young and mature cocoa leaves subjected to mechanical stress caused by wind. Plant Physiol Biochem 155:851–867

Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A et al (2002) Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89(5):380–386

Motamayor JC, Risterucci AM, Heath M, Lanaud C (2003) Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. Heredity 91(3):322–330

Motamayor JC, Lachenaud P, da Silva e Mota JW, Loor R, Kuhn DN et al (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). PLos One 3(10):e3311

Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Iii DL et al (2013) The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol 14(6):r53

Motilal LA, Zhang D, Mischke S, Meinhardt LW, Boccara M et al (2016) Association mapping of seed and disease resistance traits in *Theobroma cacao* L. Planta 244(6):1265–1276

Muller E (2016) Cacao swollen shoot virus (CSSV): history, biology, and genome. In: Bailey BA, Meinhardt LW (eds) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, Cham, pp 337–358

Mustiga GM, Gezan SA, Phillips-Mora W, Arciniegas-Leal A, Mata-Quirós A et al (2018) Phenotypic description of *Theobroma cacao* L. for yield and vigor traits from 34 hybrid families in Costa Rica based on the genetic basis of the parental population (Original Research). Front Plant Sci 9(808)

N'Goran JAK, Laurent V, Risterucci AM, Lanaud C (1994) Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. Heredity 73(6):589–597

N'Guessan KF, N'Goran JAK, Eskes AB (2008) Resistance of cacao (*Theobroma cacao* L.) to *Sahlbergella singularis* (Hemiptera: Miridae): investigation of antixenosis, antibiosis and tolerance. Int J Trop Insect Sci 28(4):201–210

Naganeeswaran S, Elain Apshara S, Manimekalai R, Vasu A, Malhotra SK (2015) Cocoa EST database: comprehensive database of cocoa expressed sequence tags (ESTs). IJIRCCE 3(11):10441–10444

Nevo E (1988) Genetic diversity in nature. In: Hecht MK, Wallace B (eds) Evolutionary biology. Springer, US, Boston, MA, pp 217–246

Niemenak N, Kaiser E, Maximova SN, Laremore T, Guiltinan MJ (2015) Proteome analysis during pod, zygotic and somatic embryo maturation of *Theobroma cacao*. J Plant Physiol 180:49–60

Niogret J, Ekayanti A, Kendra PE, Ingram K, Lambert S et al (2020) Host preferences of the cocoa pod borer, *Conopomorpha cramerella*, the main threat to cocoa production in Southeast Asia. Entomol Exp Appl 168(3):221–227

Noah AM, Niemenak N, Sunderhaus S, Haase C, Omokolo DN et al (2013) Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L. J Proteom 78:123–133

Nyadanu D, Akromah R, Adomako B, Kwoseh C, Lowor ST et al (2012) Inheritance and general combining ability studies of detached pod, leaf disc and natural field resistance to *Phytophthora palmivora* and *Phytophthora megakarya* in cacao (*Theobroma cacao* L.). Euphytica 188(2):253–264

Nyassé S, Efombagn MIB, Kébé BI, Tahi M, Despréaux D et al (2007) Integrated management of *Phytophthora* diseases on cocoa (*Theobroma cacao* L.): impact of plant breeding on pod rot incidence. Crop Prot 26(1):40–45

Ofori A, Padi FK, Ameyaw GA, Dadzie AM, Lowor S (2015) Genetic variation among cocoa (*Theobroma cacao* L.) families for resistance to cocoa swollen shoot virus disease in relation to total phenolic content. Plant Breed 134(4):477–484

Oliveira TAS et al (2014) Fatores epidemiológicos de *Phytophthora palmivora* afetando a severidade da podridão-dos-frutos do mamoeiro na pós-colheita. Summa Phytopathol 40(3):256–263

Olufemi AK, Olatunde FA, Adewale AS, Mohammed I, Osasogie U et al (2020) Effect of high density planting on the vigour and yield of *Theobroma cacao* L. in the Southwest of Nigeria. Na J Adv Res Rev 8(1):217–223

Orea MDD, Romero-Cortes T, Lopez-Perez PA, Perez Espana VH, Ramirez-Lepe M et al (2017) Current status of cocoa frosty pod rot caused by *Moniliophthora roreri* and a phylogenetic analysis. Plant Pathol J 16:41–53

Oro F, Mississo E, Okassa M, Guilhaumon C, Fenouillet C et al (2012) Geographical differentiation of the molecular diversity of cacao swollen shoot virus in Togo. Arch Virol 157(3):509–514

Osorio-Guarín JA, Berdugo-Cely J, Coronado RA, Zapata YP, Quintero C et al (2017) Colombia a source of cacao genetic diversity as revealed by the population structure analysis of germplasm bank of *Theobroma cacao* L. (Original Research). Front Plant Sci 8(1994)

Osorio-Guarín JA, Berdugo-Cely JA, Coronado-Silva RA, Baez E, Jaimes Y et al (2020) Genome-wide association study reveals novel candidate genes associated with productivity and disease resistance to *Moniliophthora* spp. in cacao (*Theobroma cacao* L.). Genes Genomes Genet 10(5):1713–1725

Padi FK, Domfeh O, Takrama J, Opoku S (2013) An evaluation of gains in breeding for resistance to the cocoa swollen shoot virus disease in Ghana. Crop Prot 51:24–31

Patrocínio NGRB, Ceresini PC, Gomes LIS, Resende MLV, Mizubuti ESG et al (2017) Population structure and migration of the witches' broom pathogen *Moniliophthora perniciosa* from cacao and cultivated and wild solanaceous hosts in southeastern Brazil. Plant Pathol 66(6):900–911

Phillips-Mora W (2003) Origin, biogeography, genetic diversity and taxonomic affinities of the cacao (*Theobroma cacao* L.) fungus *Moniliophthora roreri* (Cif.) Evans et al. as determined using molecular, phytopathological and morpho-physiological evidence. Ph.D., University of Reading

Phillips-Mora W, Wilkinson MJ (2007) Frosty pod of cacao: a disease with a limited geographic range but unlimited potential for damage. Phytopathol 97(12):1644–1647

Phillips-Mora W, Castillo J, Krauss U, Rodríguez E, Wilkinson MJ (2005) Evaluation of cacao (*Theobroma cacao*) clones against seven Colombian isolates of *Moniliophthora roreri* from four pathogen genetic groups. Plant Pathol 54(4):483–490

Phillips-Mora W, Castillo J, Arciniegas A, Mata A, Sánchez A et al (2009) Overcoming the main limiting factors of cacao production in Central America through the use of improved clones developed at CATIE. Paper presented at the 16th International cocoa research conference, COPAL, Bali, Indonesia

Phillips-Mora W, Mata-Quirós A, Arciniegas-Leal A (2018) Generation of cacao clones with durable resistant against frosty pod rot (*Moniliophthora roreri* (Cif. & Par.) Evans et al.) Paper presented at the international symposium on cocoa research (ISCR), Lima, Peru

Pinheiro TT, Litholdo CG Jr, Sereno ML, Leal GA Jr, Albuquerque PSB et al (2012) Establishing references for gene expression analyses by RT-qPCR in *Theobroma cacao* tissues. Genet Mol Res 10(4):3291–3305

Pires JL (2003) Avaliação quantitativa e molecular de germoplasma para o melhoramento do cacaueiro com ênfase na produtividade, qualidade de frutos e resistência a doenças. Ph.D., Universidade Federal de Viçosa

Ploetz RC (2007) Cacao diseases: important threats to chocolate production worldwide. Phytopathol 97(12):1634–1639

Posnette AF, Robertson NF, Todd JM (1950) Virus diseases of cacao in West Africa. Ann Appl Biol 37(2):229–240

Pound FJ (1938) Cacao and witches' broom disease of South America, with notes on other species of Theobroma: report on a visit to Equator, the Amazon valley and Colombia. CABI Publisher, Port of Spain, p 58

Pound FJ (1943) Cacao and witches' broom disease (*Marasmius perniciosus*). In: Rhodes AL (ed) Report on a recent visit to the Amazon territory of Peru, p 14

Prior C (1979) Resistance of cocoa to vascular-streak dieback disease. Ann Appl Biol 92(3):369–376

Pugh T, Fouet O, Risterucci AM, Brottier P, Abouladze M et al (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. Theor Appl Genet 108(6):1151–1161

Purdy L, Schmidt R (1996) Status of cacao witches' broom: biology, epidemiology, and management. Annu Rev Phytopathol 34(1):573–594

Queiroz VT, Guimarães CT, Anhert D, Schuster I, Daher RT et al (2003) Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. Plant Breed 122(3):268–272

Ram A, Valle RR, Gardini EA (2004) Monília do Cacaueiro. Fundação Cargil, 35 pp

Ray S, Satya P (2014) Next generation sequencing technologies for next generation plant breeding (opinion). Front Plant Sci 5(367)

Reges BM, Maia AM, Sarmento DHA, Silva MS, dos Santos SML et al (2021) Quality variables for technological application of cocoa clones from the Brazilian semiarid region. Rev Bras Eng Agric Ambient 25(3)

Resende MDV (2002) Genética biométrica e Estatística no melhoramento de plantas perenes. Embrapa Florestas, Colombo, p 975

Risterucci AM, Paulin D, Ducamp M, N'Goran JAK, Lanaud C (2003) Identification of QTLs related to cocoa resistance to three species of phytophthora. Theor Appl Genet 108(1):168–174

Rodrigues GS, Pires JL, Luz EDMN (2020) Association of genes from different sources of resistance to major cacao diseases. Revista Ceres 67(5):383–394

Roivainen O (1976) Transmission of cocoa viruses by mealybugs (Homoptera: Pseudococcidae). Agric Food Sci 48(3):203–304

Romero Navarro JA, Phillips-Mora W, Arciniegas-Leal A, Mata-Quirós A, Haiminen N et al (2017) Application of genome wide association and genomic prediction for improvement of cacao productivity and resistance to black and frosty pod diseases (Original Research). Front Plant Sci 8(1905)

Rorer JB (1918) Enfermedades y plagas del cacao en el Ecuador y metodos modernos apropiados al cultivo del cacao. Guayaquil, 80 pp

Royaert S, Phillips-Mora W, Arciniegas Leal AM, Cariaga K, Brown JS et al (2011) Identification of marker-trait associations for self-compatibility in a segregating mapping population of *Theobroma cacao* L. Tree Genet Genomes 7(6):1159–1168

Royaert S, Jansen J, da Silva DV, de Jesus Branco SM, Livingstone DS et al (2016) Identification of candidate genes involved in witches' broom disease resistance in a segregating mapping population of *Theobroma cacao* L. in Brazil. BMC Genomics 17:107–107

Samuels GJ, Ismaiel A, Rosmana A, Junaid M, Guest D et al (2012) Vascular streak dieback of cacao in Southeast Asia and Melanesia: in planta detection of the pathogen and a new taxonomy. Fungal Biol 116(1):11–23

Sanches CLG, Pinto LRM, Pomella AWV, Silva SDVM, Loguercio LL (2008) Assessment of resistance to *Ceratocystis cacaofunesta* in cacao genotypes. Eu J Plant Pathol 122(4):517–528

Santana Silva RJ, Micheli F (2020) RRGPredictor, a set-theory-based tool for predicting pathogen-associated molecular pattern receptors (PRRs) and resistance (R) proteins from plants. Genomics 112(3):2666–2676

Santos AA (2019) Variabilidade morfoagronômica e resistência a doenças em variedades locais de cacau no sul da Bahia. Ph.D., Universidade Estadual de Santa Cruz

Santos RMF, Lopes UV, Bahia RC, Machado RCR, Ahnert D et al (2007) Marcadores microssatélites relacionados com a resistência à vassoura-de-bruxa do cacaueiro. Pesqui Agropecu Bras 42(8):1137–1142

Santos RC, Pires JL, Correa RX (2012a) Morphological characterization of leaf, flower, fruit and seed traits among Brazilian *Theobroma* L. species. Genet Resour Crop Evol 59(3):327–345

Santos RF, Lopes U, Silva SVM, Micheli F, Clement D et al (2012b) Identification of quantitative trait loci linked to Ceratocystis wilt resistance in cacao. Mol Breed 30(4):1563–1571

Santos R, Silva S, Sena K, Micheli F, Gramacho K (2013) Kinetics and histopathology of the cacao-*Ceratocystis cacaofunesta* interaction. Trop Plant Biol 6(1):37–45

Santos ESL, Cerqueira-Silva CBM, Mori GM, Ahnert D, Mello DLN et al (2015) Genetic structure and molecular diversity of cacao plants established as local varieties for more than two centuries: the genetic history of cacao plantations in Bahia, Brazil. PLoS ONE 10(12):e0145276

Sarillana ZC, Fundador EO, Fundador NG (2021) Synthesis of ZnO nanoparticles using *Theobroma cacao* L. pod husks, and their antibacterial activities against foodborne pathogens. Int Food Res J 28(1):102–109

Saunders JA, Mischke S, Leamy EA, Hemeida AA (2004) Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao.* Theor Appl Genet 110(1):41–47

Schneider M, Andres C, Trujillo G, Alcon F, Amurrio P et al (2016) Cocoa and total system yields of organic and conventional agroforestery vs. monoculture systems in a long-term field trial in Bolivia. Exp Agric 53(3):351–374

Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W et al (2007) Development of a marker assisted selection program for cacao. Phytopathol 97(12):1664–1669

Schultes RE (1958) A synopsis of the genus Herrania. J Arnold Arbor 39:217–295

Scollo E, Neville D, Oruna-Concha MJ, Trotin M, Cramer R (2018) Characterization of the proteome of *Theobroma cacao* beans by nano-UHPLC-ESI MS/MS. Proteomics 18(3–4):1700339

Scotton DC, Azevedo MDS, Sestari I, Da Silva JS, Souza LA et al (2017) Expression of the *Theobroma cacao* Bax-inhibitor-1 gene in tomato reduces infection by the hemibiotrophic pathogen *Moniliophthora perniciosa.* Mol Plant Pathol 18(8):1101–1112

Sena K, Alemanno L, Gramacho KP (2014) The infection process of *Moniliophthora perniciosa* in cacao. Plant Pathol:1272–1281

Shi Z, Maximova SN, Liu Y, Verica J, Guiltinan MJ (2010) Functional analysis of the *Theobroma cacao NPR1* gene in arabidopsis. BMC Plant Biol 10(1):248

Shi Z, Zhang Y, Maximova SN, Guiltinan MJ (2013) *TcNPR3* from *Theobroma cacao* functions as a repressor of the pathogen defense response. BMC Plant Biol 13:204

Silva S, Luz E, Almeida O, Gramacho K, Bezerra J (2002) Redescrição da sintomatologia causada por *Crinipellis perniciosa* em cacaueiro. Agrotropica 1:1–23

Silva SDVM, Paim MC, Castro WM (2004) Cacau "Jaca" resistente a *Ceratocystis fimbriata* na região cacaueira da Bahia, Brasil. Fitopatol Bras 29(5):538–540

Silva SDVM, Mandarino EP, Damaceno VO, Santos Filho LP (2007) Reação de genótipos de cacaueiros a isolados de *Ceratocystis cacaofunesta.* Fitopatol Bras 32:504–506

Silva SDVM, Lopes UV, Damaceno VO, Júnior AWDOR (2013) Seleção de clones de cacaueiros resistentes à murcha-de-ceratocystis em condições de campo. Agrotrópica 25:163–170

Silva Monteiro de Almeida D, Oliveira Jordão do Amaral D, Del-Bem L-E, Bronze dos Santos E, Santana Silva RJ et al (2017) Genome-wide identification and characterization of cacao WRKY transcription factors and analysis of their expression in response to witches' broom disease. PLos One 12(10):e0187346

Silva Neto PJ, Matos PGG, Martins ACS, Silva AP (2001) Sistema de produção de cacau para a Amazónia brasileira. CEPLAC/CEPEC/SEFIT, 125 pp

Soesilo AW, Mawardi S, Witjaksono, Mangoendidjojo W (2015) Pod characteristics of cocoa (*Theobroma cacao* L.) related to cocoa pod borer resistance (*Conopomorpha cramerella* Snell.). Pelita Perkebunan, Coffee and Cocoa Res J 31(1):1–3

Sounigo O, Umaharan R, Christopher Y, Sankar A, Ramdahin S (2005) Assessing the genetic diversity in the International Cocoa Genebank, Trinidad (ICG, T) using isozyme electrophoresis and RAPD. Genet Resour Crop Evol 52(8):1111–1120

Spence J, Bartley B (1966) Testing of breeding material of *Theobroma cacao* for resistance to black pod disease (*Phytophthora palmivora*). In: Second session of FAO technical working party on cocoa production and protection, Rome, Italy

Steven WF (1936) Swollen shoot and die-back-a new disease of cocoa. Gold Coast Farmer 5:144

Strickland AH (1950) The dispersal of pseudococcidae (Hemiptera-Homotepta) by air currents in the gold coast. Proc R Entomol Soc London Ser A, Gen Entomol 25(1–3):1–9

Suarez Contreras LY, Rangel Riaño AL (2013) Aislamiento de microorganismos para control biológico de *Moniliophthora roreri.* Acta Agron 62(4):370–378

Talboys PW, Brian PW, Garrett SD (1972) Resistance to vascular wilt fungi. Proc R Soc B: Biol Sci 181(1064):319–332

Tan G-Y, Tan W-K (1988) Genetic variation in resistance to vascular-streak dieback in cocoa (*Theobroma cacao*). Theor Appl Genet 75(5):761–766

Teh C-L, Pang JT-Y, Ho C-T (2006) Variation of the response of clonal cocoa to attack by cocoa pod borer *Conopomorpha cramerella* (Lepidoptera: Gracillariidae) in Sabah. Crop Prot 25(7):712–717

Teixeira PJPL, Thomazella DPT, Reis O, do Prado PFV, do Rio MCS et al (2014) High-resolution transcript profiling of the atypical biotrophic interaction between *Theobroma cacao* and the fungal pathogen *Moniliophthora perniciosa.* Plant Cell 26(11):4245–4269

Thatikayala D, Jayarambabu N, Banothu V, Ballipalli CB, Park J et al (2019) Biogenic synthesis of silver nanoparticles mediated by *Theobroma cacao* extract: enhanced antibacterial and photocatalytic activities. J Mater Sci Mater Electron 30(18):17303–17313

Thevenin J-M, Rossi V, Ducamp M, Doare F, Condina V et al (2012) Numerous clones resistant to Phytophthora palmivora in the "Guiana" genetic group of *Theobroma cacao* L. PLoS ONE 7(7):e40915

Thresh JM, Owusu GLK, Ollennu LAA (1988) Cocoa swollen shoot: an archetypal crowd disease/Afrikanische Sproßschwellung an Kakaobäumen: Eine archetypische "crowd"-Krankheit. J Plant Dis Prot 95(4):428–446

Trebissou CI, N'Guessan RK, Tahi MG, Aïdara S, Akaffou SD et al (2020) Response of forty (40) high-producing cacao genotypes to cacao swollen shoot disease. J Exp Agric Int 42(10):1–10

Utro F, Cornejo OE, Livingstone D, Motamayor JC, Parida L (2012) ARG-based genome-wide analysis of cacao cultivars. BMC Bioinform 13(19):S17

Vanhove W, Vanhoudt N, Van Damme P (2016) Biocontrol of vascular streak dieback (*Ceratobasidium theobromae*) on cacao (*Theobroma cacao*) through induced systemic resistance and direct antagonism. Biocontrol Sci Technol 26(4):492–503

Verica J, Maximova S, Strem M, Carlson J, Bailey B et al (2004) Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. Plant Cell Rep 23(6):404–413

Wang B, Motilal LA, Meinhardt LW, Yin J, Zhang D (2020) Molecular characterization of a cacao germplasm collection maintained in Yunnan, China using single nucleotide polymorphism (SNP) markers. Trop Plant Biol 13(4):359–370

Ward JA, Ponnala L, Weber CA (2012) Strategies for transcriptome analysis in nonmodel plants. Am J Bot 99(2):267–276

Warren J, Misir S, Kalai (1995) Isozyme markers for self-compatibility and yield in *Theobroma cacao* (cacao). Heredity 74(4):354–356

Went FAFC (1904) Krulloten en Versteende vruchten van de cacao in Suriname. Amsterdam Müller, 40 pp

Wickramasuriya AM, Dunwell JM (2018) Cacao biotechnology: current status and future prospects. Plant Biotechnol J 16(1):4–17

Wilde J, Waugh R, Powell W (1992) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor Appl Genet 83(6–7):871–877

Yamada MM, Faleiro FG, Flores AB et al (2009) Microsatellite diversity and heterozygosity of parents of cocoa breeding population. Crop Breed Appl Biotechnol 9:17–22

Yamada MM, Faleiro FG, Lopes UV, Pires JL (2015) Genetic parameters and resistance of cacao progenies to ceratocystis wilt under natural infection. Agrotrópica 27(3):215–218

Yang JY, Motilal LA, Dempewolf H, Maharaj K, Cronk QC (2011) Chloroplast microsatellite primers for cacao (*Theobroma cacao*) and other *Malvaceae.* Am J Bot 98(12):e372-374

Younis A, Ramzan F, Hwang Y-J, Lim K-B (2015) FISH and GISH: molecular cytogenetic tools and their applications in ornamental plants. Plant Cell Rep 34(9):1477–1488

Zarrillo S, Gaikwad N, Lanaud C, Powis T, Viot C et al (2018) The use and domestication of *Theobroma cacao* during the mid-Holocene in the upper Amazon. Nat Ecol Evol 2(12):1879–1888

Zhang D, Boccara M, Motilal L, Mischke S, Johnson ES et al (2009a) Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers. Tree Genet Genomes 5(4):595–607

Zhang D, Mischke S, Johnson ES, Phillips-Mora W, Meinhardt L (2009b) Molecular characterization of an international cacao collection using microsatellite markers. Tree Genet Genomes 5(1):1–10