Cultivating cacao (*Theobroma cacao* L.) has a long tradition in Colombia where cocoa is widely consumed as a beverage. It has become a priority of the Colombian government, as it is one of the crops promoted in the development of programs aimed at favoring peace in post-conflict regions and replacing crops formerly used for illicit purposes (Abbott *et al.* 2018).

*T. cacao* is native to the humid tropical forests of South America from the Amazon and Orinoco regions (Rojas and Sacristán 2013). Cacao trees grown from seeds produce their first flowers two years after germination, in the case of early genotypes (Mossu 1990). The pods usually require four to seven months to reach maturity; the number per tree is highly variable, usually ranging from zero to 200 (Wibaux *et al.* 2017). The percentage of flower-setting pods ranges from 0.5% to 5% (De Almeida and Valle 2010), and natural pollination is ensured by insects (Adjaloo and Oduro 2013, Glendinning 1972). Some authors consider dependence on wild pollinators as a limiting factor for cacao production (Bridgemohan *et al.* 2017), while others reject this view, asserting that self-incompatibility and physiological wilt are the more limiting factors (Hardy 1960).

In the absence of permanent shade, cacao trees usually reach stable production five years after planting (Paulin *et al.* 1993). Some authors recommend an evaluation of yield three years after stabilization of production (Carvalho *et al.* 2002). Clement *et al.* (1999) reported the existence of an association between early yielding genotypes and a rapid decline in their level of yield. The release of such *a priori* promising genotypes to farmers would lead to disappointing results in cacao farms, where cacao trees can be cultivated...
over a span of up to fifty years (Jagoret et al. 2011). Thus, the trade-off to be applied is between reliability and duration in evaluating a variety before its release to farmers.

The breeding strategy in several cocoa-producing countries has been based on mass selection of clones and on hybrid selection (Paulin and Eskes 1995). Several techniques have been adopted for the release of clones to farmers, such as rooted cuttings, grafting on adult trees in the field, and grafting on young plants in the nursery (Sena et al. 2015).

In many places, the potentially high yield level of improved varieties is strongly challenged by the parasitic pressure from pests and diseases, some of them controlled, but at a high environmental and economic cost, while others, especially the cocoa swollen shoot virus that occurs in African countries, remain totally out of control (Ploetz 2007).

The species T. cacao was first subdivided in two morpho-geographic subspecies, spp. cacao and spp. sphaerocarpum (Cuatrecasas 1964, Whitkus et al. 1998), corresponding respectively to the two main genetic groups, Criollo and Forastero (Whitkus et al. 1998). The genetic group known as Trinitario was generated by hybridization between Forastero and Criollo individuals in Trinidad (Motamayor et al. 2003). A more recent molecular study showed that cacao diversity could be classified in ten different genetic groups of which the Criollo group is the most genetically differentiated (Motamayor et al. 2008).

According to Lopez de Velasco, cited in Buriticá (1985), the first cocoa exporting region of the Americas was Maracaibo in northeastern Venezuela. The earliest known cacao plantation in Colombia dates from 1622 near the city of Cali, department of Valle del Cauca (Patiño 2002). The main type of cacao grown in Colombia until 1885 was Criollo. At that time, another type, Forastero, called “pajarito” (little bird), was introduced in Antioquia (Chavarriaga and Ochoa 1940). Due to the incidence of diseases, Criollo began to be replaced by Forastero trees. This is how cacao “pajarito”, which seemed to be more resistant to diseases, began to be distributed in the rest of the country (Garcés 1939).

The T. cacao cultivated area increased in Colombia from the second half of the eighteenth century (Alvim 1956). In 1889, the basin of the Palo River in the northern region of the department of Cauca, was one of the main cocoa-producing areas in Colombia (Palau 1889, cited in Patiño 2002). Until the early 1950s, the main cacao-growing regions of Colombia were Cauca, Valle del Cauca, and Huila, leading the production of cacao in the country (Garcia 1955). Fig. 1A shows the production of cacao in Colombia according to Alvim (1956).

According to Garcia (1997), in the mid-1950s, cacao plantations in Cauca and Valle del Cauca were extensively affected by attacks of the Ceratocystis sp.–Xyleborus sp. complex. The disease, known as “mal de machete” (machete disease), caused by Ceratocystis sp., was registered for the first time in the department of Huila in 1940 (Barros 1970). In 1955, the disease destroyed more than 50% of the plantation trees established in the departments of Valle del Cauca and Cauca, whose production represented, at that time, 75% of the national production (Barros 1970). As a consequence, production was greatly reduced in both departments.

Later, other departments came to be recognized as cacao bean producers. By 1983, the department of Santander was already recognized as the largest producer of cacao in the country, followed by the departments of Huila, Nariño and Antioquia (Cubillos 1983). According to information published by Fedecacao (National Cacao Producers Federation) (Fedecacao 2018), the most productive cacao departments in 2017 were, in the following descending order, Santander, Antioquia, Arauca, Huila, Tolima, and Nariño (Fig. 1B). A production of 60,535 tons was registered in 2017, reflecting an increase of 6.6% in relation to production of the previous year (Fedecacao 2018).

This review provides details about the history and perspectives of the cacao breeding program in Colombia, describing different cacao collecting expeditions carried out, as well as cacao breeding strategies that have been used throughout the history of cacao cultivation in Colombia. While Colombia is thought to be part of the center of the origins of cacao, it has made extensive use of genetic material introduced from other countries and only limited use of regional materials in breeding programs; therefore, the current breeding strategy seeks to make better use of available genetic resources. Recent developments in cacao breeding programs in African and American countries have been documented (Ahnert and Eskes 2018), but, despite being the fourth largest cacao producer in the Americas, cacao breeding in Colombia was not included, as the information was not previously available. To our knowledge, this is the first time that the whole history of cacao breeding programs in Colombia, as well as the current breeding strategy, is being documented and presented to both the national and the international scientific community.

![Fig. 1. Cocoa-producing departments in Colombia in the years 1956 (A) and 2017 (B). Numbers indicate cocoa production per department in thousands of metric tons.](image-url)
Collecting and conserving germplasm

The first *T. cacao* germplasm collecting project was carried out in the early 1940s. As a result, a collection of clones was obtained and coded SCP (Selection Cacao Palmira), and established in the Experimental Station of Palmira, Valle del Cauca (Oicatá 1986). Collecting trips continued to around 1948, resulting in new collections coded SCC (Selection Cacao Cauca) and SCT (Selection Cacao Tuluá) (García 1997).

In the early 1950s, Colombia’s Ministry of Agriculture organized the National Cacao Campaign. One of the activities of the initiative was to promote collecting expeditions. Between 1952 and 1953, the Colombian government, in collaboration with the Imperial College of Tropical Agriculture, organized the Anglo–Colombian cacao collecting expedition with the objective of collecting both wild and cultivated samples of *T. cacao* (Baker et al. 1953). The expedition was led by F.W. Cope, Professor of the Imperial College of Tropical Agriculture (ICTA), with the participation of Professor R.E.D. Baker, also from ICTA, D.J. Taylor, and R.E. Schultes, Professor at Harvard University, as well as some Colombian scientists. Cacao germplasm was predominantly collected along the banks of the Apaporis, Caquetá, Cagüán, Guaviare, Inírida, Putumayo and Vaupés rivers that cross the departments of Amazonas, Vaupés, Caquetá and Putumayo (Baker et al. 1953). Some areas of Valle del Cauca and El Chocó departments were also explored. As a result of the expedition, around 191 cacao accessions were obtained, corresponding to both spontaneous and cultivated trees, of which 63 were shipped to Trinidad. In addition to *T. cacao* germplasm, seven other species of the genus *Theobroma* and ten of the genus *Herrania* were collected (Zhang et al. 2011).

In 1960, the first cacao germplasm collection was established in Colombia by the Department of Agricultural Research (DIA) (no longer in existence) in the Experimental Station of Palmira, Valle del Cauca, the headquarters of the national cacao research program (Oicatá 1986). After establishing the germplasm collection, the first materials from Trinidad and Ecuador were introduced. In 1962, the Colombian Agricultural Institute (ICA) was created with the objective of coordinating research and extension activities in the agricultural sector. The germplasm collection was transferred to ICA in 1964, and the accessions previously coded SCP (Selection Cacao Palmira), SCC (Selection Cacao Cauca) and SCT (Selection Cacao Tuluá), were renamed as SC (Selection Colombia) (García 1997). In 1969, ICA established a germplasm collection of Criollo cacao at the Caribia Research Center, in Magdalena department on the north coast of the country (Oicatá 1986).

In 1977, Buriticá, Arboleda, and Correa carried out a collecting expedition, funded by ICA, in the Caguan river basin in Caquetá department (Buriticá 1985). The main objective was to collect material resistant to “witches’ broom” disease, caused by *Moniliophthora perniciosa*, and to “mal de machete”, caused by *Ceratocystis* sp. (Oicatá 1986). *T. cacao* was collected from wild trees and included in the ICA genebank in Palmira, Colombia, as ABC (Arboleda, Buriticá, Correa) accessions. In this collecting expedition, groups of wild trees were found concentrated in small areas visited by inhabitants of the region at harvest time, an interesting fact that suggests that the inhabitants of the region were “cacao harvesters” rather than “producers” (Buriticá 1985). New accessions were also collected in Huila, Putumayo and Amazonas departments.

In 1983 and 1984, collecting expeditions were organized by ICA in the Pacific region, in Nariño and Choco departments. Other species of genus *Theobroma*, such as *T. gileri* and *T. obovatum*, as well as *Herrania* species, such as *H. nycterodendron*, *H. albiflora* and *H. purpurea*, were also collected (Oicatá 1986).

In the mid-1980s, a collecting expedition in the basins of the Caquetá river was carried out with funding from the Colombian government, and coordinated by now defunct Administrative Department of Municipalities (DAINC) and the Araracuara Corporation (Ocampo 1985). The expedition covered the banks of the Caquetá river from Puerto Santander to Isla Colombia on the border with Brazil, and a stretch of the Apaporis River. The objective was to collect germplasm from *T. cacao* and other species of genus *Theobroma*. Samples of *T. cacao*, *T. bicolor*, *T. microcarpum*, and *Herrania* sp. were collected mostly from wild trees that were over 40 years old and 11 meters high on average (Ocampo 1985). Around 151 accessions were collected and coded EBC (Caquetá Botanical Expedition). Most of the EBC accessions were held in the ICA genebank at Palmira, Colombia, while a few were sent to the San Carlos genebank at the INIAP Napo-Payamino Experimental Station, Ecuador (Allen 1988).

In 1986–1987, new collecting expeditions were organized by the ICA Cacao Research Program in the following departments: Cundinamarca, Antioquia, Magdalena, Guajira, Huila, Norte de Santander and Tolima. These expeditions resulted in the collection of accessions coded from ICA 861307 to ICA 861769 (Turnbull and Hadley 2018).

In 1992, the Colombian government issued a decree aimed at restructuring ICA, and as a result, the Colombian Agricultural Research Corporation (Corpoica) was created. Projects and research centers were transferred to Corpoica from ICA, and the latter assumed responsibility for ensuring the health of crops and farm animals, as well as technical control of imports, exports, and the use of agricultural inputs in the country. In 1995, ICA sent budwoods from a collection of 210 accessions of cacao to Corpoica’s research center, La Suiza, located in the department of Santander. Currently, there are about 500 accessions in each of the germplasm collections established in Palmira and Santander, the increase in number of accessions being the result of activities organized and financed by Corpoica. Larger collections are maintained by foreign institutes, such as the Brazilian cocoa research center (CEPEC/CEPLAC), which has 1,300...
accessions from different countries in addition to regional and improved material (Lopes et al. 2011).

In 2018, Corpoica adopted the name Agrosavia while maintaining its commitment to research and transfer of technological innovation. In order to increase the genetic diversity of the Colombian T. cacao collection, new wild accessions were obtained during a recent germplasm collecting expedition, organized in September 2018, in a project called “Expedición Cacao-Bio”, funded by the Colombian Administrative Department of Science, Technology and Innovation (Colciencias). A team from Agrosavia and the Universidad de los Andes, sampled the diversity of cacao and its wild relatives in the departments of Caquetá and Putumayo. The aim of the study was to expand knowledge about the diversity associated with cacao and its wild relatives in a region recognized as a hotspot of diversity of the plant genera Theobroma and Herrania. To achieve this goal, sampling included not only vegetal material but also arthropods, phytopathogens, and microbiota associated with soil, mycorrhiza, trunks, and leaves. Researchers traveled on the Caguán and Caquetá rivers and selected five sites from which 151 samples of the genera Herrania and Theobroma were collected. Fig. 2 shows the approximate locations where germplasm collecting expeditions have taken place from the 1940s to the present.

Cacao breeding in Colombia: assessing germplasm and creating improved cultivars

Initial breeding strategies

Cacao selection programs began in 1930 in Trinidad and Nigeria (Paulin and Eskes 1995). In Colombia, the first reported cacao breeding effort consisted of assessment, in the experimental station of Palmira, of cacao trees collected during expeditions realized in the 1940s. The first selection criteria were the number of pods harvested per tree and the weight of fresh cacao beans per tree. For breeders, these selection criteria were insufficient for the selection of promising trees. For this reason, since 1948, additional criteria, such as the level of self-incompatibility, number of flowers per flower cushion, yield of dry cocoa per year, and percentage of healthy pods versus pods damaged by pathogens were included (García 1997). Within the selected trees in the Experimental Station of Palmira, SCP5 and SCP6 stood out as the most productive of that collection (Alvim 1956), although SC clones were never widely used in breeding programs.

The objective of the aforementioned National Cacao Campaign was to increase cocoa production in the country (García 1955). The campaign encouraged the establishment of some clonal cultivars. Resources were invested in the rooting of cacao cuttings and release of clonal material to farmers (Alvim 1956). However, this type of plant propagation proved to be expensive and difficult in some areas of Colombia (Alvim 1956). In the 1960s, two private Colombian chocolate companies, Luker and Nacional de Chocolates, promoted the establishment of new plantations and the introduction of cacao hybrids to Colombia (Sáenz et al. 1986).

As part of ICA’s research program on diseases having negative effects on cacao cultivation, an experimental plot was established in 1967 to assess hybrids obtained from...
different combinations, using both regional and introduced material. The natural incidence of _Ceratocystis_ sp. was monitored for ten years, and it was observed that during that period, the full-sib progeny issued from the cross PA46 × IMC67 was not affected by the disease (Ocampo and Correa 1978). The evaluation also indicated that progenies issued from SC13, SC24 and SC26 (Selection Colombia), ICS1, ICS 39, and ICS 95 as the female parent showed the highest incidence of the disease (Ocampo and Correa 1978).

The National Cacao Program of ICA, established in the experimental station of Palmira, set up trials to assess the yield of full-sib progenies in the 1980s. To our knowledge, all clones used as parents had been introduced from other countries (Peru, Trinidad and Ecuador). The criteria considered for the selection of the most promising full-sib progenies were: seed index (average weight of one dried seed), pod index (number of pods to produce one kilo of dry cocoa), size and number of beans per pod, number of ovules, level of self-compatibility and general combining ability (Ocampo et al. 1987). Hybrids were then delivered to farmers as planting material. The combinations selected and distributed according to Ocampo et al. (1987) are shown in Table 1.

The use of hybrids as planting material was promoted until the end of the 1990s; at that time, the recommendation changed to clonal cultivars (Perea et al. 2013). Other countries in America, such as Costa Rica, Ecuador and Brazil, also currently promote the use of clones as planting material (Lopes et al. 2011, Phillips-Mora et al. 2012, Sánchez-Mora et al. 2014).

As for pathogens limiting the cultivation of cacao in Colombia, according to Barros (1970), the disease caused by _Moniliophthora roreri_, commonly known as “frosty pod”, was one of the main constraints at that time in several cacao growing regions. In San Vicente del Caguán, Caquetá department, several plantations were abandoned due to the attack of this pathogen. Witches’ broom was registered in the country in 1928, in the region of Tumaco in the south-department, several plantations were abandoned due to the cacao growing regions. In San Vicente del Caguán, Caquetá pod”, was one of the main constraints at that time in several by et al. (Lopes also currently promote the use of clones as planting material. The combinations selected and dis-

<table>
<thead>
<tr>
<th>Crosses</th>
<th>Seeds</th>
<th>Pods</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA46 × IMC67</td>
<td>P7 × ICS6</td>
<td>P7 × ICS6</td>
</tr>
<tr>
<td>ICS1 × SCA6</td>
<td>ICS6 × P7</td>
<td>ICS6 × P7</td>
</tr>
<tr>
<td>TSH812 × IMC67</td>
<td>ICS39 × P7</td>
<td>ICS39 × P7</td>
</tr>
<tr>
<td>ICS6 × TSHA664</td>
<td>ICS1 × P7</td>
<td>ICS1 × P7</td>
</tr>
<tr>
<td>ICS6 × IMC67</td>
<td>IMC67 × EET62</td>
<td>IMC67 × EET62</td>
</tr>
<tr>
<td>ICS60 × SCA12</td>
<td>EET62 × IMC67</td>
<td>EET62 × IMC67</td>
</tr>
<tr>
<td>IMC67 × ICS6</td>
<td>ICS6 × IMC67</td>
<td>ICS6 × IMC67</td>
</tr>
<tr>
<td>TSH792 × IMC67</td>
<td>EET400 × IMC67</td>
<td>EET400 × IMC67</td>
</tr>
<tr>
<td>ICS40 × IMC67</td>
<td>P7 × ICS60</td>
<td>P7 × ICS60</td>
</tr>
<tr>
<td>TSH565 × IMC67</td>
<td>ICS1 × IMC67</td>
<td>ICS1 × IMC67</td>
</tr>
<tr>
<td>P7 × ICS39</td>
<td>P7 × IMC67</td>
<td>P7 × IMC67</td>
</tr>
<tr>
<td>EET400 × ICS1</td>
<td>EET400 × ICS39</td>
<td>EET400 × ICS39</td>
</tr>
<tr>
<td>EET400 × ICS6</td>
<td>EET400 × ICS60</td>
<td>EET400 × ICS60</td>
</tr>
<tr>
<td>ICS95 × IMC67</td>
<td>EET96 × IMC67</td>
<td>EET96 × IMC67</td>
</tr>
<tr>
<td>ICS8 × IMC67</td>
<td>P7 × ICS1</td>
<td>P7 × ICS1</td>
</tr>
</tbody>
</table>

As for pathogens limiting the cultivation of cacao in Colombia, according to Barros (1970), the disease caused by _Moniliophthora roreri_, commonly known as “frosty pod”, was one of the main constraints at that time in several cacao growing regions. In San Vicente del Caguán, Caquetá department, several plantations were abandoned due to the attack of this pathogen. Witches’ broom was registered in the country in 1928, in the region of Tumaco in the south-west of the country, and from 1962, it began to spread to other cocoa producing areas (Ocampo 1977). In a witches’ broom-resistance assessment study, conducted by Mejía and Rondón, and cited by Grisales (1983), the full-sib progenies issued from the following crosses were evaluated: SCA6 × ICS39, P7 × ICS6, EET 95 × SCA 6, SCA12 × ICS6, PA46 × IMC67, and SCA6 × IMC67. The study indicated that progenies issued from SCA series (Scavina) were not very productive and that they did not have desirable pod and seed indices. In addition, resistance to witches’ broom was not durable (Grisales 1983).

**Mass selection in local plantations**

Mass selection in local plantations has been applied in several cocoa-producing countries, such as Colombia, Brazil, Ecuador, Trinidad, Ivory Coast, Ghana, Nigeria and Cameroon (Ahnert and Eskes 2018).

In Colombia, in the 1990s, Corpoica (now Agrosavia), Fedecacao, and local private companies initiated a new breeding strategy based on mass selection of promising trees in cacao farms. Between 1996 and 1999, Corpoica and the National Program of Agricultural Technology Transfer (Pronatta), carried out a project that aimed to select trees from local farms in Santander department (Arguello et al. 1999). Individual plants were characterized _in situ_ and selected based on traits, such as plant architecture, number of pods per tree, pod index, seed index, resistance to pests and diseases, kilograms of cocoa per tree per year and, distribution of floral cushions on the trunk and primary branches. As a result, 50 outstanding trees were identified for their agronomic traits and coded SCC (Selection Colombia Corpoica). In this project, the producers were also trained in the propagation of plants using grafting techniques (Arguello et al. 1999). A total of 19 SCC clones are being currently evaluated, by Agrosavia, in replicated blocks in experimental centers having contrasting environmental characteristics.

In 2004, Fedecacao and Corpoica signed a technical–scientific cooperation agreement financed by the Ministry of Agriculture and Rural Development (MADR) and the Inter-American Institute for Cooperation on Agriculture (IICA), aiming at working in partnership in a five-year cacao breeding program to increase productivity. This collaborative project was called “Cacao de Colombia” (Colombian cacao); one of its objectives was the selection of trees with outstanding agronomic characteristics in local plantations in the main cacao-growing regions of the country (Aranzazu et al. 2009). During this project, Criollo cacao trees were collected from the Serranía del Perijá and the Sierra Nevada de Santa Marta, in 2006 and 2007 (Aranzazu et al. 2009). These materials were collected and characterized for the following traits: number of beans per pod, weight, length and diameter of the pod and bean index.

As a result of tree selection programs in local plantations, four clonal cultivars were delivered by Corpoica, now Agrosavia, to farmers as planting material in 2014, TCS01 and TCS06, and in 2017, TCS13 and TCS19, showing an average yield between 1.8 kg/tree/year (TCS19) and 3.3 kg/tree/year (TCS01) (Agudelo et al. 2017, Palencia et al.
2014). TCS01 is outstanding mainly due to its very high seed index (3 g per dry bean). These clones are currently used as parents in the Agrosavia breeding program.

Since the year 2000, Fedecacao initiated a tree selection program in local plantations in Santander, Arauca and Antioquia departments. Trees were evaluated and selected for resistance to diseases, productivity and physical and organoleptic qualities. As a result of the selection of trees in local plantations, Fedecacao delivered eight clonal cultivars to farmers: FLE2, FLE3, FSV41, FEC2, FTA2, FSA11, FSA12 and FEAR5 (Perea et al. 2013). The yield of these clones ranges between 1.4 kg/ha/year (FEC 2) and 1.8 kg/ha/year (FLE 3), while their seed and pod indices range between 1.3 (FEC2) and 2.0 (FSV41), and from 13 (FLE2) to 21 (FSA 12) (Perea et al. 2013).

Another mass selection effort was conducted in the municipality of Tumaco in Nariño department. Cacao research programs in this region strongly encouraged selection of regional genotypes to preserve the quality of local material. Studies of morpho-agronomic variability and genetic diversity using molecular markers were conducted in local plantations, keeping outstanding material, most of the time, in situ or in local clonal gardens (Ballesteros et al. 2015, Ruiz 2014). A report from 1983, prepared by the private chocolate companies Luker and Compañía Nacional de Chocolates, highlights the presence of local Criollo cacao trees with outstanding characteristics in the region of Tumaco, Nariño department (Gutierrez and Moreno 1983). Fig. 2 shows the approximate locations where mass selection of cultivated trees in local plantations has been carried out.

**Current breeding strategy**

A long-term strategy was adopted in 2012 to create improved populations of parents to be used as sources of favorable alleles in a Recurrent Selection (RS) program. The first step of this program aimed mainly at creating parents for a subsequent cycle, as well as clonal cultivars, accumulating alleles favorable in terms of yield components. It consisted in the creation of six full-sib progenies issued from crosses between productive parental clones. These progenies have been assessed in trial plots set up in the Palmira and La Suiza Research Centers. In 2015, the 19 trees with the highest early yield were cloned and are currently under assessment in two clone trial plots set up in those research centers.

Other cocoa-producing countries, such as Ivory Coast, Brazil and Trinidad and Tobago, also implemented recurrent selection as a breeding strategy long before Colombia (Clement et al. 1993, Lopes et al. 2011, Maharaj et al. 2011). In the case of Brazil, a recurrent selection program was implemented in the early 1990s, mainly as a response to the arrival of witches’ broom in the country (Lopes et al. 2011).

The RS program is continuing with a broader level of diversity and aiming at accumulating favorable alleles for yield and resistance to the four diseases causing most damage in Colombia: 1) moniliasis, caused by Moniliophthora roreri, accounts for 40–70% of fruit loss, depending on the area and agronomic practices (Arguello et al. 2008); 2) witches’ broom, caused by Moniliophthora perniciosa, affects both vegetative development and productivity (Evans et al. 2013); 3) black pod disease, caused by Phytophthora sp., causes significant fruit losses in some cocoa production areas in Colombia (Ramirez 2016), including the death of young cacao trees when trunks are attacked (Guest 2007); and 4) mal del machete, caused by Ceratoctysis sp., results in the death of trees a few weeks after the appearance of symptoms (Engelbrecht et al. 2007).

Phytophthora can be controlled with fungicides, resulting in both economic and environmental costs (Guest 2007). The damage associated with the other three diseases can be mitigated through proper agronomic practices (Jaimes and Aranzazu 2010); however, these are not systematically adopted by farmers and are not effective against the most aggressive pathogenic strains. In such conditions, the most effective way to control these diseases appears to be the creation of resistant varieties, with positive results obtained for witches’ broom in Brazil (Pimenta et al. 2018) and for moniliasis in Costa Rica (Phillips-Mora et al. 2012). In Colombia, further efforts need to be committed to improve the level of resistance of the cacao varieties released to farmers, as some of the highest yielding types appear to be susceptible to at least one major disease affecting the country.

A total of 160 full-sib progenies are currently being created, resulting from crosses between 60 parental clones selected for yield and resistance to the four diseases cited above. These 160 progenies will be shared by four research centers. The 40 progenies received by each research center will result from crosses using an incomplete North Carolina II mating design that allows the estimation of genetic parameters (general and specific combining ability, narrow sense heritability) associated with yield and disease resistance (Tahi et al. 2019). Combining ability of these variables has also been used for parent selection in Brazil (Dias and Kageyama 1995). The progenies will be assessed for their level of resistance to the diseases found in each locality, using early inoculation methods on seedlings in nursery (Nyassé et al. 1995, Silva et al. 2007), and later, assessing the incidence of the diseases in the field. Some crosses performed to create full-sib progenies will involve one or two parents with a high level of yield, others will involve parents resistant to two different diseases and, other crosses will involve parents resistant to the same disease. A combined progeny–tree selection step will allow selection of the most promising trees from among the most promising progenies to be cloned for future purposes. Fig. 3 shows the current breeding strategy:

- Assessment of trees as clones for future selection of locally adapted varieties: trees with high yield associated with moderate to high level of resistance to disease. These promising trees are expected to be found mostly within the progenies issued from crosses involving at
Molecular tools to support conventional cacao breeding program in Colombia

Molecular studies may be useful in accelerating breeding programs, especially in perennial plants like *T. cacao* that have an extended juvenile phase. Since the first genomes of *T. cacao* were published in 2011 and 2013, representatives of two contrasting genetic groups, Criollo (Argout et al. 2011) and Amelonado (Motamayor et al. 2013), genomics has provided breeders a new set of tools to facilitate the direct study of the genotype and its relationship with the phenotype.

Colombia is defined as having one of the highest expected levels of diversity for this species (Thomas et al. 2012). In this context, the study of Colombian genetic resources appears very interesting, especially in terms of providing solutions within the genetic breeding program to overcome the principal limitations affecting the crop in Colombia. Recently, the cacao genebank, managed by Agrosavia, was analyzed using 96 Single Nucleotide Polymorphism (SNP) markers (Osorio-Guarin et al. 2017). The study revealed a wide genetic diversity and suggested the existence of two new populations, mostly represented by Colombian accessions and, different from the ten genetic groups previously identified using molecular markers by Motamayor et al. (2008), that proposed the classification of the currently known *T. cacao* germplasm. Based on this study and aiming to support the current breeding program at Agrosavia, *T. cacao* accessions were selected and are currently being evaluated for disease resistance, productivity, cadmium uptake, sexual compatibility and cocoa quality traits.

*Genome Wide Association Studies (GWAS)*

In the Agrosavia cacao breeding program, phenotypic data collection for disease incidence and productivity are currently under process from the germplasm collection and recombinant progenies. A first GWAS was conducted on 200 cacao accessions in order to associate disease resistance, compatibility and productivity traits with SNP markers obtained by Genotyping by Sequencing. The first results of the GWAS were promising, showing positive association between SNP markers, localized in *T. cacao* chromosomes 5 and 1, and disease resistance to witches’ broom and frosty pod rot (Osorio-Guarin et al. 2018). Further analyses are currently being undertaken to validate these associations and to identify and characterize genes involved in associated regions.

*Transcriptomic studies*

Another molecular approach that complements GWAS for the detection of genes associated with agronomical traits, consists of the study of the differential gene expression of the plant under different conditions. Transcriptomic studies using RNAseq Next Generation Sequencing technologies allow detection of a large spectrum of genes expressed at a time point. By comparing different time points and different conditions, it is possible to detect genes expressed or repressed and link this information to the observed phenotype.

With the objective of developing molecular markers that are located nearby or inside genes involved in specific phenotypes, transcriptomic studies are currently underway for several traits, including heavy metals uptake (cadmium) and resistance to black pod (*Phytophthora palmivora*) and frosty pod rot (*Moniliophthora rorei*). A study was conducted in 2017 to determine the differential expression of genes of salicylic acid, jasmonic acid and ethylene metabolic
pathways (Delgadillo 2017), previously described as being involved in disease resistance (Rangel et al. 2010) during the infection of *T. cacao* by *P. palmivora*. Results indicated that genes related to the salicylic acid pathway show greater activation in the susceptible plant compared to the tolerant plant. In contrast, genes from the jasmonic acid and ethylene pathways were activated only in the tolerant plant during the first 48 hours of infection (Delgadillo 2017).

**Genomic selection**

Genomic selection is a marker-assisted selection method published in 2001 (Meuwissen et al. 2001) that aims to predict the genetic breeding value of selected candidate individuals within a breeding population. Genomic selection is based on genome-wide dense markers, usually SNP molecular markers, in which all markers are used, differing from previous methods of marker-assisted selection in which only markers associated with a significant effect are used.

The use of genotype-based selection, instead of phenotype-based selection, has the potential to overcome lengthy cocoa breeding cycles. In addition, genomic selection has the potential to increase the accuracy of selection of juvenile individuals in cacao breeding. Ribeyre et al. (2017) showed that genomic breeding values (GEBVs), computed by Genomic Selection, could be used to predict adult performance at the seeding stage for cacao bean weight and resistance to *Phytophthora megakarya*, increasing significantly the number of seedlings that can be accurately phenotypically evaluated.

In the Recurrent Selection program initiated by Agrosavia, full-sib progenies created are expected to cover a broad genetic diversity, based on the results of a molecular study conducted on the clones used currently as parents (Osorio-Guarín et al. 2017). Phenotypic and genotypic assessments will provide a good framework to carry out genomic prediction. In a first stage, genome wide association studies (GWAS) will be performed to characterize the genetic effects on the agronomic traits assessed. Depending on GWAS results, different Genomic Selection models will be evaluated, based on their predictive ability. The prediction models will be used to assist breeders in selecting cacao seedlings to be evaluated in the field and to select superior clones for the next breeding cycle.

**Conclusions**

One of the most important principles of any breeding program is the use of wide genetic diversity with the aim of increasing the probability of obtaining materials resistant to diseases and highly productive. Although Colombia is part of the center of the origin of the species, cacao breeding programs in this country have not made good use of all the genetic diversity available in germplasm collections, but instead, have mainly promoted the use of introduced accessions from Ecuador, Peru, Brazil, Costa Rica and Trinidad.

From the 1990s, Colombia initiated mass selection in local plantations, looking for cultivated trees with increased disease resistance and productivity. As a result, clonal cultivars were released as planting material since 2014. As part of a new breeding program adopted in Agrosavia, a Recurrent Selection strategy is now being implemented aimed at accumulating favorable alleles for yield and resistance to the most limiting diseases in Colombia, using, in addition, a large sample of the genetic diversity existing in local genebanks. To speed up the breeding program, phenotypic data from the trees assessed in the first cycle will be combined with molecular data obtained from the same trees, with the objective of developing strategies of marker-assisted and genomic selection.

**Author Contribution Statement**

CR-M conceived the manuscript and its components. All authors researched the literature. CR-M, XA, OS and RY conceived and designed the Figures. CR-M, XA, OS and RY wrote the manuscript. All authors read and approved the final document.

**Acknowledgments**

The authors would like to thank the Colombian government for financing the cacao breeding program that is currently developed by the Corporación Colombiana de Investigación Agropecuaria, Agrosavia. We extend our thanks to the National Germplasm Banks (SBGNAA) for the list of accessions sent to the La Suiza Research Center. We especially thank all the researchers and partner institutions who throughout history have contributed to the research on cacao breeding in the country.

**Literature Cited**


Cacao breeding in Colombia, past, present and future


Motamayor, J.C., K. Mockaitis, J. Schmutz, N. Haiminen, D. 9

Breeding Science Preview