UNIVERSITY OF MODENA AND REGGIO EMILIA

Ph.D. School in Agri-food Sciences, Technologies and Biotechnologies

XXXV cycle

Optimising the genetic resources of a *Capsicum* spp. (peppers) proprietary collection through effective pre-breeding approaches

Candidate: Dr. Marco Martelli

Tutor: Prof. Enrico Francia

Co-tutor: Dr. Justyna Anna Milc

Coordinator of the Ph.D. course: Prof. Alessandro Ulrici

...Des chercheurs qui cherchent, on en trouve;
des chercheurs qui trouvent, on en cherche.

Charles de Gaulle, 25 Mars 1959

TABLE OF CONTENTS

PREFACE

This PhD research falls within the frame of modern multidisciplinary plant breeding field embracing molecular tools and methodologies with innovative approaches for improvement of *Capsicum* species.

Furthermore, this PhD project is a prominent example of how the collaboration between public and private research can converge on common objectives by jointly sharing knowledge for the increase of know-how with concrete implications at the community level.

Hopefully, our industrial PhD work can serve and represent an inspiration and a guidance for orienting modern breeders.

As a significant leap forward the economics of the breeding program in pepper for the company Esasem S.p.A., the results obtained are innovation elements that confer a relevant impact on the efficacy and efficiency of the plant selection methods for interested traits.

Overall, this doctorate embodies a scientific progress carried out for the genetic resources optimization through effective pre-breeding approaches related to pepper (*Capsicum*) species.

ABSTRACT

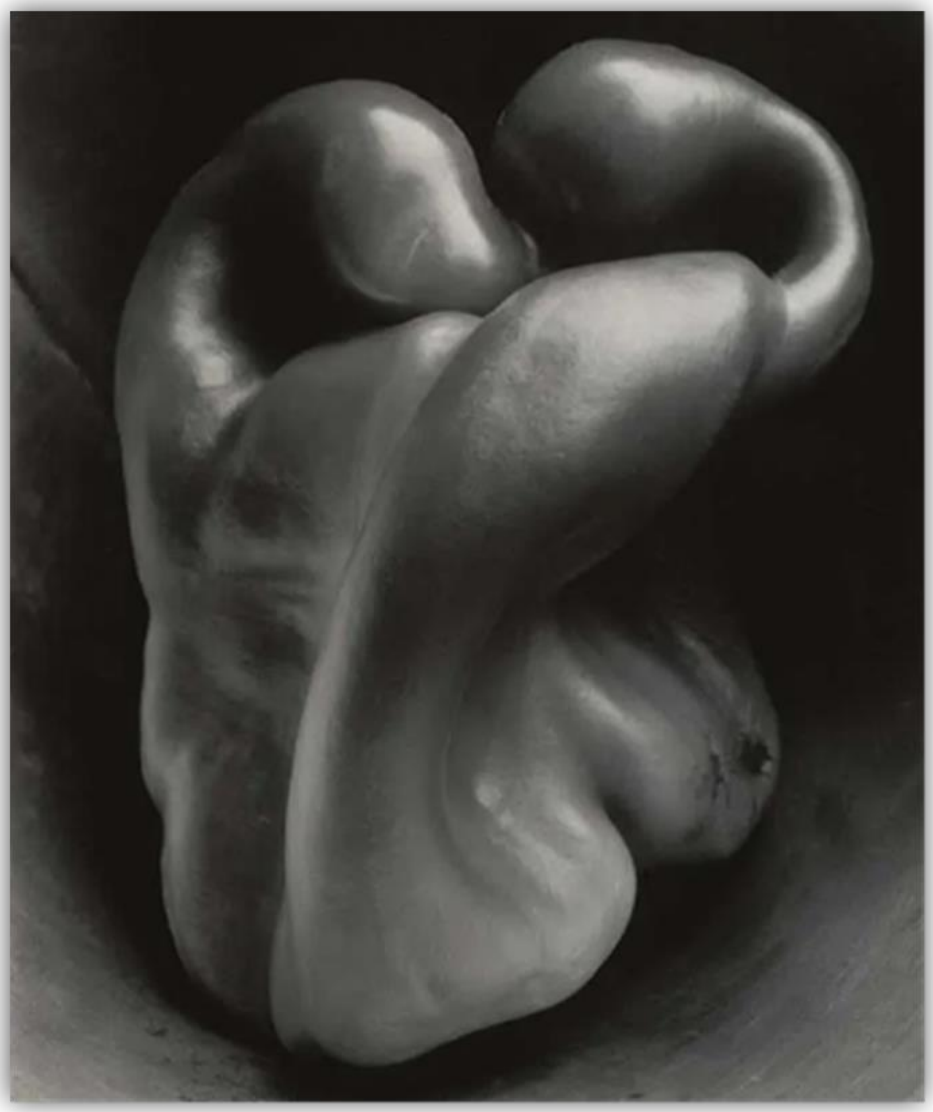
For over a century, plant breeding played a key role for the improvement of crops productivity. Evolution and success of this science has relied, mostly on the last decades, on an integrated multidisciplinary approach targeted on both identification and combination of favorable genetic traits. Plant breeder must constantly cope with ever-changing aspects such as evolution of agricultural practices, growing needs of specific agronomic characteristics, new environments adaptability and finally consumer preferences as market requirements. Within this frame, varietal release process contributed to the development of modern societies. Novel genotypes with desirable characteristics and obtained through introduction of associated genes for biotic and abiotic stress resistance from unadapted, landraces, wild relatives and close related species, still represent challenging tasks. However, choice of breeder strategy in terms of selection methodologies adopted may determine the success or failure of a process that lasted years.

Continuous innovation in plant selection approaches is necessary, and therefore, in our industrial PhD thesis we explored this aspect using pre-breeding approaches for the optimization of genetic improvement activities.

Facing the superior genotypes constitution, the research carried out for a throughout characterization of a proprietary collection of *Capsicum* spp. (peppers) germplasm brought to:

- selection and application of genotypic assays for associated genes/loci related to pathogens resistances (viruses, nematodes, fungi, bacteria) useful for marker-assisted breeding;
- exploitation of genetic variability in the germplasm collection by re-sequencing selected genotypes as promising progenitors for the future breeding programs. The genomes of identified elite material was characterized and polymorphisms detection revealed the core collection genetic distance, allowing at the same time the obtaining of a company genotypic fingerprinting profile as well the identification of private SNPs for future investigation of new potential candidate genes.

Predictive information derived from the research activity has been applied directly towards gene pyramiding flow improving the efficiency and efficacy of selection process and setting, as long-term objective, to realize an efficient breeding program following a novel scheme for the segregating populations selection based on a multi-environment phenotyping (Inter-placing method).



Edward Henry Weston (1886-1958)

Chapter 1

General introduction

1.1 Pepper (*Capsicum* spp.)

1.1.1 Origin and distribution

Peppers (*Capsicum* spp.) are one of the most important worldwide vegetable crop. Originated in the area of American tropics where was domesticated as far back 6000 B.C., the name *Capsicum* derive from the Greek word "*Kapsimo*", which means "to bite" (Sushil *et al*., 2016).

With a growing economic importance due to the widely consumed (fresh, dried or processed products), this species is characterized by an amazing variability in plant and fruit characteristics with innumerable uses, both food and non-food (Tripodi, P. and Kumar, S., 2019). According to annual world production, in 2020 (Fig. 1) peppers are among the top 10 most important vegetables (Faostat, 2020).

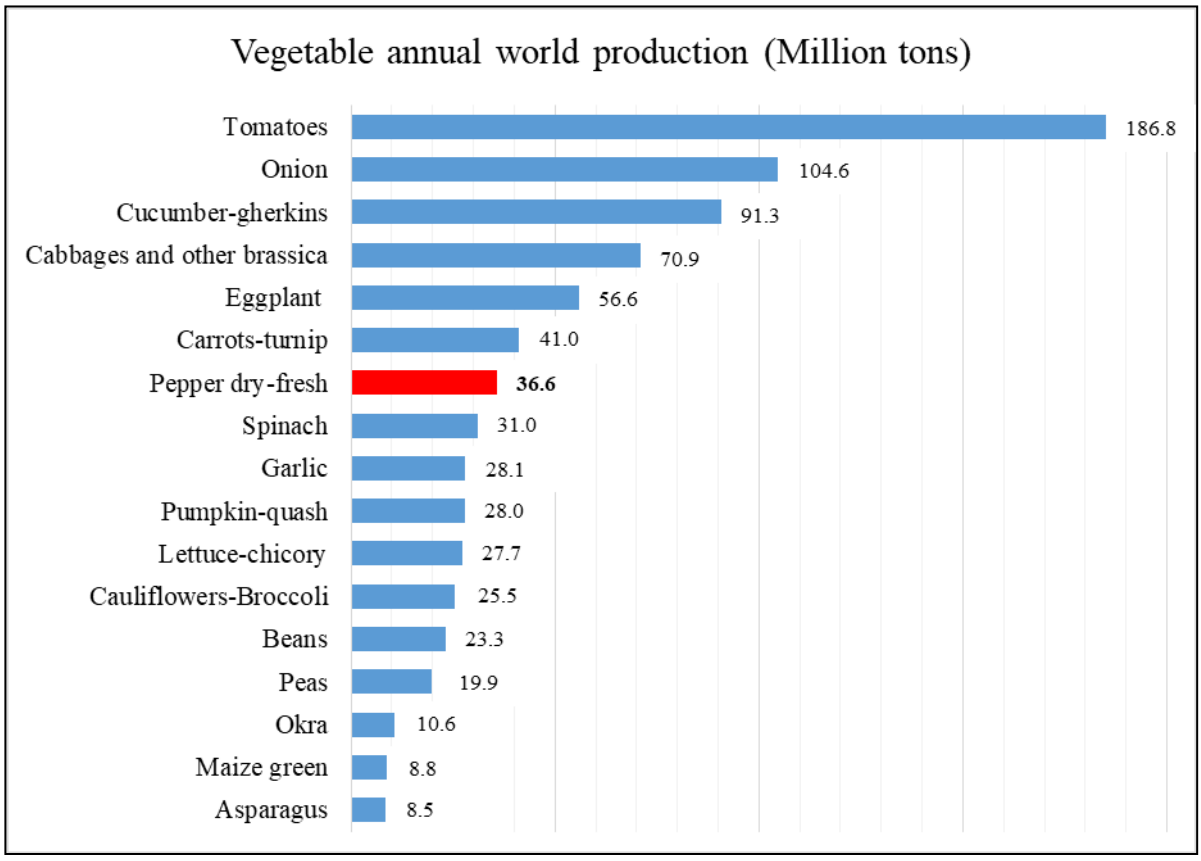


Figure 1. Vegetable annual world production (source Faostat, 2020)

In total 138 countries cultivate pepper with a world surface area of about 3.7 million of hectares and a global production estimate from 2 to about 4.5 million tons of dry types and 10 to 38 million tons as fresh (Fig. 2-3).

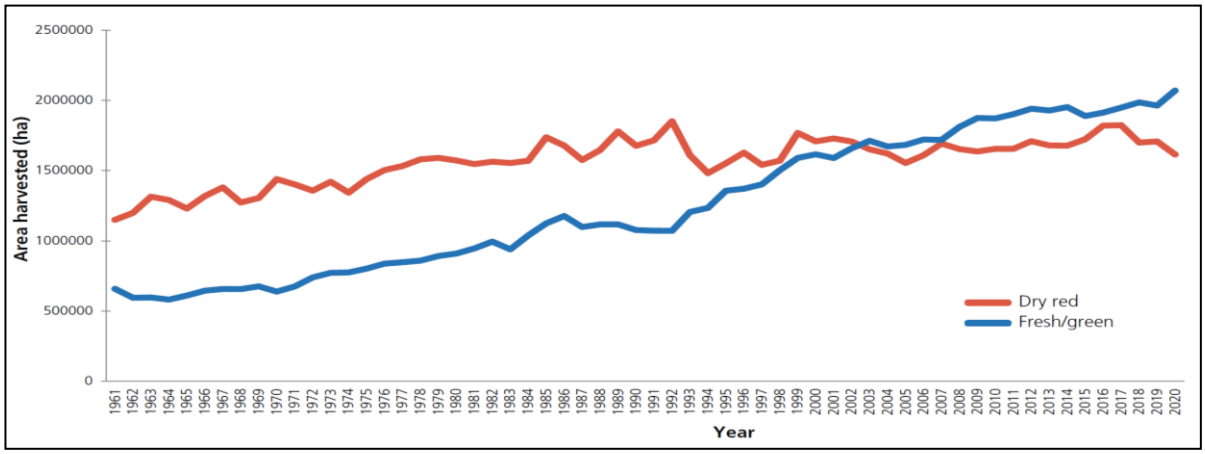


Figure 2. World area harvested (ha) of pepper, from 1961 to 2020 (source Faostat 2020)

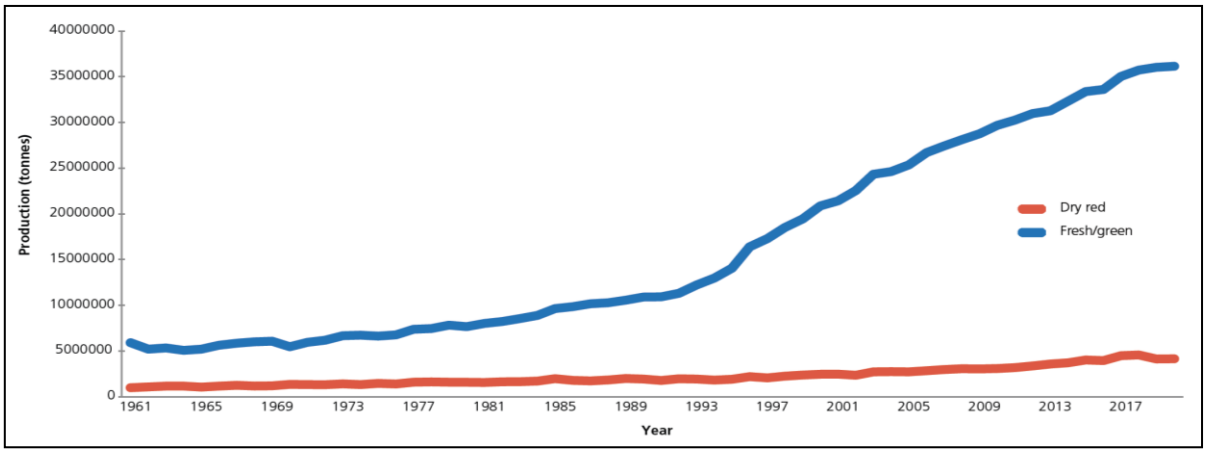


Figure 3. World production (tons) of pepper, from 1961 to 2020 (source Faostat 2020)

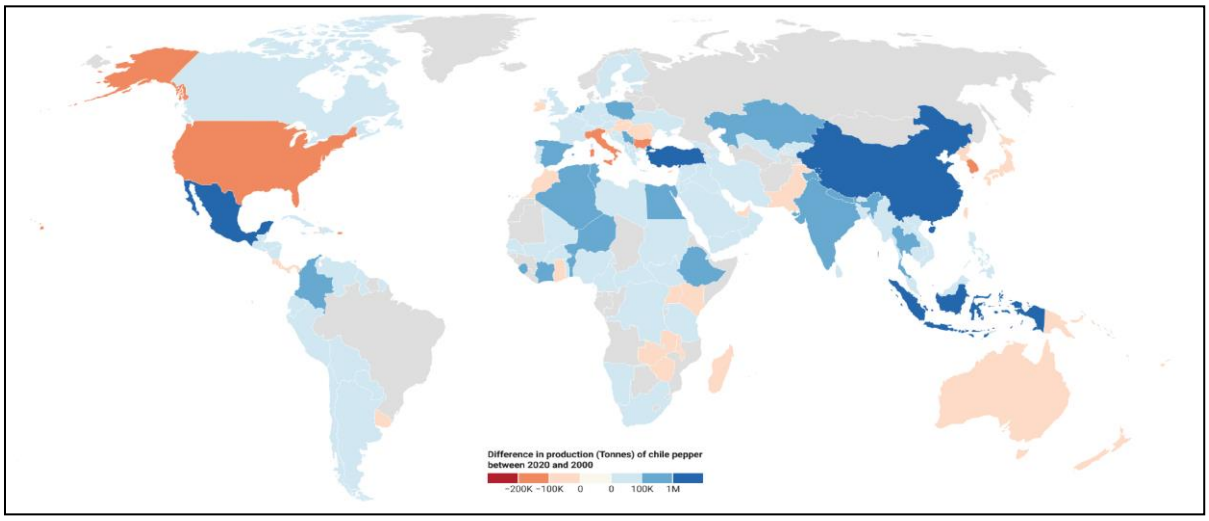


Figure 4. Difference in production (tons) of pepper (summed value of red dry and fresh green) between 2020 and 2000. The darker blue scale color show the greater the increase in production that has occurred during the last 20 years; the darker the red, the greater the decrease (source Faostat 2020).

As shown in Figures 5 and 6, China represents the main country for fresh cultivation with 798,877 hectares while India is the first for dry pepper with 780,000 hectares (Faostat, 2019).

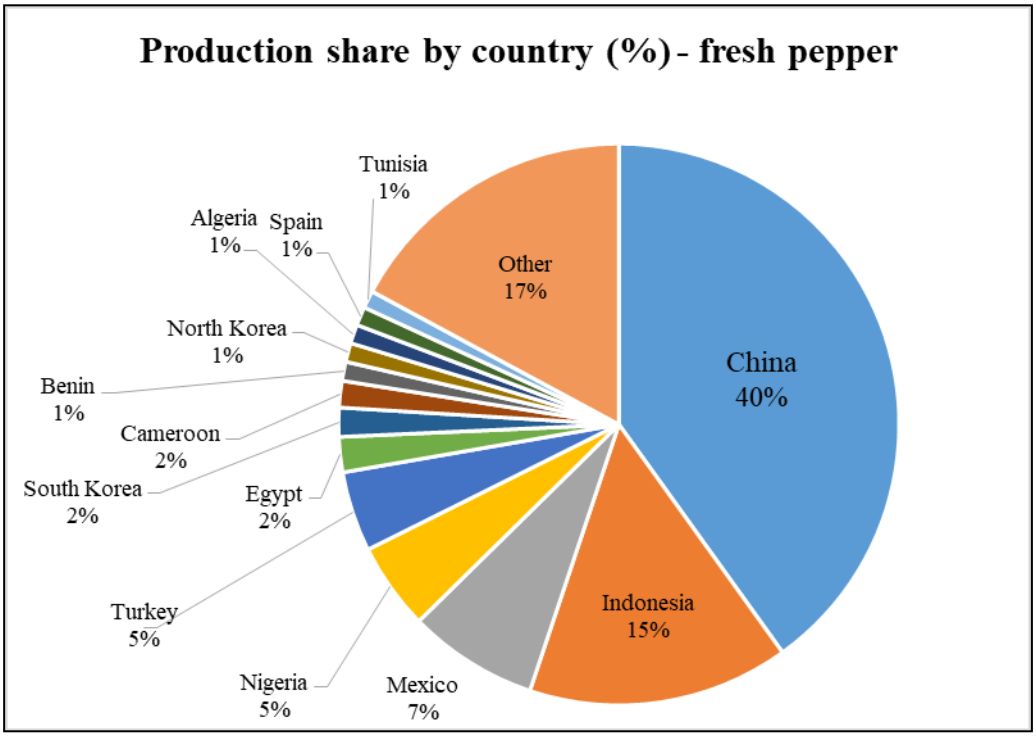


Figure 5. Production share by country (%) of fresh pepper (source Faostat, 2019)

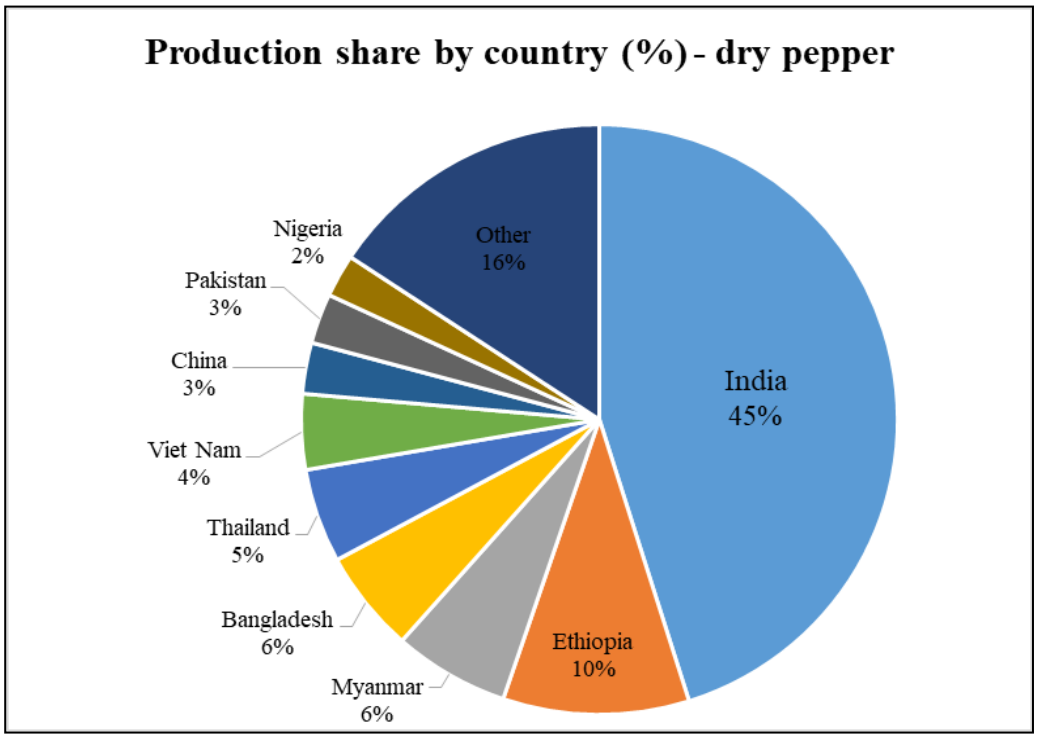


Figure 6. Production share by country (%) of dry pepper (source Faostat, 2019)

Italy confirms its important role regarding pepper market both in national and international trading of pepper. In 2021, indeed, Italian pepper production has been estimated in 180,034.9 tons with a cropping area of 8,018 hectares for production of pepper in full air, and 71,128.7 tons for 1,652.42 hectares related the greenhouse production (ISTAT, 2021).

1.1.2 Genus *Capsicum*: taxonomy and morphology

Cultivated pepper belongs to the family of Solanaceae that includes 98 genera and 2,716 species of important vegetables such as potato, tomato and eggplant.

Originated from the American continent, the genus *Capsicum* have its genetic center in the central region of Bolivia. Including many species, genus is usually classified both in cultivated plants and wild species particularly widespread in Bolivia, Brazil, Perù, Venezuela, Colombia, northern Argentina and even in all Central American countries (Csilléry, G., 2006, McLeod, M. J., 1982).

Genus includes at least 32 species (Tab. 1), five of which are domesticated (Tab. 2): *C. annuum* L., *C. baccatum* L., *C. chinense* Jacq., *C. frutescens* L., and *C. pubescens* (Qin *et al.*, 2014).

C. annum	*C. ciliatum*	*C. leptodum*
Var. aviculare	*C. coccineum*	*C. minutuflorum*
Var. annum	*C. cornutum*	*C. mirabile*
C. baccatum	*C. dimorphum*	*C. parvifolium*
Var. baccatum	*C. dusenii*	*C. praetermissum*
Var. pendulum	*C. eximium*	*C. pubecens*
C. buforum	*Var. tomentosum*	*C. schottianum*
C. campylopodium	*C. galapagoense*	*C. scolnikianum*
C. cardenasii	*C. geminifolium*	*Var. flexuosum*
C. chacoense	*C. hookerianum*	*C. tovarii*
C. chinense	*C. lanceolatum*	*C. villosum*

Table 1. *Capsicum* species (source IBPGR, 1983)

Domesticated species	Common Name
C. annum	Cayenne, Jalapeno, Paprika, Bell pepper, Chilpetin
C. baccatum	Brown and Aji pepper
C. chinense	Scotch Bonnet, Datil, Trinidad Scorpion, Habanero
C. frutescens	Tabasco, Melagueta, Bird's eye
C. pubescens	Rocoto, Tree pepper, Quechuan

Table 2. Domesticated *Capsicum* species (source Qin *et al*., 214)

Some wild species are still collected and used by indigenous populations. All wild species share similar characters: small, bright red fruits, erect, with a tender pedicel that allows them to be easily detached from the calyx when ripe. This characteristic facilitates the action of birds which, attracted by the color of the fruits and not sensitive to capsaicinoids, guarantees their dissemination. The different species are divided into three groups based on morphological and cytological characteristics. Each group includes both domesticated species and their wild ancestors. Species belonging to the same group are intercrossable, vice versa hybridization between species of different groups is hampered by various incompatibility mechanisms. The groups represent a sort of *gene pool* of the *Capsicum* genetic variability origin (Bosland, P.W., and Votava, E.J., 2000).

- ***Capsicum annuum* L.**

C. annuum L. is the most extensively cultivated species for the greatest economic importance, compared with the other four cultivated pepper species (Pickersgill, B., 1997). The *C. annuum* group includes three domesticated species: *C. chinense*, *C. frutescens* and *C. annuum*, which share a common ancestral gene pool. They are the most important species since commercial point of view. Each was independently domesticated: *C. chinense* in the Amazon or in Perù, *C. frutescens* in central-southern America and *C. annuum* in Mexico (Smith *et al.*, 1987; Bosland, P.W., 1992). The different varieties of *C. annuum* are usually classified according to their characteristics of fruits: spiciness, color, shape, flavor, size and use (Smith *et al.*, 1987). The most probable progenitor of cultivated *C. annuum* is the wild "chiltepin" (*C. annuum* variety *aviculare*), characterized by a wide diffusion, from southern America to Arizona. It is believed to have been domesticated in Mexico and Central America where, before the Spanish colonization, the Aztecs had already developed a dozen of variety. Undoubtedly, "chiltepin" are the forerunners of a large number of varieties today grown in Mexico (Fig. 7).

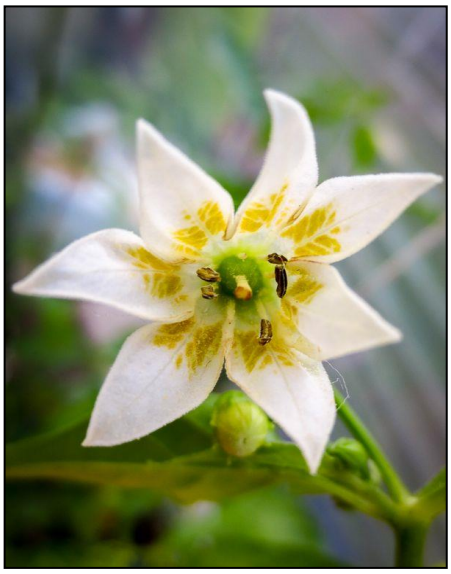
Figure 7. Wild type “chiltepin” (source: Esasem)

The *C. annuum* includes many of the Mexican “chile peppers”, the spicy peppers mainly grown in Africa and Asia while different sweet pepper cultivar grown in the temperate regions of Europe and North America.

- ***Capsicum baccatum***

C. baccatum, called "Ajì", is a species widespread from the South, in Brazil, up to in the West, on the shores of the Pacific Ocean. It is believed that its domestication center includes Bolivia, Ecuador, Perù, and Chile. In South America it is the largest species commonly cultivated. *C. baccatum* has cream colored flowers with yellow, brown and dark green spots on the corolla (Fig. 8).

Two botanical varieties are distinguished: *C. baccatum* var. *baccatum*, which is the form wild, and C. *baccatum pendulum* variety, which includes the domesticated forms. Fruits vary according to shape, color, spiciness and express peculiar properties organoleptic (Fig. 9). The *C. baccatum* group also includes the species *C. praetermissium* and *C. tovarii* (Tong, N., 1998).



Figures 8-9. *C. baccatum* flowers and fruits (source: Esasem)

- ***Capsicum chinense***

The *C. chinense* species, like all the others, is native to America. It was the physicist Dutch Kikolaus von Jacquinomist who gave the name "chinense" to the species in 1776, but it is not yet clear why he thought that China was his center of origin (Jacquin N.J.,1776).

It is the most widespread species in Brazil (Cheng S.S., 1989) and in tropical regions, above all the Caribbean ones.

Like as for *C. baccatum*, the fruits show great variability in shape and size degree of spiciness, "habanero" (Fig. 10) stands out as the hottest of peppers in the world.

Figure 10. *C. chinense* fruits (source: Esasem)

- ***Capsicum frutescens***

The species *C. frutescens* (synonyms: *C. minimum* and *C. fastigiatium*) includes a minor number of varieties compared to *C. chinense* and *C. annuum*.

The famous cultivar in Brazil is "Malagueta" which, however, is not related to Aframomum melegueta, "melegueta" or "pepper of Guinea ". *C. frutescens* (Fig. 11) includes moreover the well-known “Tabasco” variety, used for the sauce of the same name. Some varieties are grown in Africa, India and Asia where they are called "bird peppers" and grow as shrubs and perennials.

Figure 11. *C. frutescens* fruits (source: Esasem)

- ***Capsicum pubescens***

The *C. pubescens* group consists of relatively unknown species. *C. pubescens* was initially described by Ruiz and Pavon (1790) as a species native to the highlands of Bolivia and, that according to Heiser (1976), was domesticated around 6000 BC in the Andes, where it is often referred as “locato” or “rocoto”.

It grows in the area between Mexico and Perù, in the Andes of South America and in the highlands of Central America in small plots family members. In the rest of the world it is not very widespread. The C. *pubescens* plant has purple flowers (Fig. 12), numerous pubescent leaves and black seeds, which makes it distinguishable from other species (Fig. 13). It is a herbaceous plant, highly branched and it can grow up to 12m tall living over 10 years in the conditions climatic of tropical America.

Fruits vary in shape (from elongated to spherical) and for color (red, orange or yellow) but they do not show the same range of variability as *C. annuum.* The fruits are characterized by a dense pulp, difficult to dehydrate and store.

Figures 12-13. *C. pubescens* flowers and fruits (source: Esasem)

1.1.3 Nutritional value

Pepper is largely appreciated for its nutritional value (Parisi *et al*, 2020). Pepper fruits are rich of antioxidants and nutrients for human diet such as carotenoids (pro-vitamin A), vitamin A (which destroy free radicals), vitamins C, vitamin E, flavonoids and capsaicinoids (Maga, J.A., 1975, Simonne *et al.*, 1997).

Carotenoids and flavonoids are especially responsible for the colors of the fruits while capsaicinoids cause the taste characteristic pungency that marks the plant (Antonio *et al*., 2018).

Plants are also used in traditional medicine as folk remedies and as resource for pharmacological industry for diarrhea, arthritis, muscle cramps, asthma, dropsy and colic toothache due to the nutraceutical and anticancer effects (Ravishankar *et al*., 1988). One of the most important compounds of some *Capsicum* species are the capsaicinoids.

Causing the spicy flavor, pungency is due to the complex of capsaicin (unsaturated and saturated amides) and dihydrocapsaicin. Capsaicin can be used as pain relieving agent in chemo-radio therapy and neuropathic pain (Berger *et al*., 1996).

Moreover, different species of the genus *Capsicum* are used as ornamentals for the esthetically appreciated thanks to their characteristic traits.

In addition, recent investigations have shown that pepper seeds could be considered a valuable source of edible oil, fiber-rich flour and protein after processing.

According to quality analysis (nutritional, chemical, sensory and antioxidant characteristics), pepper seed oil is a high-quality and is suitable for the use in the food and non-food industries (pharmaceutical, chemical, cosmetic industries) (Cvetkovi *et al*., 2022).

1.1.4 Widespread diseases

Severe diseases, caused by several pathogens that are diffused worldwide, threaten the production of pepper. Approximately 87 pathogens and diseases of pepper are estimated (Kim *et al.*, 2019).

The biological range of these pathogens is broad, including viruses such as *Potyviruses* (e.g., *Potato virus Y, Tobacco etch virus*, *Pepper mottle virus*), *Tobamoviruses* (e.g., *Tobacco mosaic virus*, *Tomato mosaic virus*), *Cucumoviruses* (e.g., *Cucumber mosaic virus*), *Tospoviruses* (e.g., *Tomato spotted wilt orthotospovirus*), nematodes (*Meloidogyne* spp.), fungi (e.g., *Phytophthora capsici*, *Rhizoctonia solani*, *Verticillium dahliae*, *Colletotrichum scovillei* and *truncatum*, *Leveillula taurica*, *Fusarium* spp.), bacteria (e.g., *Xanthomonas* spp.) and insects (e.g., mites, aphids, lepidoptera and thrips).

To ensure healthy and profitable product, cropping methods and pesticides are applied, however, nowadays the use of resistant plants to protect pepper cultivation against biotic stresses represent the main strategy (Sarath *et al.*, 2011; Djian-Caporalino *et al.*, 2014).

The limited use of soil fumigants imposed in the recent year, addressed interest in resistance introduction in rootstock and cultivars of soilborne pathogens such as *Meloidogyne* spp. and *Phytophthora* spp. (Ros *et al.*, 2014).

- **Fungal diseases: powdery mildew**

Caused by an obligate fungal plant pathogen that belongs to the ascomycetes, the powdery mildew of pepper is a disease known also as *Leveillula taurica* (that is the asexual stage of *Oidiopsis taurica)* (Rajesh, R.W. and Madhukar, S. W., 2018).

This worldwide diffused disease can cause serious yield losses of up to about 80% of the crop in areas with dry warm or humid warm climates (Parisi *et al.*, 2020). Producing grayish-white patches on the ventral side of the leaves and light green-yellow lesions on the dorsal side (Fig. 14), the fungal infection causes a premature defoliation with a photosynthesis reduction and makes fruits not suitable for marketing.

As endophytic fungus *L. taurica* is able to hinder the efficacy of chemical treatments (Elad *et al.*, 2007).

Figure 14. *Leveillula taurica* on leaf (source Esasem)

- ***Phytophthora capsici* root rot and foliar blight**

With over 70 species in the world, *Phytophthora capsici* is one of the main pathogens of pepper (Fig. 15). The infection, that may have devastating effects on cultivations, usually occurs during periods characterized by strong rainfall and high humidity levels with temperature ranging from 15°C to 23°C (Quirin *et al.*, 2005).

Typically soil-borne, the oomycete can infect very broad range of host, including Solanaceae, Cucurbitaceae and the other plants (Kim *et al.*, 2019).

Depending on the site of infection, *P. capsici* is able to infect all parts of the plant causing several different disease syndromes (root rot, stem rot, fruit rot and foliar blight). Among of these, *Phytophthora capsici* root rot is the most harmful disease because usually it is responsible of plant death reducing drastically crop yield (Xu *et al.*, 2016).

In the past, soil sterilization with the methyl bromide, was the main method to control the disease; however, it was banned in 2005. Due to capacity of the pathogen to overwinter for long periods in the soil crop rotation has also resulted ineffective.

For these reasons, chemical control and cultural practice are unsafe and ineffective methods to control the *P. capsici* infection (Lamour *et al.*, 2003).

Figure 15. *Phytophthora capsici* root rot (source Esasem)

- **Bacterial diseases: bacterial spot of pepper**

Bacterial spot in pepper is caused by four species of the genus *Xanthomonas*: *X. vesicatoria*, *X. perforans*, *X. euvesicatoria* and *X. gardneri* (Parisi *et al.*, 2020).

Water-soaked, black spots and necrotic lesion on the fruit, flowers, stems and leaves are the typical damage caused by this disease that generally afflicts the cultivations of pepper in tropical and subtropical production areas in the world (Truong *et al.*, 2011).

Formation of circular water-soaked spots is followed by forming large yellow areas when the infection progresses (Fig. 16). These dry out becoming roughened and raised appearing cracked. Often, during the development of the disease it is possible to observe on the stem the formation of raised and elongated tumors with circular, small, slightly lesions on green fruits (Parisi *et al.*, 2020).

Lesions on the fruit cause the yield reduction making them unmarketable (100% of losses of the harvest). In addition the dropping of leaves exposes the fruits to sunscald reducing the productivity.

Copper-based pesticides have been used against bacterial spot disease for several decades as chemical methods available (Giannessi, L.P. and Reigner, N., 2005).

Figure 16. Bacterial spot on fruit (source Esasem)

- **Viral diseases - thrips transmitted viruses: *Orthotospoviruses***

Orthotospoviruses, also known as *Tospoviruses*, are pepper viral pathogens belonging to the Tospoviridae family. The viruses are composed of three molecules S, M and L, an own tripartite single-stranded RNA genome with negative or ambisense polarity (Molly *et al.*, 2000). They can infect about 1,090 plant species (industrial, vegetable and ornamental crops) belonging to 90 families (Moury *et al.*, 2000; Riley *et al.*, 2011). At least seven species of thrips, primarily *Frankliniella occidentalis*, are responsible of virus transmission. Among *Orthotospoviruses*, the most damaging for the pepper crop is *Tomato spotted wilt orthotospovirus* (TSWV) which infection results in stunting of plants and severe quality reduction (Fig. 17). The disease is usually first seen on fruit where it causes distorted growth (bumps), uneven ripening, ring spots, line patterns and necrotic patches (Nigam, D. and Garcia-Ruiz, H., 2020).

Figure 17. Ring spot on fruit caused by *Tomato spotted wilt orthotospovirus* (source Esasem)

- **Viral diseases - aphid transmitted viruses: *Potyviruses***

Potyviruses have a positive-sense single-stranded RNA genome (approx. 10 kb) and belong to the Potyviridae family. *Potyviruses* are worldwide widespread (mainly in warmer areas) and use aphids as transmission vector (Kang *et al.*, 2005; Caranta *et al.*, 1999; Kim *et al.*, 2015).

Main economically important crops, such as tomato, pepper and potato are especially affected by Potato virus Y (PVY) (Arnedo-Andrés *et al.*, 2002).

PVY of Mediterranean regions are been divided in three pathotypes: 0, 1 and 1,2 in base of the plant resistance genes recognizing the corresponding virus isolate antigen. Symptomatology in host is highly variable and is characterized by foliar necrosis, blistering and mosaic symptoms (Caranta *et al.*, 1999).

To control the yield losses related to *Potyvirus* infections, chemical control of vector organisms has been used with low efficiency against these viruses (Devran *et al.*, 2015).

- ***Cucumoviruses***

Cucumber mosaic virus (CMV) that belong to the genus *Cucumovirus* and *Bromoviridae* is one of the most virulent plant viruses diffused worldwide and is mainly present in the temperate and tropical zones.

The host range of the *Cucumber mosaic virus* is very broad and includes 1,200 plant species belonging to 100 plant families (also monocotyledons and dicotyledons, such as pepper). Virus can be transmitted through both parasitic plant dodder (*Cuscuta* spp.), through seeds and by about 75 species of aphids, primarily *Myzus persicae* and *Aphis gossypii* (Choi *et al.*, 2018).

CMV genome is very complex and an individual particle of virus contains three genomic RNAs: RNA 1 (3.3 kb), RNA 2 (3.0 kb), RNA 3 (2.2 kb) and a subgenomic RNA 4 (1.0 kb). This is the reason for the high diversity among CMV isolates (typically classified in subgroup I and subgroup II) (Eun *et al.*, 2016). Leaf distortion, fruit lesions (Fig. 18), dwarfing, chlorosis and systemic necrosis affected pepper (mainly subgroup I) with a yield reduction greatly in the early stages up to 80% (Mochizuki, T. and Ohki, S. T., 2012). Due to high complexity of CMV genome and broad variety of transmission vectors, disease control is very difficult (Avilla *et al.*, 1997).

Figure 18. CMV on fruit (source Esasem)

- **Viruses transmitted by contact: *Tobamoviruses***

The genus *Tobamovirus* is one of the most damaging virus group for the plants of the Solanaceae family.

Belongs to the Virgaviridae, the *Tobamovirus* species involved on the pepper infection are *Tobacco mosaic virus* (TMV), *Tomato mosaic virus* (ToMV), *Pepper mild mottle virus* (PMMoV), *Tobacco mild green mosaic virus* (TMGMV), *Bell pepper mottle virus* (BPeMV), *Paprika mild mottle virus* (PaMMV) and *Obuda pepper virus* (ObPV) (Parisi *et al.*, 2020).

Depending on the presence of different L resistance genes, these viruses are categorized in four pathotypes: P0, which includes TMV and ToMV, P1, formed by PaMMV, and P1.2/ P1.2.3, formed by PMMoV. ToMV and PMMoV are more diffused in Southeast Asia while TMV, ToMV, PMMoV and TMGMV are predominant in southern Europe (Kenyon *et al.*, 2014). Tobamoviruses are very stable viruses that present a single-positive-sense RNA genome of about 6.3-66 kb encoding for four viral proteins.

Infection causes misshapen and discolored fruit with necrotic patches while leaf suffer distortion and chlorotic mottle or mosaic (Fig. 19). These symptoms generally cause severe yield losses and quality reduction of harvest. Transmission of *Tobamovirus* is mechanically favored by either contact or seed-borne attached to their coat (Dombrovsky, A. and Smith, E., 2017).

Figure 19. Damage caused by *Tobamovirus* on fruit (source Esasem)

- **Root-knot nematodes**

Root-knot nematodes belong to the genus *Meloidogyne* including about 98 species. It is a soil-dwelling microscopic roundworm, parasite of several plants (vegetables, fruits, ornamentals and field crops) including some Solanaceae species (Seebolf, K., 2014).

Root-knot caused by these nematodes (Fig. 20) are known also as galls and affect the use of water and nutrients of plant host with an effect of reduction of shoot growth and photosynthetic capacity (Celik *et al.*, 2016).

Most aggressive species are *M. incognita, M. javanica, M. arenaria and M. hapla* (Wang *et al.*, 2018). The use of chemicals (nematicides) has been a common strategy to control nematodes for a long time but recently their use has been reduced in response to the negative impact on human health and environment (Pinar *et al.*, 2016).

Figure 20. Damage caused by root-knot nematodes (source Esasem)

1.1.5 Genetic and genomic resources of *Capsicum* spp.

Several institutes and organizations are engaged in maintaining and collecting genetic resources of *Capsicum* species for basic research.

AVRDC (Asian Vegetable Research and Development Center of Taiwan), with a total of 8,170 accessions belonging to wild and cultivated *Capsicum* species from around the world, has the largest collection of *Capsicum* germplasm.

United States Department of Agriculture (USDA) has 6,067 accessions while NBPGR (National Bureau of Plant Genetic Resources of New Delhi) has collections of 2,774 accessions.

Listed in the Tab. 3 and 4, are the main genetic-resource centers active for germplasm maintenance and bioinformatics databases of *Capsicum* species (Sushil *et al*., 2016).

GENEBANK	WEBSITE
Banco de Germoplasma de Hortalicas (BGH) - Brazil	https://www.thechileman.org/search.php
Centre for Genetic Resources (CGN) - Netherland	https://www.wur.nl/en/research-results/statutory-research-tasks/centre-for-genetic-resources-the-netherlands-1.html
The Chile Pepper Institute (New Mexican State University) - Mexico	https://cpi.nmsu.edu/
The German Research Centre for Biotechnology (CAP) - Germany	https://www.thechileman.org/search.php
United Stated Department of Agriculture Research Service (USDA) - United States of America	https://www.ars-grin.gov/npgs/acc/acc-queries.html
The Asian Vegetable Research and Development Center (ARVDC) - Taiwan	https://www.arvdc.org
National Bureau of Plant Genetic Resources - India	https://www.nbpgr.ernet.in/pgrportal
ECPGR Pepper Database, Aegean Agricultural Research Institute, AARI - Turkey	https://www.atae.gov.tr/eng/Default.aspx
The Chile variety database - Chile	https://www.g6csy.net/chile/database.html

Table 3. List of resources and germplasm databases of *Capsicum* (source Sushil *et al*., 2016)

ONLINE RESOURCES	HOLDER
Capsicum Transcriptome DB	Cinvestav Irapuato - Mexico
Kazusa Marker Database	Kazusa DNA Research Institute - Japan
PEPPER EST Database	Korea Research Institute of Bioscience and Biotechnology - Korea
Sol Genomics Network	Boyce Thompson Institute for Plant Research - United States of America
Plant Geneome Database of Japan	Kazusa DNA Research Institute - Japan
SolCyc Database	Boyce Thompson Institute for Plant Research - United States of America

Table 4. Databases list for bioinformatics support of *Capsicum* genetics (source Sushil *et al*., 2016)

Genetic maps

Large-scale genomic resources availability allowed an accelerated development of high-density genetic linkage maps in *Capsicum*.

Phenotypic information related segregating mapping population and the analysis of these genetic maps has facilitated the investigation of molecular marker linked to genes and QTLs associated with different agronomic characters.

Many pepper populations were used for the construction of pepper genetic maps, and are represented by recombinant inbred line (RIL), double haploid (DH), F2 or backcross mapping populations with anonymous markers (RFLP, AFLP, RAPD) and isozyme marker (Livingstone *et al*., 1999; Kang *et al*., 2001; Rao *et al*., 2003; Lee *et al.*, 2004; Ben-Chaim *et al*., 2006).

The first genetic map was constructed using an inter-specific cross of *C. annuum* and *C. chinense*, with wide genome coverage contained 85 RFLP markers (Tanksley *et al.*, 1988). Thereafter, Prince *et al.* (1993) improved the map using a cross of same *C. chinense* parent with different *C. annuum* using 192 markers.

Using the same population Livingstone *et al*. (1999) constructed a comparative map (pepper-tomato) using AFLP, RFLP and RAPD markers covering a total of 1,245.7 cM distance.

Inter-specific maps

The development of maps from inter-specific crosses offers the advantage of obtain segregant populations characterized by high polymorphism level. From the other side, inter-specific offspring are subjected often to low fertility level, distorted segregation and new chromosomal structure set-up arise during the species differentiation (Lantieri S. and Pickergill B., 1993). Mostly inter-specific maps are developed through the molecular analysis of *C. annum* x *C. chinense* population due to the fertile cross and high polymorphic progeny.

Integrated genetic maps

Single genetic maps, developed in *Capsicum,* are developed using different molecular marker types and this represent a limit on the extrapolation of linkage information from one population to others, for this reason, consensus maps, using common molecular marker shared between single maps are constructed (Fig. 21). Furthermore, a core-map for new crosses can be enriched with new marker in order to saturate the genetic map and clone identified genes.

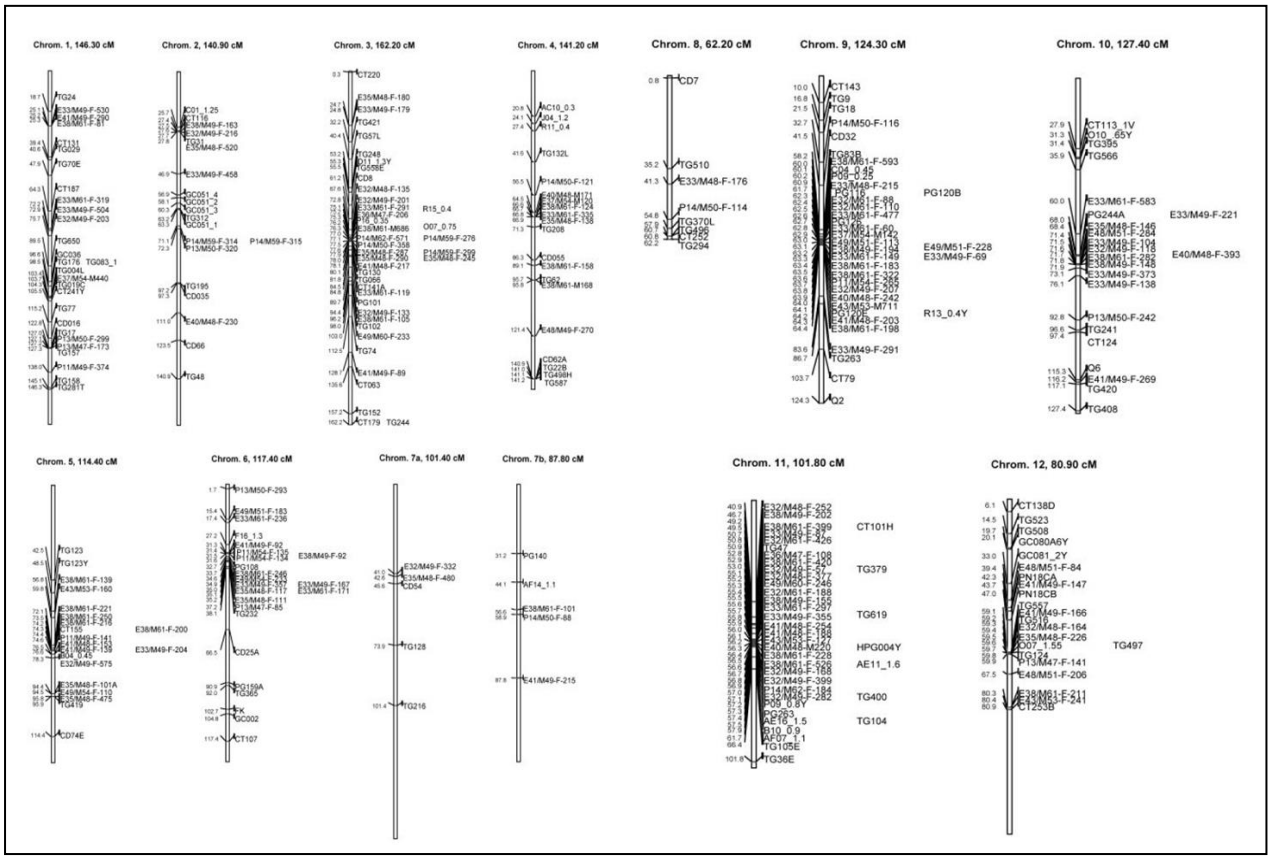


Figure 21. Example of integrated map (source Sushil *et al*., 2016)

Available genome sequences

Large scale genomic resources provided from new genotyping technologies and next generation sequencing are fundamental tools to understand molecular mechanisms related to important traits of pepper genome. They represent an essential instrument for genetic improvement as crucial information about divergence, evolution and domestication of various species (Sehgal, D., 2016). The most cultivated pepper (*C. annnum*) is diploid and consist of n=12 chromosomes. Based on chromosome numbers, the *Capsicum* species were characterized into two groups, respectively 2n = 24 and 2n = 26. Only one species, *C. annuum* var. *glabriusculum*, is tetraploid (2n = 4x = 48). Genome size is one of the largest within Solanaceae family (~3.5 Gb) with an estimate of 75-80% of repetitive elements and includes 35 thousand of genes (Hulse-Kemp *et al*., 2018). Nowadays, most updated pepper genomes are available in Sol Genomics Networks (SGN) (https://solgenomics.net). Furthermore Criollos de Morelos CM334 (Mexican landrace hot pepper) genome is available on https://peppergenome.snu.ac.kr/, while the genomic information of Zunla-1 (widely cultivated accession) and Chiltepin genomes (wild progenitor) is https://peppersequence.genomics.cn/ (Fig. 22).

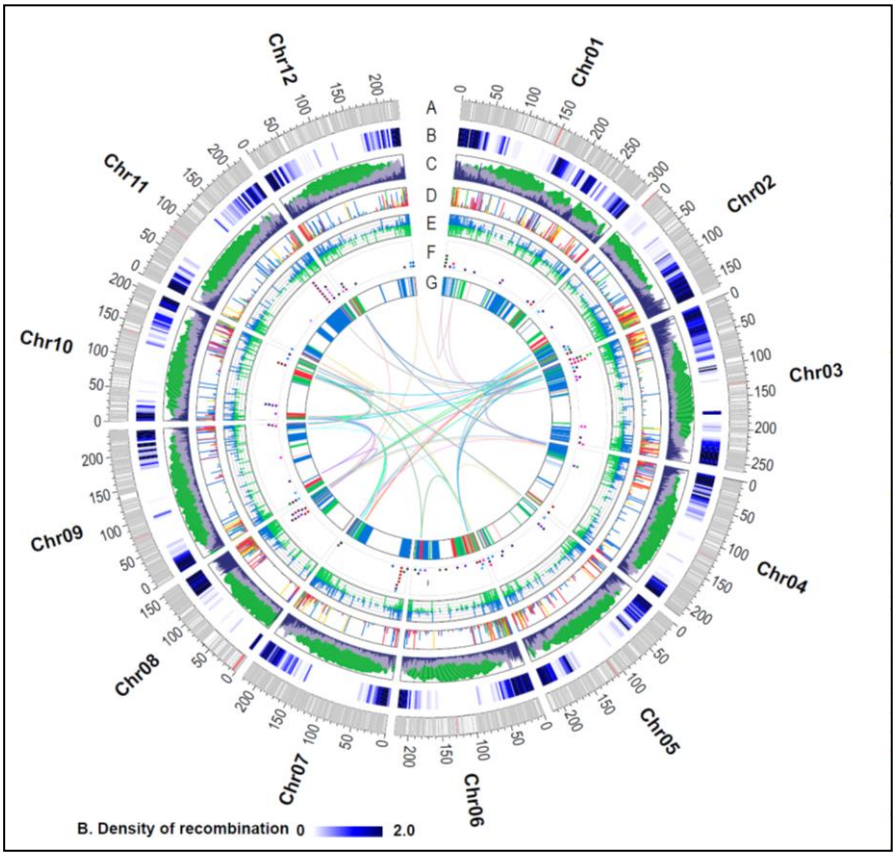


Figure 22. Global view of the pepper genome (source Qin *et al*., 2014).

1.2 Plant breeding: an overview

Selective breeding has been employed for thousands of years to introduce into plants qualities and specific traits that were considered desirable to consumer. Selection of the best plants to provide seed for their next crop was the beginning of selective breeding carried out from early farmers. Without the knowledge of genetics, they have been utilizing it to improve their crop for commercial purpose through elimination of undesirable characters altering the natural process of evolution. They naturally tried to capture eligible characteristic (disease resistance individuals or especially prolific genotypes) by crossing them into other plant changing through selection of desirable plant. Mutant alleles for desirable cultivation behavior, but unnecessary for wild survival, were selected. Producing food was the focus at this stage of agriculture rather than finding nature mechanism. In this way, for 10,000 years, informal selection provided by farmers led, by the end of 18th century to the creation of thousand different cultivars generating an expansion of the existing genetic base. Some essential elements of plant breeding were already known by this time but knowledge of the genetic variation was still poor (Xu, Y., 2010). Several definitions of plant breeding have been proposed during the last century such as "evolution directed by the will of man" (N.I. Vavilov) or "the art and science of improving the heredity of plants for the benefit of humankind" (J.M. Poehlman). Most universal description has been proposed by Bernardo R., 2002, as "Plant breeding is the science, art, and business of improving plants for human benefit". During 20th century classical genetics revolutionized plant breeding while, at the beginning of 21th century, genomics is leading to a new revolution (Pérez-de-Castro, A.M. *et al*., 2012).

Human intervention for centuries has altered relationships among ecosystems used for fuel, feed, food, shelter and medicines. Acceleration of the rates of extinction of species caused by human colonization, intensification of agriculture and industrialization, is one of the most profound and irreplaceable changes that humans have impressed. Multiplicity of factors are responsible for erosion in biodiversity such as loss, degradation and fragmentation of habitats. The contribution of Green Revolution to the loss of biodiversity is partly due to the narrowing of genetic diversity related to extensive mono-cropping. Irrigation and chemical inputs, which rely on Green Revolution technology, have reduced the fertility of soil rendered in some cases inhospitable to other species. Selection pressure and genetic drift, in addition, produced cumulative genetic erosion that taking place in the field (Esquinas-Alcázar, 1993). Depending

on point of view, the reduction in genetic diversity in crop plants include the decrease in number of different crop species being grown and, at the same time, the relative diminution of genetic diversity. The transition from primitive to 'improved' cultivars through plant breeding is considerable one of the major factors for genetic erosion.

Selection for closely defined aims and relative uniformity in plant ideotypes are therefore processes leading to a reduction of genetic variation (i.e. related to gene pool restriction).

1.2.1 Selection index and genotypic frequencies

The need of improving multiple traits at a time recurs frequently in most plant breeding programs. At the same time, it is important consider that the improvement of one single trait may have a deterioration effect on other associated genes, so all traits must be considered simultaneously in a crop species. Index selection, independent culling and tandem selection are the three appropriate selection methods for the traditional improvement of two or more traits. Selection index represent a single score of minus and plus of target traits (Smith, H.F., 1936). Relying on this index, selection among individuals provide a method for improving multiple factors in a breeding program. Independent culling is based on minimum levels of score for each trait. Below the culling level the individuals will be eliminated from the population allowing the selection of only genotypes meeting requirements.

For the tandem selection, traits are managed one after the other (e.g. selection of the second trait will be carried in a population selected for the first trait and subsequent screenings will be managed in the same way). A group of individuals that can mate or be crossed to each other to produce fertile progeny is the definition of biological population.

1.2.2 Response to selection towards combining ability

Different definition have been used for the 'response to selection' such as genetic progress, genetic advance, genetic gain, predicted progress or gain (denoted as R, GS and ΔG respectively). Response to selection (R) is a measure of difference between selected population and their progeny. Heretability of trait determines how much the selection differential is achieved in the offspring population (Nyquist, W.E., 1991). Breeder's equation, $R = \frac{ir\sigma A}{t}$, summarizes the response to selection and is used to increase the genetic gain of their program. Factors that constitute it are: i (selection intensity), σA (square root of the additive genetic variance), r (selection accuracy) and t (cycle time) (Falconer *et al*., 1996; Cobb *et al*., 2019). Phenotypic selection is usable for increase of response to selection but is difficult for

genetic traits with low heritability. Different molecular genetics approaches such as marker assisted selection (MAS) or genomic selection (GS) can be used to maximize the genetic gain through the increasing of the selection accuracy in a breeding cycle (Rutkoski *et al*., 2015). Breeding programs optimization for GS offer several opportunities to increase the genetic gain. Among the main aspect that can lead to a benefit for genetic response there are also factors such as increasing the number of evaluated lines, re-setting of field design, phenotypic data collected within different growing environment, selection accuracy and intensity (Endelman *et al*., 2014; Hoefler *et al*., 2020).

How two inbred lines can be combined together to produce a productive hybrid or to breed new inbred lines are key concepts in plant breeding which is summarized with the term of 'combining ability'. Selection of parental lines with this strong capacity is one of the most important breeding aims. Sprague and Tatum (1942) described two categories of combining ability: general combining ability (GCA) and special combining ability (SCA). Between parental lines, these two parameters are the incident factors that determine the hybrid performance. GCA is measured as the average performance of all hybrids made with that inbred line and could be defined as the inbred line attribute. So, higher is the GCA value of an inbred line, higher is the average of its hybrid performance. SCA is instead related for specific parent combinations and is measured by the hybrid performance deviation from the expected behavior related to GCA of the parent. Heterosis is the measurement of hybrid performance, that will be determine by GCA of the parent and SCA of the cross. Two components ascribable to GCA and SCA compose the total variation among hybrids where lines are crosses between several other lines in a systematic manner. The mean performance of a cross ($\bar{x}_{AB}$) between two inbred lines A and B can be represented as: $\bar{x}_{AB}=GCA_A+GCA_B+SCA_{AB}$. GCA of the parents A and B, respectively GCA_A and GCA_B and the cross of A × B is expected to have a result equal to the sum ($GCA_A + GCA_B$) of the GCA of their parents. Additive and non-additive genetic variances are respectively the responsible factor for differences of GCA and SCA of lines (Xu, Y., 2010).

1.2.3 Marker assisted hybrid prediction

Usually breeders are unable to check all parent combination available so they are forced to test a limited number of cross with a percentage of new commercial hybrids very low. Therefore, prediction hybrid performance is an important topic in all hybrids breeding programs. Combination of markers and phenotypic data is the aim of numerous studies developed to achieve a reliable method for hybrid predicting performance. Genetic

differences or heterozygosity between the parents generate heterosis, this is the reason why hybrid prediction has been widely used for the evaluation of genetic differences of parentals genotypes. In this frame, molecular marker have been used to study correlation between parental genetic distance (GD) and hybrids performance, providing a strong support for the hybrid prediction.

However, while different studies showed a positive correlation between GD of parentals and hybrid performance (Lee *et al*., 1989; Smith *et al*., 1990; Stuber *et al*., 1992; Reif et al., 2003), other demonstrated no significant correlation (Godshalk *et al*., 1990; Dudley *et al*., 1991).

Generally, relationship between molecular marker and single cross performance have been low in term of predictive value (Melchinger *et al*., 1990; Dudley *et al*., 1991).

Estimations based on molecular GD failed to predict hybrid performance in different crop such as soybean (Gizlice *et al*., 1993), chickpea (Sant *et al*., 1999) and pepper (Geleta *et al*., 2004). Using a set of DNA markers chosen for good genome coverage is a common practice to determine GD but not for linkage to genes related heterosis for specific trait. There is no correlation between increasing of marker density and improving estimation capability. Other point of view is the selection of markers tightly linked to genes affecting heterosis of target traits. However, there is no association between allelic differences at marker loci and allelic differences at linked loci for heterosis (Stuber *et al*., 1999). Heterosis may depend from the balance of favorable and unfavorable gene interaction rather than heterogenic gene combination. Specific gene combination from the two parents could be responsible for a positive influence in heterosis. For these reason, in the field of parentals improvement and hybrid prediction, investigation of specific combinations of genes that are associated to the heterosis improvement should be more important than focus on any single gene or QTLs.

The term “breeding by design” is used for describe the possibility to predict the performance of a cross considering information generated from molecular markers (Peleman, J.D. and Van der Voort, J.R. 2003). Step foreseen in this process start with mapping all the loci for all relevant traits and next the evaluation of respective allelic variation, concluding with breeding by design. For the first step, linkage disequilibrium (LD) mapping or GWAS studies avoid the limitations of biparentals population and for these reasons has been suggested.

Once obtained information of phenotypic value assigned to each locus, contribution of specific alleles allows creation of crosses to generate superior genotypes for presence of favorable traits. This procedure is used also for patent application (e.g. higher quality in maize) (Pérez-de-Castro *et al.*, 2012).

1.2.4 Plant Breeding for resistances

Yield losses due to diseases are estimated every year in ca. 13% of global harvest. Main areas are tropical or sub-tropical zones where the pathogens development is favored (Nelson *et al.*, 2018). For over the past half century chemicals and pesticides have been used to reduce the diseases effects on the crops allowing the farmers to intensify production systems without increasing yield losses.

Disadvantages of these methods are several such as the development of pesticide resistances by pests, reduction of plant biodiversity, risk for food security and human health (Stoytcheva, M., 2011). Therefore, it was necessary to employ alternative approaches, as breeding for resistance, which allows, through plants selection, to exploit plant defenses system. Use of genetic resistance for managing pepper diseases is the most cost-effective control method. Thousands of dollars and years of research are spent for characterizing and introgressing these resistances in cultivated varieties. Aim of breeder is to introduce these resistances from wild relative species into novel adaptive germplasm (Mundt, C. C., 2015).

Qualitative and quantitative resistance are typical mechanisms of resistance against main plant diseases. These two typologies are distinguished for type of genes at the basis of resistance control, kind of associated inheritance and phenotypic expression (Niks *et al.*, 2015). R genes controlled the qualitative resistance (Poland *et al.*, 2009). Through pathogen recognition, these genes encode for proteins (known as R proteins), which guarantee complete or near-complete resistance.

Usually, the qualitative resistance is characterized by dominant R genes while recessive resistance is generally related to following loss of function of dominant gene and consequently susceptibility to disease (Jamann *et al.*, 2014). On the other hand, more genes with small effect characterize the quantitative resistance, known as quantitative trait loci (QTL) or quantitative resistance locus (QRL).

Generally, R-genes that can be rapidly overcome by new variant of pathogens and consequently quantitative resistance is considered be more durable than qualitative resistance. At the same time, quantitative resistance reduces the pressure against the creation of new variants (Borelli *et al.*, 2018). Another important aspect of the response to pathogen are the so-called S-genes, responsible for the loss of susceptibility that leads to resistance.

The terms of ‘plant S-gene’ has been introduced in the 2002 by Eckardt (2002). Cases studies of S-gene functions are reported such as *PMR6* gene in 2002 that was discovered for coding a susceptibility factor for promoting growth of powdery mildew in *Arabidopsis*. In 2005 was

suggest to exploit susceptibility genes associated for nematode resistance as an alternative in breeding after the identification of genes like *Mlo* and *eIF4E* (De Almeida Engler *et al*., 2005). Well-characterized examples of S-genes is the transmembrane MLO (Bhat *et al*., 2005; Panstruga R., 2005; Hardham *et al*., 2007). In barley and arabidopsis, loss of function mutations in *Mlo* result associated with the resistance to adapted powdery mildew (Piffanelli *et al*., 2004; Humphry *et al*. 2006). Also in tomato, has been show that recessive allele *ol-2* is responsible for resistance to powdery mildew fungus *Oidium neolycopersici* through a null mutation of the tomato *SlMlo1* (Bai *et al*., 2008; Pavan *et al*., 2008). Also for pepper, in susceptible plants, effector protein AvrBs3, is involved in this process (Kay *et al*., 2007). In resistant plants, AvrBs3 activates pBs3 (pepper *Bs3* gene promoter), which leads to the specific expression of the R-gene *Bs3* and disease resistance. The improvement of the exploitation of S-genes is a current challenge to provide an alternative breeding strategy that can be use complementary to R-gene methods.

1.3 Modern techniques for pepper breeding population refinement

For over the last 10,000 years, plant breeding has gradually transformed from art to science. The development of molecular tools created the present molecular design-based science making plant breeding easier, quicker, more efficient and effective (Phillips, R.L., 2006). Current plant breeding now foresees advanced approach for the identification and creation of the genetic variability, understanding at the same time the breeding population structure related to genes details (function, position, their relationship with environment and other genes).

Other integrated aspects allow the breeding of today to improve the efficiency toward recombination of novel favorable alleles within cultivar or hybrids selecting the best genotype for specific characteristic. GenBank databases repositories are enlarging the data available very fast and now sequencing data for many species are available (Phillips, R.L., 2008). Great step forward is taking place about the field of gene interaction and specific function.

Multiple gene study of the expression, DNA chips and further technology are been developed such as bioinformatics tools and high throughput robotics. DNA-based molecular markers allowed the intense chromosomes genetic mapping and consequently, prediction of gene location increasing therefore the selection efficiency for agronomic traits and the pyramiding of interested genes.

1.3.1 Marker-assisted breeding methods

To recover an ideal individual with specific genetic characteristic (the 'ideotype') applying crossing-and-selection strategies may require large populations for both sizes and structure (Bonnet *et al*., 2005).

Choice of the most effective strategy can affect dramatically the reduction of resources needed to combine a set of target traits into a superior genotype. Factors such as choice of best MAS methods of most appropriate cross (e.g. single cross, BC or top-cross) could determine a positive boost obtainable by the breeders.

On this frame, aspects often faced by breeder can relate to the use of incomplete linked markers to merge target alleles or add linked alleles in repulsion with segregating crosses for other unfavorable alleles.

The use of genetic population theory can be used as general rules for understand the best crossing strategies or, where there is no recombination between marker and allele of interest, the inbreeding level for enhance the marker implementation efficiency (Wang *et al*., 2007).

The efficiency of MAS depends on many factors such as nature of the phenotyping, number of markers used, genomic region containing the desired QTLs, proportion of additive genetic variance explained by the marker, design and analysis of the experiment and the distance between marker loci.

Time and resource allocations are the main positives factors offered by MAS to the private breeding programs.

However, in the attempt to combine more target traits, often a general loss of breeding gain with an elevated number of breeding cycle is observed. MAS is a techniques that can improve both the speed and the precision of this process (Xu, Y., 2002).

Molecular markers

Molecular markers are genetic loci linked with a trait chromosome defined as sequences DNA differences that can be inherited between different individuals and that can be investigated through the polymorphisms (Kesawat, M. S. and Das Kumar, B., 2009).

Specific events such as deletion, insertion, translocation and point mutations, which do not necessarily interfere with gene function, are at the evolution base.

For these reasons, the molecular markers are fundamental tools within the science of genetic improvement to investigate genetic diversity between the individuals and the populations.

A molecular marker is defined ideal when is not influenced by the environment, polymorphic in coding and non-coding sequences, highly reproducible, stable, co-dominant, analyzed with a simple, fast and automatable way (Mondini, L. and Noorani, A.P, 2009).
Often, not all these characteristics may be present together in a single molecular marker, therefore the most suitable marker is chosen through the application characteristics.
Molecular markers are divided based on (Semagn, K. and Bjørnstad, A. N., 2006):

- number of loci per marker (single or multiple) and inheritance (dominant or codominant);
- transmission mode: biparental nuclear inheritance or uniparental nuclear inheritance, maternal organelle inheritance, paternal organelle inheritance;
- requirement of prior sequence information;
- detection method: hybridization-based, PCR-based, next-generation technology.

Based on the method of detection, molecular markers are classified as hybridization-based markers and PCR-based markers. In the first group, DNA is digested through a specific restriction enzyme, after that the resulting fragments are hybridized with a labelled probe or a DNA fragment of known sequence that allow the DNA profiles identification. In the second group, sequence DNA amplification is done using both thermostable DNA polymerase enzyme and specific primers.

- **Restriction Fragment Length Polymorphism (RFLP)**

Restriction fragment length polymorphism (RFLP) marker was the first molecular marker able to detect the difference between DNA fragment sizes in individual belonging the same species. The polymorphisms resulting as gain/loss of restriction sites due to insertion/deletions (InDels), point mutation, translocations, duplications and inversions, prevent the restriction enzyme cut.
Usually applied for diversity and phylogenetic studies among individuals within populations and gene mapping studies, the RFLP are highly reproducible, co-dominants, locus-specific, random located. Nowadays, RFLP are rarely used due to high DNA requirement (quality and quantity), time-consuming, labor-intensive and time cost (Kesawat, M. S. and Das Kumar, B., 2009).

- **Random Amplified Polymorphic DNA (RAPD)**

Random amplified polymorphic DNA markers are DNA fragment amplified able to amplify from one to ten sites simultaneously through PCR using short random primers (10 bp).
Polymorphisms are detect via presence or absence of specific size band related to the primer annealing sites variation (John *et al.*, 1990).
RAPD markers are randomly distributed and abundant into the genome, dominants, easy to assay, highly reproducible. Moreover, they require less time consuming and low quantity of DNA. On the other hands RAPD are not locus specific and little reproducible. For this reason require standardized analyzes (Kesawat, M. S. and Das Kumar, B., 2009).

- **Amplified Fragment Length Polymorphism (AFLP)**

To overcome the limitations of RFLP and RAPD techniques, Amplified Fragment Length Polymorphism markers (AFLP) were developed (Vos *et al.*, 1995). These combine PCR amplification technology with restriction fragments technique (RLFP).
The technique is based on the use of restriction enzymes for DNA digestion, followed by oligonucleotides ligation (adapters) to the digestion products. Then, adapters are recognized by primers, which are amplified via PCR.
AFLP markers are randomly distributed into the genome, economic, abundant and prior sequence information are not necessary. Disadvantages are related to the dominance of alleles and the difficult to the interpretation of results (high number with different produced bands intensity) (Lynch M., 1998).

- **SSRs or microsatellites**

Simple sequence repeats or microsatellites (SSRs) are short tandem repeated DNA stretches, very abundant in the genome of different organisms, highly polymorphic (Rostoks *et al.*, 2006). SSRs are originated in response to the recombination of double-strand DNA and transfer of mobile elements (retrotransposons, mismatches and slippage of single-strand). The most common SSRs are mononucleotides A, T, dinucleotides AT, GA, trinucleotide AGG and tetranucleotide AAAC.
Primers for PCR are designed on flanking regions of SSRs. These regions are favorable to recognized and highly conserved. SSR markers are characterized by co-dominance, low request of DNA quality and quantity, high reproducibility, high abundance in the genome. SSRs are used mainly in plant mapping studies (Kalia *et al.*, 2011).

- **Cleaved Amplified Polymorphic Sequences (CAPS)**

Cleaved Amplified Polymorphic Sequences (CAPS) markers are DNA fragment amplified by PCR followed by digestion with restriction enzymes (Jarvis *et al.*, 1994). The CAPS products are DNA fragments with different length causing from variation in the restriction sites, which are identified by agarose or acrylamide gel electrophoresis. CAPS markers require low quantities of DNA, are codominant and highly reproducible. Moreover, this technique is characterized by simple and relatively economical equipment and the easy result interpretation. CAPS markers are used in molecular identification, genotyping and map-based cloning studies (Weiland J., 2003).

- **Sequence-Characterized Amplified Regions (SCAR)**

Sequence characterized amplified region (SCAR) markers derive from the RAPD markers, but are more specific and reproducible. These markers are identified genomic fragments by PCR amplification with primers designed from nucleotide sequence of RAPD fragment. Codominant markers, mono-locus, reproducible, reliable and easy to use. Often are applied in marker assisted selection, physical mapping, comparative mapping or homology investigation among related species and gene mapping studies (Yang *et al.*, 2013).

- **Single-Nucleotide Polymorphism (SNP)**

Comparing DNA sequences from different individuals in the same population allows the identification of single nucleotide polymorphisms (SNP), single base-pair changes caused by substitutions such as transitions (purine/purine), transversion (purine/pyrimidine), or by insertion or deletion, which form the so-called InDel category (Kesawat, M. S. and Das Kumar, B., 2009). SNPs are random distributed throughout the genome of genotypes belonging to several species with frequency between one SNP in every 100-300 bp.

Moreover, SNPs can be present with different frequencies in coding, non-coding or intragenic regions (Xu Y., 2010) highly reproducible, cost effective, codominant. SNPs are excellent markers to study the differences between genotypes and usually applied in mapping studies, marker-assisted breeding as well as map-based cloning.

To date, methods available able to make SNP markers more efficient are related the high-throughput genotyping. Among these there are Next Generation Sequencing (NGS), allele-specific PCR, chip-based NGS and genotyping by sequencing (GBS) (Agarwal M. and Shrivastava N., 2008).

1.3.2 Gene pyramiding for major genes

Different breeding programs are based on the introgression of more than one gene or QTLs associated with interest traits into one single genetic background, in a process named pyramiding.

Gene storage from different sources is one of the most applicative aims of gene pyramiding. With the traditional phenotypic selection, plants are evaluated for multi-traits; therefore, it may complicate the evaluation of segregant population for gene using destructive bioassay. For these reasons, DNA marker can facilitate this process considering that can be test without phenotyping aspects. Additional positive factor related to gene pyramiding under the plant genomics approach regards the number of plants that must be analyzed.

In a pyramiding program this factor will increase with the number of introduced loci.

The genotyping platform can provide a larger screening efficiency and ensure a probability that a positive allele is present in a population, (Pérez-de-Castro *et al.*, 2012). MAS is particularly used for pyramiding of diseases resistances.

Alleles associated for resistance to biotic stress and genes for agronomic or seed quality can be accumulated together to maximize the MAS improving genetic background. Multiple genes for resistance represent one of the most important tools for extending the cultivar lifespan.

Reason of this is the importance of development of durable resistance, considering that pathogens frequently could overcome them with the emergence of new races. Reducing the ability of pathogens to adapt to resistance gene, combination of multiple genes can provide a broad spectrum of resistance (Kloppers F.J., and Pretorius Z.A, 1997; Shanti *et al*., 2001; Singh *et al*., 2001). Fixing all genes of interests cannot be realized in a single selection step so is necessary to select genotypes with complementary sets of homozygous loci and cross them.

A general framework to optimize breeding schemes in order to accumulate identified genes from multiple parents was developed by Servin *et al*. (2004). Accumulation of favorable genes related to different traits in a single genotype could be summarized with the term of gene pyramiding, one of the oldest plant breeder challenges.

Critical concept is represented by the possible “direction” of the stored allele at different traits. For instance, negative alleles could be positive for some traits but also favorable alleles are positive for others.

Consequently, to reach the breeding aims there could be the need to combine positive QTLs alleles for some traits with the negative ones of others.

Considering multiple traits, each of these must be checked in various development stage or different environments of a breeding program as in phenotypic selection. While applying gene pyramiding for different traits, is important give attention to the trait correlation.

In fact, positive relationship will increase the pyramiding process success about the selection of alleles with the same favorable direction. At the same time will hinder the screening for alleles with different positive direction as for negative correlation (Xu *et al*., 1998).

- **Marker assisted recurrent selection (MARS)**

Marker-assisted recurrent selection (MARS) was proposed in 1990 (Edwards M. and Johnson L., 1994; Lee M., 1995; Stam P., 1995). Using at each generation molecular markers, to mark all importance traits, the goal of MARS is to obtain individuals with as many accumulated favorable alleles as possible.

Usually genetic information achieved on experimental population from QTLs analysis. Once defined QTL mapping of a biparental population, both parents can contribute positive alleles, resulting the superior genotype as a mosaic of chromosomal segments from the two parents. This technique can be applied through various versions in relationship of which molecular markers information are incorporated.

Availability of genetic data at each selection generation for all the progeny with the integration of genotypic and phenotypic data, represent the key advantages of these breeding method.

Anyway, all studies conducted revealed that MARS can be considered superior (between 3% and almost 20%) to phenotypic selection for accumulating interested alleles in one single individual (Van Berloo R. and Stam P., 2001; Charmet *et al*., 1999). Compared to phenotypic selection, MARS is advantageous when the selection population is larger or more heterozygous including BC1 or F2 populations.

Practical MARS definition could refer to the improvement of segregant population (F2) by one MAS cycle associated to phenotypic data followed by next marker-based selection (only marker score) (Johnson, G.R., 2004).

This method showed that in ten generation and with fifty detected QTLs in a population of 200, frequency of positive alleles could arrive up to reach up to 100% when markers were located perfectly on the QTL (Hospital *et al*., 2000).

- **Marker assisted backcrossing (MABC)**

Backcrossing has been a widely used technique for almost a century, as a method used in the plant breeding to incorporate one or a few genes within an elite variety, described for first time in 1922 (Stoskopf *et al*., 1993).

Efficiency of selection greatly increases with the use of DNA markers in backcrossing creating a new level of selection named marker-assisted backcrossing (MABC) (Holland, J.B., 2004). This technique use marker to replace phenotypic screening for target gene or QTLs. Particularly indicated for traits that have laborious or time-consuming procedure, MABC can be used to select alleles in the seedling stage, allowing best plants selection for next backcrossing. Recessive alleles can also be detected, which could be difficult to identified using conventional method. Second step is referred to 'recombinant selection' with the screening of selecting BC progeny for target gene using flanking marker. Reducing size of donor chromosome segment keeping target locus is a purpose of recombinant selection. This aspect is referred to the concept of 'linkage drag' (undesirable gene that are linked to target allele coming from donor parent).

Using traditional breeding, often donor chromosome segment remains large even different BC generation while using flanking gene marker (e.g. less than 5 cM on either side), linkage drag can be reduce (Hospital *et al*., 2000). Third level of MABC is referred to as 'background selection' because it allows to selection of BC progeny with unlinked marker to the target locus but associated with recurrent parent (RP).

In this way, it is possible to enhance the genome proportion of recurrent genotype. Traditional backcrossing need for recover RP a minimum of six BC (genome recovery would be 99.2%) while using marker it can reach by BC4 or even BC2 with a great time-consuming reduction (Visscher *et al*., 1996; Hospital F. and Charcosset A., 1997; Frisch *et al*., 1999). Different advantages are offered by MABC over conventional breeding such as speed up for recovery of recurrent parent genome, reducing linkage drag and biosafety for the absence of pathogen inoculation.

- **Marker-assisted reverse breeding (MARB)**

Breeders have to create new hybrids through the cross of homozygous parental line. In this frame, the genetic improvement of parental lines represents the tool to improve hybrid performance.

Methods for the easy regeneration of heterozygosity in genotypes is one of the challenges in breeding hybrids oriented. For long time breeding strategy, started from hybrid material, has been strongly related to breeder experience to obtain parental lines. Breeding cycle is generally long (six-seven years) with low recovery efficiency of positive alleles.

Reconstitute elite parent plants from the original hybrids was proposed in 2006 by Dirks *et al*. Here homozygous parental lines are ri-created from heterozygous plants through the suppression of meiotic recombination.

Basic procedure consists of inactivation of the chromosome recombinase to eliminate the recombination in the segregant progenies of the hybrids. Producing complementing parental lines for any heterozygous plant through achiasmatic meiosis (meiosis without crossovers) is a modern plant breeding method named reverse breeding (RB). Generated gametes are the developed as doubled-haploid (DH) plants each of them can be crossed to reconstitute the initial hybrids. Beginning method foresees a screening of DH plants for RNAi construct absence and then crosses for the hybrids generation (RMRB-Rnai mediated reverse breeding). Marker-assisted reverse breeding (MARB) is an innovative, simple and fast molecular breeding method applied to reverse breeding field. Such technique allows to reverse any hybrid into homozygous inbred lines with a high similarity level to its original parent line. Reverse breeding has been applied so far in different crops but only at a research stage (Lusser *et al*., 2012). Usually, a small number of molecular markers are necessary to verify the absence of crossover event (marker being placed at the two distal ends and a third marker placed near the centromere). Current method (MARB) is chip-based where genotyping performed to lead the selection for similarity to the parent on homozygous plant (Guan *et al*., 2015). Issues to consider and reverse breeding limitation can be listed as:

- usable only in crops where double haploid technology is applicable;
- limited to species with maximum 12 or less chromosome because in bigger genome the number of non-recombinant double haploids needed for complementary searching that regenerate the original hybrid, would be too high and practically difficult (Lusser *et al*., 2011);
- complete homozygosity of plants generated represent a limit to genetic variation (Van Dun and Dirks *et al*., 2006). Anyway, RB and now MARB could be used as an intermediate step of breeding procedure with a huge implication in varietal genetic improvement as is able to generate, from complex genotypes, the homozygous parental lines. Self plants with a substituted chromosome are an interesting breeding aspect that will allow to breeder a better efficacy at single chromosome level for the creation of lines with specific introgressed traits.

- **Genomic selection and prediction**

First described in 2001 (Meuwissen *et al*., 2001), genomic selection is based on the availability of both phenotypic and genotypic data for a reference population.
With a simultaneous estimation of markers available, haplotypes and loci phenotyping, this approach differs from the other MAS method because do not rely on pre-marker selection. Data collected on a training population are used to create a model where phenotype differences are explained by markers given information. Once generated the model, each selected candidate will be associated with a genetic value depending on the genome-estimated breeding values (GEBVs).
This prediction model combines markers with phenotypic data in order to increase the accuracy of prediction.
As compared to MAS, GEBV is dependent on all markers including major and minor markers effects and a minimal marker genome coverage is required (Lorenz *et al*., 2011).
In general, the decreasing genotyping cost brings to a growing utilization of this approach also for the convenience in cost-effective (Pérez-de-Castro *et al.*, 2012).

1.4 Additional factors interplay for breeding improvement

1.4.1 Enviromics

Phenotypic variation is a product of genetic and environmental variability observed across diverse environments. Especially for multiple environments, enviromics represent a viable solution to this bottleneck combining both the application of modern prediction tool and also traditional phenotypic selection. In this frame, innovative approaches have integrated genomic and phenomic information considering different enviromics background as crop growth model (CGM) outputs (Heslot *et al*., 2014; Rincent *et al*., 2019), CGM integrated with GS (Cooper *et al*., 2016; Messina *et al*., 2018; Robert *et al*., 2020), historical weather records to predict cultivars in years to come (De Los Campos *et al*., 2020), and linear and nonlinear reaction-norm models (Jarquín *et al*., 2014; Morais-Júnior *et al*., 2018; Millet *et al*., 2019; Monteverde *et al*., 2019).
A significant strategy in this direction was proposed by De los Campos *et al*. (2020): a method for the analysis of multienvironment trials evaluating yield and stability capacity. Procedure proposal evaluate the patterns GxE levering from field trials and predicting, in correlation to an environment, the expected performance of a genotype.

In order to carry out and improve prediction accuracy in new environments (characterized with the same environmental characteristics), recently (Rogers *et al*., 2021) emphasized the importance of incorporating high throughput environmental data into genomic prediction models.

1.4.2 Phenomics and genomics interconnection

Time and cost necessary for field measurement of interesting trait are the most important limitation for determining accurate phenotypes. Phenomics, that study plant phenotypes under a range of environmental conditions, can be used in efficient mode to collect data on agronomic traits (Atkinson *et al*., 2018). Named also high-throughput phenotype (HTP), its objective is to reduce the cost of data increasing the prediction accuracy early in the crop-growing season. Genomic models take in consideration together multiple environments and multiple traits, giving a great exploitation potential between variable and derived effects. Field of phenotyping is the crucial step necessary to achieve a large and inter-operable phenomics dataset. These are the tools for the characterization of genetic materials in different environment. These data, in addition to pedigree and genomic information, have the potential to increase the prediction accuracy of a breeding program multi-trait and multi-environment oriented. Multi-trait data from the multi-environment trials should be collected during breeding programs to exploit the possible correlation among specific traits of interest.

The use of genomics with enviromics gives bases for the prediction of cultivar across different growing condition (Jarquín *et al*., 2014; Messina *et al*., 2018; Millet *et al*., 2019). Use of large-scale of envirotyping data allow to predict from a wider number of growing conditions from historical climate, soil data and train GS models.

Enviromic matched Genomic Prediction, E-GP, is an efficient approach to better use of environmental databases to link genotype-phenotype variations and explain phenotypic variations across environments. This predictive breeding tool can lead to more efficient training sets for GS contributing to the optimization of experimental networks of field trials (Rincent *et al*., 2019). In addition, prediction of single cross-hybrid breeding values of the selection candidate can be determined for the early stages of selection by genomics and enviromics study (Morais-Júnior *et al*., 2018; Costa-Neto *et al.*, 2020). Phenomics and enviromics benefits offer the capacity to improve the accuracy and efficacy of breeding pipelines. Enviromics, phenomics and genomics should be interconnected offering an opportunity to accelerate genetic gains and enhance the breeding program potential.

1.4.3 Criticity in breeding schemes

Breeding pipeline is created to point out the combination of factors such as population improvement, introgression efforts, trait discovery and seed commercialization only after a deep market study and investment impact of any company or institution (Fig. 23).

A breeding program may be specific for more market segments and can be associated with multiply breeding schemes.

Elements that define a breeding strategy on different stage are represented by the acronym CES: Crossing, Evaluation and Selection (Henryon *et al*., 2014; Yabe *et al*., 2017; Cobb *et al*., 2019; Pook *et al*., 2020; Gaynor *et al*., 2021).

CES objectives are usually visualized as illustrative tables of flow chart but often this representation cannot provide information necessary to reproduce in an effective way the breeding schemes. Such tables usually include:

- evaluation (number of locations, replication level within and among locations, number of checks, plot size, experimental design);
- selection (selection method/unit and percentage of individuals selected);
- crossing (number of parents/crosses/progeny, allocation method) (Covarrubias-Pazaran *et al*., 2022).

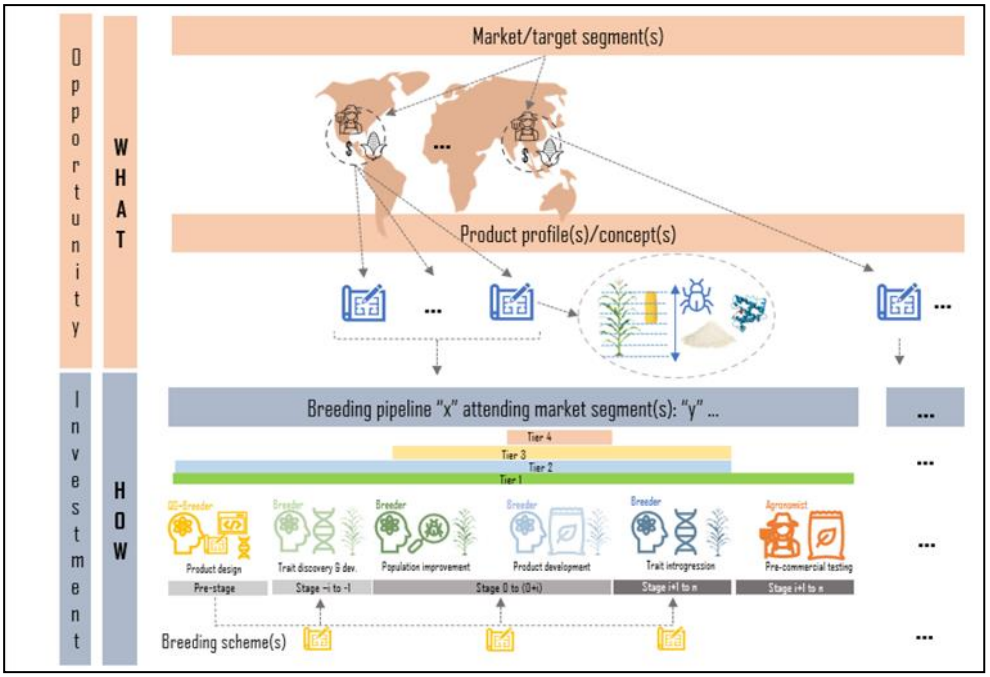


Figure 23. Representation of the relationship between target segment and pipelines of a breeding program (source Covarrubias-Pazaran *et al*., 2022)

Other elements that make the breeding schemes communication complex depends, for instance, on species biology or complexity of market product profile in response to the desired final product (Henryon *et al*., 2014).

Despite this CES decisions complexity, rarely the breeding schemes are shared via official documents. It is common for the head of a breeding department to distribute among the staff the schemes received from his own predecessor in a verbally/practically way rather than providing a writing description using documents or software (Covarrubias-Pazaran *et al*., 2022). This transfer methods, also based on free interpretation, unfortunately has led to an important losses of both germplasm and information with the disappearance responsibility of many programs in the last century (Baenziger, P.S., 2006; Gepts P. and Hancock J., 2006; Morris *et al*., 2006).

Better preservation of pedigree data and germplasm among breeder need the improving of transferring methods through an industrial processes form (Bernardo, R., 2002). In contrast to the traditional breeding approach, common during the 20th century, tools used in these methodologies are associated with statistical aspect and could be used to draft draw improving breeding schemes successfully. Use of simulations has a particular practical relevance in these complex processes (Li *et al*., 2012; Yabe *et al*., 2017). Breeding Pipeline Manager (BPM) is an example of software available as tool targeted in standardized breeding process (Covarrubias-Pazaran *et al*., 2022).

Edward Henry Weston (1886-1958)

Chapter 2

Aims of the project

In the context of climate change, growing population, increasing demands and expectation generated from whole agri-food system, the plant breeding value is constituting, in increasingly manner, a key role indispensable to satisfy the needs of the global agricultural sector aimed toward sustainability and innovation.

With a major contribution to increased yield in arable farming, plant breeding impact is facing multi-scale challenges such as improved market and trade conditions, increased food availability, additional farm income avoiding land use and reducing of gas emissions as well as loss of biodiversity. The collaboration between the University of Modena and Reggio Emilia and Esasem S.p.A. on this PhD project falls into this framework joining the commitment of public and private Italian research with a worldwide impact.

In line with company strategic objectives, this research doctorate aims at optimising the plant selection process through pre-breeding approaches related to *Capsicum* species.

To reach these goals, effective methods of genetic improvement were faced for the superior genotypes constitution.

In this view, the following main approaches have been identified and detailed:

- constitute a proprietary company collection characterized by as broad as possible genetic base, including pre-breeding crosses;
- characterize the collection with a 48 SNP assay in order to verify the level of homozygosity, study the genetic distances among accessions, study the correlations of polymorphisms with resistance genes, fruit characteristics as shape and provenience area, for the future conscious use of company genetic resources;
- develop a set of assays based on known markers for resistance genes retrieved from literature, to be used for new routine selection of the proprietary collection;
- characterize thoroughly the chosen elite genotypes useful as promising progenitors for future breeding programs and through re-sequencing identify a complete set of polymorphisms, investigate further core collection genetic composition as well as determinate a private genotypic fingerprinting profile for a possible patent development.

Predictive information derived from our research activity will be applied directly for gene pyramiding flow, improving the efficacy of the breeding process. The long-term objective of the present study is to conduct an efficient breeding program following a novel Inter-placing selection breeding scheme.

Edward Henry Weston (1886-1958)

Chapter 3

Molecular characterization of pepper (*Capsicum* spp.) collection

3.1 Introduction

The high demand for food production due to the population growth could represent one of main causes of biodiversity losses.

The study of genetic resources of a breeder germplasm collection is the most appropriate strategy for maintaining and enhance the variability of species of great importance, such as *Capsicum*. From this perspective, regarding to productivity and quality of *Capsicum*, exploiting the existing genetic variability and subsequently understanding is one of the most important strategies for plant breeding programs (Moulin *et al*., 2015). This is because the success of genetic improvement process lies in the genetic diversity available within studied germplasm. In this sense, *Capsicum* germplasm characterization through molecular markers is an effective breeder strategy that must be combined with solid phenotyping data (morpho-agronomic descriptors).

In this chapter, we describe firstly the germplasm collection constitution, which represent the basement of whole research activity. For a better genotypes choice efficiency, the EMEA (Europe, Middle East, Africa) company experimental net as well as additional in vivo resistances bioassay helpful for the accurate panel set-up, has been used.

Further activities concerned the creation of new cross combinations aimed to the broadening of genetic base of interest for this study.

A molecular characterization of the collection was carried out using KASP genotyping technology aiming at understand the genetic composition as complementary pillar toward the identification of elite materials useful for subsequent pre-breeding activities.

3.2 Materials and methods

3.2.1 Germplasm development

Plant material screening in selective environments for high pathogenic pressure was carried out to improve the phenotypic characterization and confirm the presence of genotypic variants with high resistance expression.

This activity, concentrating on advanced selection material, was developed through the monitoring of a large set of genotypes under examination within multi-location field trial distributed both throughout Italian country and abroad (Fig. 24).

Figure 24. Company experimental net for genotypes examination

Pathology bioassay (Fig. 25 and Fig. 26) for the characterization of resistances related each genotype has been conducted through the associated company Budakert Ltd (Hungary), supervised by Senior Scientist Dr. Gabor Csillery.

Artificial inoculation related *Tobamovirus*, *Potyviruses*, *Cucumber mosaic virus*, *Tomato spotted wilt virus* and Bacterial spot was conducted following technical guideline of CPVO, Community Plant Variety Office (https://cpvo.europa.eu/en/applications-and-examinations/technical-examinations/technical-protocols). Details of in vivo bioassay are reported below.

Tobamovirus

- Pathogen: *Tobacco mosaic virus* and *Pepper mild mottle virus*.
- Source of inoculum: Geves (FR), Naktuinbouw (NL).
- Isolate:
  - ✓ Tobacco mosaic virus Pathotype 0 (TMV: 0) strain mat/ref/06-05-01-01 (Geves);
  - ✓ Pepper mild mottle virus Pathotype 1-2 (PMMoV: 1-2) strain zb203 (Naktuinbouw);
  - ✓ Pepper mild mottle virus Pathotype 1-2-3 (PMMoV: 1-2-3) strain mat/ref/06-01-02-02 (Geves);
- Test facility: climate room.
- Temperature: 20-25°C.
- Light: 12 hours or longer.
- Growing method: sowing and raising of seedlings in boxes.
- Inoculation:
  - ✓ Preparation: 1 g leaf with symptoms with 10 mL PBS (Phosphate Buffered Saline). Homogenize, add carborundum to buffer.
  - ✓ Method of inoculation: rubbing of cotyledons with a virus suspension.
  - ✓ Plant stage at inoculation: cotyledon stage.
  - ✓ Inoculation method: rubbing with the virus suspension.
- Observations:
  - ✓ Method: visual.
  - ✓ Observation scale: TMV: 0, susceptibility: mosaic, growth reduction, death of plants. Resistance: local necrotic lesions, which can lead to leave drop, systemic necrosis, vein necrosis, stem necrosis.
  - ✓ First observation: 4-7 days post-inoculation for observation of local necrotic lesions which can lead to cotyledon drop. After this date these necrosis can hardly be seen on fallen cotyledons.
  - ✓ Second observation: two weeks post-inoculation for observation of symptoms of susceptibility.
  - ✓ Final observations. Susceptibility: mosaic (green), growth reduction. Resistance: local necrotic lesions which can lead to cotyledon drop, systemic necrosis.

Validation of test: evaluation of variety resistance was calibrated with results of resistant and susceptible controls.

Potyviruses

- Source of inoculum: Geves (FR).
- Isolate: for PVY: 0 strain 571, PVY Pathotype 1, PVY Pathotype 2.
- Test facility: climate room.
- Temperature: 18-25°C.
- Light: 12 hours.
- Inoculation:
  - ✓ preparation of inoculum: 1 g leaf with symptoms with 4 mL PBS (Phosphate Buffered Saline) with carborundum (80mg) and activated carbon (80mg);
  - ✓ plant stage at inoculation: cotyledon stage;
  - ✓ inoculation method: rubbing with the virus suspension;
  - ✓ final observations: three weeks post-inoculation.
- Observations:
  - ✓ method: visual;
  - ✓ observation scale: susceptibility (mosaic, growth reduction, vein banding and vein necrosis), resistance (no symptoms).

***Cucumber mosaic viruses* (CMV)**

- Strain: Sp-u1 (Wageningen).
- Growing conditions: 22°C, 12 hours of light.
- Growing method: raising of plants in climatized room.
- Inoculation:
  - ✓ preparation of inoculum: crushing of 1g of fresh leaves of Vinca rosea in 4 ml of Phosphate buffer 0.03M pH 7 + DIECA (diethyl dithiocaremate de sodium) (1 for 1000) + 300 mg of activated carbon + 80 mg of carborundum;
  - ✓ plant stage at inoculation: young plants at the stage of developed cotyledons. First leaf non-pointing;
  - ✓ method of inoculation: mechanical rubbing of cotyledons with a virus solution, the plants are kept in darkness for 48 hours;
  - ✓ duration of test: from sowing to inoculation 12 to 13 days.
  - ✓ from inoculation to reading: 3 readings at 10, 15 and 21 days after inoculation.

- Observations:
  - ✓ method: visual;
  - ✓ observation scale: susceptibility (growth reduction, vein banding and vein necrosis), resistance (no symptoms).

Tomato spotted wilt virus (tswv:0)

- Maintenance of pathotypes: freezing at -70C°.
- Execution of test: growth stage of the plants (one or two leaves expanded).
- Temperature: day, 20°C, night, 20°C.
- Growing method: glasshouse.
- Inoculation medium: 0.01 M PBS (Phosphate Buffered Saline) with 0.1% sodium sulfite freshly added.
- Method of inoculation: mechanical, rubbing with carborundum on cotyledons.
- Special conditions: keep inoculum suspension cool during inoculation.
- Duration of test: 20 days from sowing to inoculation, 14 to 20 days from inoculation to reading.
- Remarks: resistance will break down when temperature is higher than 25°C.

Bacterial spot

- Pathogen: *Xanthomonas campestris* pv. *Vesicatoria (Xcv).*
- Execution of test: growth stage of plants 6th to 8th true leaves.
- Temperature: 24°C night, 25°C day.
- Relative humidity: 80%.
- Light: 30 000 lx, day length 16 hours.
- Method of inoculation: infiltration into abaxial surface of a leaf in 13-15 mm diameter spots.
- Duration of the test: 10-14 days.

Figures 25-26. Classical symptoms comparison between resistant and susceptible genotypes for *Tabacco* M*osaic Virus* (left) and T*omato spotted wilt virus* (right)

Powdery mildew

- Pathogen: *Leveillula taurica.*
- Inoculum preparation: white powdery mildew spores collected from naturally-infected plants were used to inoculate the abaxial sides of leaves via the dropping method.
- Disease evaluation (Fig. 27): presence or absence of white fungal hyphae on infected leaves at 60 days after seedling transplant was used as a measure of disease infection. Powdery mildew infection was scored by the appearance of mycelia growth on the leaf surface of plants. Susceptible genotypes was also used as check. Intensity of white fungal hyphae observed on infected leaves (lower face) was visually scored using following disease scale: 1 = no sign of disease; 2 = minute necrotic lesions with no detectable sporulation; 3 = few large sporulation lesions; 4 = numerous large sporulation lesions.

Figure 27. Comparison between susceptible (left) and resistant (right) genotypes using artificial inoculation condition for selection to *Leveillula taurica* (Hungary)

Root-knot nematodes and *Phytophthora capsici*

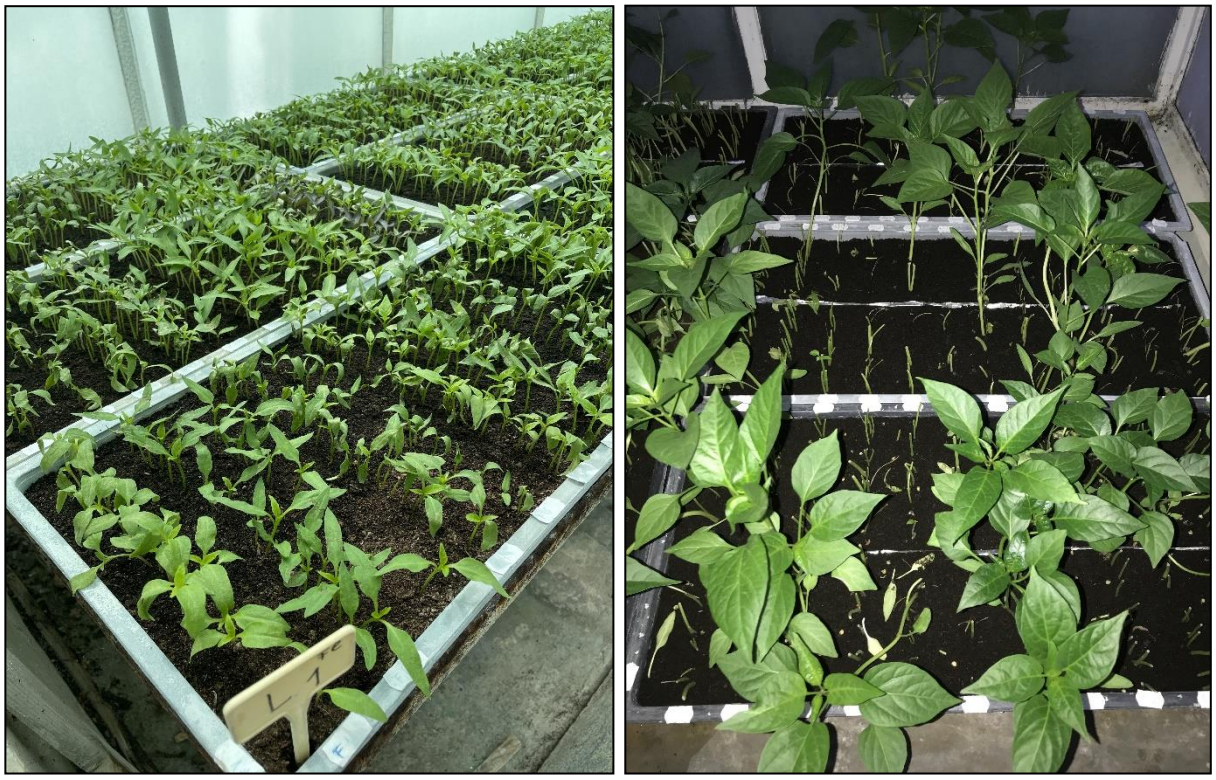
Evaluation of genotypes for resistance/susceptibility to *Meloidogyne* and *Phytophthora capsici* was conducted using infected soil with such pathogens within the company facilities. The root system of each plants were examined through visual score at the end of crop cycle. RKN resistance phenotype was evaluated using a root galling score.

According to the assessment, the plants with a score 0-1 were considered as resistant, with a score of 2-3 as medium sensitive while 4-5 as susceptible. The scale was arranged as follow: 0= no gall, 1= presence of 1-2 galls, 2= presence of 3-10 galls, 3= presence of 11-30 galls, 4= presence of 31-100 galls, 5= presence of more than 100 galls.

Evaluation of *Phytophthora capsici* was assessed using following scale: 0 = no visible symptoms; 1 = light brownish lesions on stems, leaves not wilted or slightly wilted; 2 = stem lesions extended to 1–2 cm in length, leaves wilted irrecoverably; 3 = stem lesions extended to be larger than 2 cm, leaves wilted or defoliated; 4 = long, brownish lesions on stems

extended, all leaves defoliated, plant almost dead; 5 = plant dead). Plants with a score of 0 or 1 were considered resistant while the plants with an higher score were considered susceptible.

For segregant material, laboratory activity concerned also recovering resistant plant survivors in order to advance and maintain interested genes carriers (Fig. 28 and Fig. 29).
Further activities were developed to broaden the genetic base of the collection. Pre-breeding crosses for the introgression of related resistances genes in new segregant populations was performed in greenhouse condition (figures 30-31).



Figures 28-29. Recovering process of survivor plants to *Cucumber Mosaic Virus* (CMV) inoculation

Figures 30-31. New pre-breeding cross combinations

3.2.2 Genetic investigation using KASP genotyping technology

Seed of the collection were sent to external company (Incotec group BV) for DNA extraction and subsequent analysis with Dna UniSNP kit.

The 48 SNP markers used for genotyping were chosen for their effectiveness on a worldwide germplasm as they were derived from a vast initial set and validated, from the service provider, on varieties of different origins whether the background is closely related or very different.

The wide applicability can be summarized as follows:

• reliable genotypic profile identification: the selected set is large enough to distinguish between an unlimited number of individuals;

• selected SNPs from this set - identified for each specific variety - can be use to perform hybrid purity testing whereby both inbreds and off-types can be identified;

• purifying parental lines: save years of work by cutting out several generations. This can be done by searching for homozygous positions and disregarding heterozygous ones to generate homogeneous parental lines;

• selection of homozygous plants: for backcrossing or other kinds of breeding.

Founded on 48 SNP detection via Kompetitive Allele Specific PCR (KASP), the genotyping assays are based on competitive allele-specific PCR, which enables bi-allelic scoring of single nucleotide polymorphisms (SNPs) and insertions and deletions (Indels) at specific loci.
The KASP method introduces fluorescence resonance energy transfer (FRET) for signal generation, where two fluorescent cassettes are used for the identification of allele-specific amplification for a single bi-allelic SNP.
The process begins with the first round of the PCR, where one of the allele-specific primers matches the target SNP and, with the common reverse primer, amplifies the target region. The common reverse primer also binds and PCR proceeds with the allele specific primer becoming incorporated into the template.
Competitive allele-specific PCR achieves bi-allelic discrimination through the competitive binding of the two allele-specific forward primers. If the genotype at a given SNP is homozygous, only one of the two possible fluorescent signals will be generated.
If the genotype is heterozygous, a mixed fluorescent signal will be generated (LGC, biosearch technologies, https://www.biosearchtech.com/).

Cluster analysis

For the construction of the similarity tree, the software PAST (PAleontological STatistics) ver. 4.03 (https://www.nhm.uio.no/english/research/resources/past/) was used.
The relationships between individuals of the germplasm collection within the dendrogram were determined using the Unweighted Group Arithmetic Means (UPGMA) algorithm.
The choice of the genetic similarity coefficient to be used was made by testing 3 different indices: Gower, Euclidean and Correlation.
The difference evaluated in terms of representation of the observed variability was determined by the use of Euclidean index both for the creation of the dendrogram and for the PCO (Principal Coordinates) graph.

Principal Coordinate Analysis - PCoA

Principal coordinate analysis, also carried out using the PAST software, is a sorting technique based on a similarity index that can be used to describe the genetic variability of a population.
Like cluster analysis, PCoA also starts from a similarity matrix, whose overall variability is broken down into a defined series of latent variables.
The first component found describes the largest percentage of the total variability found.

The second component has two important characteristics:

• summarize most of the variability present in the matrix, left unresolved by the first component;
• is not related to the first component.

The PCoA proceeds with new components, where each progressively accounts for an increasingly smaller part of the remaining variability.
Finally, the components can thus be correlated to known or presumed stratification sources for the species of interest.

3.3 Results

3.3.1 Collection set-up

A definitive panel consisting on 185 genotypes was set up and listed in Tab. 5. Seeds of germplasm collection were recovered from breeding fields placed in different country (Spain, Italy, Hungary).
Considering the aims for subsequent activities, genotypes choice was made creating individuals groups related to pathogen resistances and taking care to keep for each of them both resistant and susceptible genotypes.
Entirely composed by *Capsicum annum*, peppers types which constitute the collection, are extremely variable.
All the main segments cultivated in the Mediterranean area are represented such as: blocky, lamuyo, white type, bull horn, kapia, charleston, sivri, topepo and hot cayenna.

Pathogen	Species	Id	Placement	Generation	Allele	Phenotype
Leveillula taurica (LT)	C.annum	682	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	683	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	684	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	686	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	687	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	690	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	691	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	695	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	698	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	701	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	724	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	1287	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	1288	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	1289	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	1290	SP / alm	F4	rr	resistant
Leveillula taurica (LT)	C.annum	789	IT / gh-e	F5	rr	resistant
Leveillula taurica (LT)	C.annum	790	IT / gh-e	F5	rr	resistant
Leveillula taurica (LT)	C.annum	1270	IT / tun-2	F5	rr	resistant
Leveillula taurica (LT)	C.annum	1291	IT / tun-2	F5	rr	resistant
Leveillula taurica (LT)	C.annum	1271	IT / tun-2	F6	rr	resistant
Leveillula taurica (LT)	C.annum	1293	IT / tun-2	F6	rr	resistant
Leveillula taurica (LT)	C.annum	1294	IT / tun-2	F7	rr	resistant
Leveillula taurica (LT)	C.annum	1251	IT / tun-2	F10	rr	resistant
Leveillula taurica (LT)	C.annum	1252	IT / tun-2	F10	rr	resistant
Leveillula taurica (LT)	C.annum	1253	IT / tun-2	F10	rr	resistant
Leveillula taurica (LT)	C.annum	1255	IT / tun-2	F11	rr	resistant
Leveillula taurica (LT)	C.annum	1257	IT / tun-2	F11	rr	resistant
Leveillula taurica (LT)	C.annum	1258	IT / tun-2	F11	rr	resistant
Leveillula taurica (LT)	C.annum	1259	IT / tun-2	F12	rr	resistant
Leveillula taurica (LT)	C.annum	1260	IT / tun-2	F12	rr	resistant
Leveillula taurica (LT)	C.annum	688	IT / gh-e	F4	ss	susceptible
Leveillula taurica (LT)	C.annum	692	IT / gh-e	F4	ss	susceptible
Leveillula taurica (LT)	C.annum	693	IT / gh-e	F4	ss	susceptible
Leveillula taurica (LT)	C.annum	721	IT / gh-e	F4	ss	susceptible
Leveillula taurica (LT)	C.annum	1391	IT / tun-2	F5	ss	susceptible
Leveillula taurica (LT)	C.annum	1383	IT / tun-2	F10	ss	susceptible
Leveillula taurica (LT)	C.annum	1384	IT / tun-2	F10	ss	susceptible
Leveillula taurica (LT)	C.annum	1385	IT / tun-2	F10	ss	susceptible
Cucumber mosaic virus (CMV)	C.annum	1350	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1351	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1352	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1353	IT / tun-2	F5	rr	resistant

Cucumber mosaic virus (CMV)	C.annum	1354	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1355	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1356	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1357	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1358	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1359	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1360	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1362	IT / tun-2	F5	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1363	IT / tun-2	F6	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1365	IT / tun-2	F6	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1366	IT / tun-2	F6	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1371	IT / tun-2	F7	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1372	IT / tun-2	F7	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1373	IT / tun-2	F7	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1374	IT / tun-2	F7	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1376	IT / tun-2	F7	rr	resistant
Cucumber mosaic virus (CMV)	C.annum	1126	IT / gh-c	F1	ss	susceptible
Cucumber mosaic virus (CMV)	C.annum	1127	IT / gh-c	F1	ss	susceptible
Nematodes (Ma,Mi,Mj)	C.annum	1530	IT / tun-1	F1	unknow	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1531	IT / tun-1	F1	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1532	IT / tun-1	F1	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1513	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1514	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1515	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1516	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1517	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1518	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1521	IT / tun-1	F5	rr	resistant
Nematodes (Ma,Mi,Mj)	C.annum	1129	IT / gh-e	F1	ss	susceptible
Nematodes (Ma,Mi,Mj)	C.annum	1039	IT / gh-e	F1	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	1047	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-2)	C.annum	1057	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-2)	C.annum	1083	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-2)	C.annum	1017	IT / gh-c	F1	rr	resistant
Tobamovirus (tm:0-2)	C.annum	1017	IT / gh-c	F1	rr	resistant
Tobamovirus (tm:0-2)	C.annum	17	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	23	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	24	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	28	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	34	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	162	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	371	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	417	IT / gh-e	F10	rr	resistant

Tobamovirus (tm:0-2)	C.annum	494	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	498	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-2)	C.annum	1006	IT / gh-c	F1	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	1039	IT / gh-c	F1	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	5	IT / gh-e	F7	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	7	IT / gh-e	F7	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	42	IT / gh-e	F7	ss	susceptible
Tobamovirus (tm:0-2)	C.annum	484	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	1129	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-3)	C.annum	1133	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-3)	C.annum	1134	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-3)	C.annum	1135	IT / gh-c	F1	h	resistant
Tobamovirus (tm:0-3)	C.annum	24	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-3)	C.annum	28	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-3)	C.annum	34	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-3)	C.annum	162	IT / gh-e	F10	rr	resistant
Tobamovirus (tm:0-3)	C.annum	1039	IT / gh-c	F1	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	1043	IT / gh-c	F1	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	1055	IT / gh-c	F1	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	276	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	387	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	464	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	492	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	494	IT / gh-e	F10	ss	susceptible
Tobamovirus (tm:0-3)	C.annum	496	IT / gh-e	F10	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	1008	IT / gh-c	F1	h	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	1081	IT / gh-c	F1	h	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	1082	IT / gh-c	F1	h	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	1129	IT / gh-c	F1	h	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	9	IT / gh-e	F7	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	317	IT / gh-e	F7	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	323	IT / gh-e	F7	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	525	IT / gh-e	F7	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	17	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	19	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	23	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	311	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	371	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	417	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	503	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	515	IT / gh-e	F10	rr	resistant
Tomato spotted wilt virus (tswv:0)	C.annum	1039	IT / gh-e	F1	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	15	IT / gh-e	F7	ss	susceptible

Tomato spotted wilt virus (tswv:0)	C.annum	216	IT / gh-e	F7	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	224	IT / gh-e	F7	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	230	IT / gh-e	F7	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	234	IT / gh-e	F7	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	308	IT / gh-e	F7	ss	susceptible
Tomato spotted wilt virus (tswv:0)	C.annum	482	IT / gh-e	F10	ss	susceptible
Phytophthora capsici (Pc)	C.annum	1530	IT / tun-1	F1	unknow	resistant
Phytophthora capsici (Pc)	C.annum	1531	IT / tun-1	F1	rr	resistant
Phytophthora capsici (Pc)	C.annum	1039	IT / gh-c	F1	ss	susceptible
Phytophthora capsici (Pc)	C.annum	1129	IT / gh-e	F1	ss	susceptible
Bacterial spot (Xcv)	C.annum	1004	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	1005	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	1006	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	1007	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	1027	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	1037	IT / gh-e	F1	unknow	resistant
Bacterial spot (Xcv)	C.annum	2001	HU / Sze	F5	rr	resistant
Bacterial spot (Xcv)	C.annum	2002	HU / Sze	F5	rr	resistant
Bacterial spot (Xcv)	C.annum	2003	HU / Sze	F5	rr	resistant
Bacterial spot (Xcv)	C.annum	2004	HU / Sze	F5	rr	resistant
Bacterial spot (Xcv)	C.annum	2005	HU / Sze	F7	rr	resistant
Bacterial spot (Xcv)	C.annum	2006	HU / Sze	F7	rr	resistant
Bacterial spot (Xcv)	C.annum	2007	HU / Sze	F6	rr	resistant
Bacterial spot (Xcv)	C.annum	2008	HU / Sze	F6	rr	resistant
Bacterial spot (Xcv)	C.annum	2009	HU / Sze	F7	rr	resistant
Bacterial spot (Xcv)	C.annum	2010	HU / Sze	F7	rr	resistant
Bacterial spot (Xcv)	C.annum	1008	IT / gh-e	F1	ss	susceptible
Bacterial spot (Xcv)	C.annum	1039	IT / gh-c	F1	ss	susceptible
Bacterial spot (Xcv)	C.annum	2011	HU / Sze	F5	ss	susceptible
Bacterial spot (Xcv)	C.annum	2012	HU / Sze	F5	ss	susceptible
Bacterial spot (Xcv)	C.annum	2013	HU / Sze	F5	ss	susceptible
Bacterial spot (Xcv)	C.annum	2014	HU / Sze	F5	ss	susceptible
Bacterial spot (Xcv)	C.annum	2015	HU / Sze	F5	ss	susceptible
Bacterial spot (Xcv)	C.annum	2016	HU / Sze	F5	ss	susceptible
Potyvirus (pvy 0,1,1-2)	C.annum	1001	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1031	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1033	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1034	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1035	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1054	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1079	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1080	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1128	IT / gh-e	F1	unknow	resistant

Potyvirus (pvy 0,1,1-2)	C.annum	1206	IT / gh-e	F1	unknow	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	2000/1	HU / Sze	F10	rr	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	2000/2	HU / Sze	F10	rr	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	2000/3	HU / Sze	F10	rr	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	2000/4	HU / Sze	F10	rr	resistant
Potyvirus (pvy 0,1,1-2)	C.annum	1004	IT / gh-e	F1	ss	susceptible
Potyvirus (pvy 0,1,1-2)	C.annum	1008	IT / gh-e	F1	ss	susceptible
Potyvirus (pvy 0,1,1-2)	C.annum	1027	IT / gh-e	F1	ss	susceptible
Potyvirus (pvy 0,1,1-2)	C.annum	1039	IT / gh-c	F1	ss	susceptible

Table 5. Pepper collection list grouped for pathogen resistances classes. Id= progressive identifier number; Placement= breeding field hosting the genotype (IT-Italy, HU-Hungary, SP-Spain) and location reference (e.g. alm-Almeria, Sze-Szentes, gh/tun-greenhouse/tunnel); Generation= selection cycle (F1 to F12); Allele= homozygous (resistant, rr; susceptible, ss), heterozygous (h), unknow data are related to hybrids used as check (known phenotypic resistance but unknow allele due to missing information related to parental lines); Phenotype= resistance or susceptibility against each pathogen group determined using bioassay previously described (paragraph 3.2.1)

3.3.2 Profiles determination and SNP informativeness

To investigate the compostion of the collection and the homozigosity level, a representative panel consisting in 128 genotypes was characterized using the Kasp technology.

Genotypic information for 48 SNP are shown in the following Tab. 6.

Achieved genotypic profiles allowed the evaluation of the similarity level using PAST software (Fig. 32).

Id	Generation	Pepper type	SNP_011	SNP_012	SNP_013	SNP_014	SNP_021	SNP_022	SNP_023	SNP_024	SNP_025	SNP_031	SNP_032	SNP_033	SNP_041	SNP_042	SNP_043	SNP_044	SNP_051	SNP_052	SNP_053	SNP_061	SNP_062	SNP_063	SNP_071	SNP_072	SNP_073	SNP_074	SNP_075	SNP_076	SNP_081	SNP_082	SNP_091	SNP_092	SNP_101	SNP_102	SNP_103	SNP_104	SNP_111	SNP_112	SNP_113	SNP_114	SNP_121	SNP_122	SNP_123	SNP_124	SNP_125	SNP_126	SNP_127	SNP_128
1001	F1	blocky white	0	1	1	1	2	0	2	2	1	1	2	1	0	2	2	0	0	0	0	0	1	1	1	1	2	2	1	0	0	2	2	1	1	1	0	0	1	2	0	2	0	0	0	0	2	1	0	-
1004	F1	blocky white	1	0	1	1	2	2	2	2	1	0	0	0	0	0	1	2	0	1	0	0	1	2	1	2	1	0	2	2	0	2	2	1	1	1	0	2	1	0	2	2	2	2	2	0	0	2	2	2
1005	F1	blocky white	1	0	1	1	1	0	2	2	1	1	0	0	0	0	0	1	0	1	0	2	1	1	1	2	2	0	1	1	0	1	2	1	0	1	1	0	1	-	0	2	1	1	0	0	0	1	0	1
1006	F1	blocky white	2	0	1	1	2	2	2	0	1	2	0	0	0	2	2	2	0	1	0	2	1	2	1	2	2	2	2	2	0	0	1	-	1	1	0	2	1	2	2	2	2	2	2	2	0	1	1	0
1007	F1	blocky white	2	0	1	1	2	0	2	2	1	1	0	0	0	0	2	2	0	1	0	0	1	0	1	2	0	0	1	0	0	1	1	1	1	1	0	2	1	0	2	1	0	0	0	2	0	1	2	1
1008	F1	blocky white	2	2	1	1	2	2	2	2	2	1	2	2	0	0	2	0	0	2	0	2	1	1	2	2	2	0	1	2	0	1	2	1	■	1	■	2	1	0	0	2	0	0	0	0	0	1	0	1
1017	F1	conical white	1	2	1	2	2	2	1	2	1	2	0	0	0	0	1	0	0	2	0	0	1	0	1	0	1	1	1	0	0	1	-	1	2	1	0	0	1	2	1	1	1	1	1	1	1	1	1	0
1027	F1	lasthochka	2	2	0	0	1	0	2	0	1	2	2	0	0	0	1	2	2	1	0	0	1	2	2	0	2	0	1	0	0	2	2	2	2	2	2	0	1	2	2	1	2	2	1	1	0	1	1	2
1031	F1	cayenna	1	1	1	0	1	2	0	2	2	2	2	0	2	0	2	1	1	1	0	2	0	1	1	2	2	2	1	0	0	1	0	2	1	2	2	0	1	2	2	2	1	1	2	1	2	2	2	1
1033	F1	cayenna	2	1	1	0	1	0	0	1	0	0	0	0	1	2	0	1	2	1	2	2	0	1	1	0	1	0	1	2	0	2	0	0	1	0	1	0	1	2	2	0	1	1	1	1	2	2	0	1
1034	F1	cayenna	1	1	1	0	1	2	0	2	0	0	2	0	2	2	0	1	2	1	2	0	0	1	1	0	2	2	1	2	0	0	2	0	1	1	2	1	2	1	1	2	2	1	1	1	1	1	2	2	1
1035	F1	cayenna	2	1	1	0	1	1	2	2	2	2	0	2	2	2	0	2	0	1	2	2	0	0	1	0	1	0	1	2	0	0	0	-	2	2	1	0	-	2	2	2	1	2	1	1	-	2	0	1
1037	F1	jalapeno	1	2	2	0	0	0	1	2	2	1	2	1	1	1	0	1	2	1	0	1	0	2	1	1	0	1	2	1	0	2	2	2	2	1	2	2	1	2	2	2	1	1	1	1	1	0	0	2
1039	F1	charleston	2	2	1	1	2	2	1	0	1	2	0	1	0	0	1	2	2	1	1	0	2	1	1	2	2	2	1	2	0	0	2	2	1	0	0	0	1	2	0	0	1	-	1	1	0	1	2	1
1043	F1	charleston	2	2	2	1	1	2	1	2	1	2	0	1	0	0	1	2	2	1	1	0	0	2	2	2	1	2	1	2	-	0	1	2	1	0	0	0	1	-	0	-	1	2	1	1	0	1	2	2
1047	F1	charleston	0	2	1	0	1	0	1	2	1	2	2	0	0	0	1	1	2	1	2	2	1	2	1	2	1	2	1	0	0	2	1	0	0	1	2	0	1	0	2	2	1	2	2	1	-	1	2	2
1054	F1	horn	0	0	1	0	0	1	1	2	1	2	1	0	1	0	1	0	2	1	0	1	2	1	2	2	1	2	1	0	0	2	1	2	0	2	1	0	2	2	1	2	0	2	0	1	1	1	2	1
1055	F1	horn	0	0	1	2	2	0	1	0	1	1	1	2	0	2	1	1	1	1	2	2	1	2	0	2	0	0	2	2	0	2	2	2	2	0	2	0	1	0	0	0	2	1	2	2	0	-	2	1
1057	F1	horn	0	0	2	1	2	2	1	0	2	1	0	0	2	0	1	1	1	1	2	0	1	0	0	2	1	2	1	0	0	2	2	-	0	2	1	0	1	0	2	2	2	2	2	1	0	-	0	1
1079	F1	lamuyo	0	0	0	0	2	1	2	2	1	0	2	2	1	1	1	0	2	1	2	2	1	1	2	1	1	0	1	2	0	0	0	0	0	2	1	0	1	1	1	0	0	0	0	1	1	1	2	1
1080	F1	lamuyo	2	1	0	1	1	0	2	2	1	2	2	2	0	0	2	-	0	1	0	2	2	1	0	1	1	0	1	2	0	0	■	0	2	1	1	0	2	1	2	0	2	2	2	2	2	1	2	1
1081	F1	lamuyo	0	2	0	2	1	2	1	2	1	2	2	0	0	2	1	0	1	1	0	1	1	1	0	2	1	0	1	2	0	1	2	0	0	2	1	1	1	2	2	0	0	0	0	1	1	2	2	1
1082	F1	lamuyo	0	2	2	0	1	1	1	1	1	2	-	0	0	2	1	0	1	2	0	2	2	1	0	1	1	0	1	0	0	0	2	0	0	2	1	0	0	1	2	0	0	0	2	1	1	1	0	1
1126	F1	lamuyo	2	2	0	2	2	1	2	2	1	0	2	2	0	1	1	0	2	1	0	2	1	0	1	2	1	0	1	2	0	2	2	0	2	1	2	2	1	1	1	0	0	0	0	2	-	-	0	1
1127	F1	lamuyo	2	2	0	2	2	1	1	1	1	0	0	1	0	1	1	-	1	1	0	2	2	0	0	1	1	0	1	0	0	0	2	1	2	2	1	1	0	2	1	1	2	2	2	0	2	0	1	2	1
1128	F1	lamuyo	2	2	0	0	2	1	1	2	1	2	2	2	1	2	2	0	1	1	0	2	1	2	2	1	1	0	1	0	0	0	2	0	0	2	1	0	1	1	1	1	0	0	0	1	2	1	2	1
1129	F1	lamuyo	2	0	-	2	0	1	1	0	1	0	1	2	0	0	1	0	1	1	0	2	1	2	2	2	1	0	1	1	0	2	2	0	0	2	1	1	1	1	1	2	0	2	1	1	2	2	0	1
1133	F1	lamuyo	0	2	0	1	2	1	2	2	1	0	0	2	0	2	1	0	0	1	0	0	0	2	0	2	1	0	1	2	0	1	1	-	2	1	1	2	2	1	1	-	2	2	0	1	0	0	0	1
1134	F1	lamuyo	2	0	0	2	2	2	1	2	1	2	2	0	0	2	2	0	2	2	0	0	2	1	2	2	2	2	0	1	2	0	1	2	2	0	2	1	1	1	2	1	0	0	0	2	2	-	1	0	1
1135	F1	lamuyo	0	0	0	2	2	1	1	2	1	0	2	2	0	0	1	0	2	1	0	2	2	2	1	2	1	0	1	2	0	2	2	2	2	1	1	2	1	1	1	2	0	0	0	2	0	1	0	1
1206	F1	cherry	0	0	0	1	0	1	2	2	0	2	1	0	0	0	2	-	0	1	1	2	2	0	2	2	2	0	2	2	0	2	0	2	1	2	1	2	2	2	1	0	1	1	1	1	0	0	2	1
1530	F1	cayenna	0	2	1	1	1	2	0	2	1	0	0	2	1	1	2	2	2	1	2	2	2	1	1	0	2	0	0	2	-	2	1	1	1	2	2	2	2	2	2	2	2	2	1	2	2	1	1	1
1531	F1	cayenna	0	1	2	2	1	2	0	2	1	0	2	0	2	2	1	1	1	1	2	2	1	2	1	0	1	0	2	2	0	2	2	-	2	1	1	2	2	1	1	1	2	2	1	2	0	1	1	1
1532	F1	jalapeno	1	2	1	2	2	2	2	2	2	1	2	2	2	1	2	1	1	1	2	2	2	0	2	1	1	1	1	2	0	2	2	0	2	2	2	2	2	2	1	2	2	2	2	1	2	-	0	1
682	F4	conical white	1	0	1	1	1	0	0	0	1	1	1	1	1	-	1	1	0	1	0	0	0	2	1	1	0	0	1	1	0	1	1	1	0	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	1
683	F4	conical white	1	2	1	0	0	0	2	-	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	0	0	0	-	1	1	0	0	1	0	1	2	1	2	0	1	2	1	-	1	1	1	-	0	■	1	1
684	F4	conical white	1	2	1	0	0	0	2	2	1	1	0	0	0	0	1	-	0	1	1	0	1	0	2	0	1	1	1	0	0	2	1	1	0	1	0	0	1	2	1	1	1	1	1	1	0	1	1	1
686	F4	conical white	1	0	1	0	0	0	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	1	0	0	-	0	1	-	1	-	0	1	0	1	1	1	1	1	1	0	-	1	2
687	F4	conical white	1	0	1	0	0	0	0	1	1	1	1	0	0	0	1	0	0	1	0	0	0	2	1	1	0	0	1	0	0	0	0	1	2	1	1	0	1	0	1	1	1	1	1	1	0	1	1	0
688	F4	conical white	1	0	1	0	0	0	1	0	1	1	1	0	0	0	1	0	0	1	2	0	0	0	0	0	2	2	1	0	0	0	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	1	0
690	F4	conical white	1	0	1	1	1	0	0	2	1	1	1	1	1	-	1	0	0	1	0	0	1	1	0	2	2	2	1	0	0	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	0	1	1	1
691	F4	conical white	1	0	1	1	1	0	0	1	1	1	1	1	1	-	1	2	0	1	0	0	1	1	0	1	0	0	1	0	0	0	1	1	-	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1
692	F4	conical white	1	0	1	1	0	0	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0	0	1	1	2	1	1	1	0	0	0	1	1	2	1	1	0	1	0	1	0	1	1	1	0	1	1	1
693	F4	conical white	1	0	1	0	0	0	2	1	1	1	1	0	0	0	1	0	0	1	2	0	0	1	2	0	2	2	1	0	0	2	1	1	1	1	0	0	1	0	1	1	1	1	1	1	0	1	1	0
695	F4	conical white	1	1	1	0	0	0	2	2	1	1	1	0	0	0	1	0	0	1	2	1	1	2	1	1	0	0	1	0	0	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	0	1	1	0
698	F4	conical white	1	1	0	2	2	0	0	0	1	1	1	0	-	0	0	1	2	1	1	0	1	0	1	2	0	0	1	0	0	2	-	1	1	1	2	0	1	0	1	0	1	1	1	1	0	1	1	1
701	F4	conical white	1	0	1	1	0	0	0	1	1	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	2	0	0	2	1	1	0	1	0	0	1	0	1	1	1	1	1	1	0	1	1	1
721	F4	conical white	1	0	1	1	1	2	2	0	1	0	1	0	0	0	1	0	0	1	1	0	1	1	0	0	1	1	1	0	0	1	1	2	0	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1
724	F4	lasthochka	1	1	1	0	1	0	-	1	1	2	1	2	0	0	1	0	0	1	0	0	0	1	2	0	0	1	0	1	0	0	2	1	1	0	1	-	0	1	2	0	0	1	1	1	1	0	1	1	1
1287	F4	sivri	0	2	2	1	1	1	2	1	0	1	0	1	0	-	2	1	1	1	1	1	1	0	1	1	1	1	1	-	0	1	1	-	0	0	0	0	0	1	0	1	2	-	1	1	0	1	0	0	0
1288	F4	kapia	0	0	0	1	0	1	0	0	1	0	0	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	1	1	0	1
1289	F4	conical white	1	0	1	0	0	0	1	0	1	1	0	1	1	-	0	1	0	1	2	1	0	0	1	0	0	1	0	1	0	1	0	0	1	1	0	1	-	0	1	0	0	0	1	1	1	0	1	1	0
1290	F4	conical white	1	1	1	0	0	0	0	0	1	1	0	1	0	1	-	1	0	0	1	0	0	0	1	0	0	1	1	1	0	0	2	1	1	0	1	1	0	1	0	1	1	1	1	1	0	-	1	-
789	F5	blocky white	1	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	1	0	0	1	0	0	0	0	1	0	-
790	F5	blocky white	1	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	1	0	0	1	0	0	0	0	1	0	1
1270	F5	cayenna	1	1	1	0	0	-	1	1	1	0	0	0	0	-	0	0	1	0	1	1	0	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	0	1	1
1291	F5	charleston	1	1	1	0	1	0	1	0	1	0	1	1	0	0	1	0	1	1	0	1	1	1	1	1	0	1	0	1	0	0	1	1	0	1	0	1	1	1	0	1	0	0	1	0	0	0	0	0
1350	F5	lasthochka	0	0	1	1	1	1	1	1	1	0	0	1	0	0	1	-	0	-	0	0	1	0	1	1	1	1	1	0	0	1	0	0	1	2	0	0	1	0	2	0	1	1	1	1	0	-	1	0
1351	F5	lasthochka	2	0	1	1	1	1	1	1	1	0	0	1	0	0	1	-	0	1	1	0	1	0	1	2	1	1	1	0	0	1	2	0	1	1	0	0	1	0	1	0	1	1	1	1	0	1	1	0
1352	F5	lasthochka	0	0	1	0	1	1	1	1	1	0	0	1	0	0	1	0	0	1	2	0	1	0	2	0	1	1	1	0	0	1	0	-	1	1	0	0	1	0	2	0	1	1	1	1	0	1	1	0
1353	F5	lasthochka	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	0	1	1	1	0	0	1	1	0	1	2	0	0	1	0	1	0	1	1	1	1	0	1	1	0
1354	F5	conical white	1	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	2	0	1	1	0	1	0	0	0	1	0	0	0	1	1	2	1	0	0	1	1	1	1	0	0	0	1	1	0	2	1
1355	F5	conical white	1	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	2	0	1	1	0	1	0	0	0	1	0	0	2	1	1	1	1	0	0	1	1	2	1	0	0	0	1	1	1	1	1

Id	Generation	Pepper type	SNP_011	SNP_012	SNP_013	SNP_014	SNP_021	SNP_022	SNP_023	SNP_024	SNP_025	SNP_031	SNP_032	SNP_033	SNP_041	SNP_042	SNP_043	SNP_044	SNP_051	SNP_052	SNP_053	SNP_061	SNP_062	SNP_063	SNP_071	SNP_072	SNP_073	SNP_074	SNP_075	SNP_076	SNP_081	SNP_082	SNP_091	SNP_092	SNP_101	SNP_102	SNP_103	SNP_104	SNP_111	SNP_112	SNP_113	SNP_114	SNP_121	SNP_122	SNP_123	SNP_124	SNP_125	SNP_126	SNP_127	SNP_128
1356	F5	conical white	1	0	1	1	1	1	1	2	1	1	0	2	0	0	0	0	0	2	0	1	1	2	1	0	0	0	1	0	0	2	1	1	1	1	0	0	1	1	2	2	1	1	1	1	1	1	1	1
1357	F5	conical white	1	0	1	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	0	1	0	0	-	1	1	2	1	0	0	1	2	1	1	0	0	-	1	0	0	0	1
1358	F5	conical white	1	0	1	1	1	0	1	1	1	1	1	0	0	0	1	0	0	1	0	0	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	0	0	1	0	0	0	1	1	1	1	0	1	1	1
1359	F5	conical white	1	0	1	1	1	0	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	1	1	1	1	0	0	2	1	0	1	0	0	0	1	0	0	1	1	1	1	1	0	1	1	1
1360	F5	conical white	1	0	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	0	0	1	1	1	-	2	0	0	1	0	1	0	1	1	1	1	0	1	1	1	1
1362	F5	conical white	1	0	1	0	1	1	1	1	1	0	-	0	0	0	1	0	0	1	0	0	1	0	0	2	1	1	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	-	1	1	0	1	1	1	1
1391	F5	charleston	1	1	1	0	2	0	1	0	1	0	1	1	0	0	1	0	1	1	1	1	1	0	1	0	1	1	1	0	0	1	1	0	1	0	1	0	1	1	0	1	0	-	1	0	0	0	0	0	0
1513	F5	kapia	0	0	0	0	1	-	0	0	1	0	2	2	0	0	1	1	1	1	1	0	2	0	2	0	1	1	1	1	0	1	0	1	0	1	1	0	1	0	0	1	1	1	1	1	1	0	0	0	1
1514	F5	kapia	0	0	0	0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	0	2	0	-	0	1	1	1	1	0	1	0	1	0	1	1	0	1	0	2	1	1	1	1	1	0	0	0	1
1515	F5	kapia	0	0	0	0	1	2	0	0	1	0	1	0	0	0	1	1	1	1	1	0	0	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0	0	0	1
1516	F5	kapia	0	0	0	0	1	2	0	0	1	0	1	0	0	0	1	1	1	1	1	0	0	0	2	0	1	0	1	0	0	1	0	1	0	1	1	0	1	2	1	1	1	1	1	1	0	0	0	1
1517	F5	cayenna	0	1	0	0	1	1	0	2	1	0	0	0	0	2	1	1	1	1	0	2	1	1	1	0	1	0	1	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1
1518	F5	kapia	0	1	0	0	1	1	0	0	1	0	0	0	2	2	1	1	1	1	1	0	1	1	1	0	1	0	1	0	0	0	0	0	0	1	1	1	0	2	1	1	1	1	1	1	0	1	1	1	1
1521	F5	lamuyo	0	1	0	1	1	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	1	0	0	2	0	1	1	0	1	1	1	1	0	-	0	1	0	1	0	0	2	2	0	0	1	0	1	0
1293	F6	sivri	0	1	1	1	0	1	1	1	1	1	0	2	0	-	-	1	1	1	1	1	0	0	0	0	1	1	1	2	0	1	1	1	0	1	0	0	1	0	1	1	1	0	1	2	0	0	0	0	0
1363	F6	conical white	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	0	1	0	0	1	0	1	0	1	0	0	1	1	0	0	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	1	1	1
1365	F6	conical white	1	0	1	1	1	0	2	0	1	1	1	0	0	-	1	0	0	1	0	0	1	0	1	-	0	1	1	0	0	1	1	1	0	1	0	0	1	0	2	0	1	1	1	1	-	1	1	1
1366	F6	conical white	1	-	1	1	1	0	-	-	1	1	1	0	0	0	1	0	0	1	-	0	1	0	1	0	0	1	1	0	0	1	-	1	0	1	0	-	1	0	2	2	1	1	1	1	0	1	1	1	1
5	F7	blocky	0	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	1	1	0	0	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	0	-	0	1	1
7	F7	blocky	2	2	1	1	0	0	1	1	1	0	1	0	0	0	1	0	0	1	0	0	1	1	0	0	1	0	1	1	-	0	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	-	0	0	1
9	F7	blocky	1	1	0	0	1	1	1	1	1	0	0	1	0	1	1	0	0	1	0	0	1	0	0	1	1	0	1	1	0	0	1	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	1	1
15	F7	blocky	0	1	0	1	0	0	1	0	1	1	0	1	0	0	0	0	0	0	1	0	-	-	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1	0	0	0	1	-	0	1	0	0	0	1
216	F7	lamuyo	0	0	0	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1	0	-	0	0	1	1	1	0	0	1	1	1	0	1	0	1	1	0	0	0	1
224	F7	lamuyo	1	0	0	0	0	1	1	0	1	0	1	0	0	0	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	1	1	1	1	0	0	0	1	1	1	1
230	F7	lamuyo	0	0	0	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	1	0	1	1	0	0	0	1
234	F7	lamuyo	0	0	0	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	1	-	1	1	0	0	0	1
308	F7	lamuyo	1	1	0	0	0	1	0	0	1	0	0	1	0	1	1	-	0	1	0	0	1	0	1	0	1	0	1	0	0	1	0	-	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	1
317	F7	lamuyo	0	0	0	0	1	1	1	1	1	0	1	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	0	1	0	0	1	0	0	0	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1
323	F7	lamuyo	0	0	0	0	1	1	1	1	1	0	1	0	0	1	1	0	0	1	0	0	1	0	1	0	1	1	1	0	1	0	0	0	1	0	0	0	1	1	1	0	1	1	0	0	0	0	0	1	1
525	F7	topepo	0	1	0	0	0	0	1	1	1	0	1	2	0	0	0	0	0	1	0	0	1	1	1	1	1	1	0	1	1	0	1	0	-	0	1	1	0	1	1	0	0	1	0	0	1	1	0	0	1
1294	F7	cayenna	0	2	1	1	1	2	1	0	1	2	2	2	2	-	2	1	2	1	1	2	2	2	1	2	0	0	2	0	0	1	2	1	2	1	1	0	1	-	2	0	2	2	1	2	-	1	1	-
1371	F7	conical white	1	0	1	1	1	1	1	1	1	1	0	1	0	0	1	0	0	1	0	0	1	-	1	1	1	1	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1
1372	F7	conical white	1	0	1	1	1	1	1	0	1	1	0	1	0	0	1	0	0	1	0	0	1	0	1	1	1	1	1	0	0	1	2	1	0	1	0	0	1	0	0	1	1	1	1	1	0	1	1	1	1
1373	F7	conical white	1	0	1	1	1	1	1	0	1	1	0	1	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	0	0	1	1	1	2	1	0	0	1	0	0	1	1	1	1	1	0	1	1	1	1
1374	F7	conical white	1	0	1	2	1	1	1	1	1	1	0	1	0	0	1	0	0	1	0	0	1	0	1	1	1	1	1	0	0	1	1	1	0	1	0	0	1	0	0	1	1	1	1	1	0	1	1	1	1
1376	F7	conical white	1	0	1	1	1	1	1	0	1	1	0	1	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	0	0	1	2	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	1	1	1
17	F10	blocky	0	0	1	1	0	1	1	1	1	0	0	1	0	1	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1
19	F10	blocky	0	0	1	1	0	1	1	1	1	0	0	1	0	1	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1
23	F10	blocky	0	0	1	1	0	1	1	1	1	0	0	1	0	1	1	0	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1
24	F10	blocky	0	0	0	1	1	0	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	1	0	0	1	1	1	1	0	1	0	0	0	0	0	1	0	1	1	1
162	F10	blocky	0	0	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	1	0	-	1	1	0	0	1	1	1	1	-	1	2	0	0	0	1	0	1	0	1	1
311	F10	lamuyo	0	0	0	0	1	1	1	1	1	0	1	0	0	1	1	0	0	1	0	0	1	0	1	1	1	0	1	0	0	1	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	1	0	1	1
371	F10	horn	1	0	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	1	0	0	0	1	1	1	0	1	1	0	1	0	0	0
387	F10	horn	0	0	0	1	0	1	0	0	1	0	1	1	0	0	1	-	0	1	1	1	1	1	1	0	1	0	1	1	0	1	2	1	0	1	1	0	1	1	1	0	0	-	1	1	1	1	0	-
417	F10	kapia	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	1	1	1	0	0	1	0	1	1	0	1	0	0	1	1	1	0	1	1	0	0	0	1	0	1	1	0	1	1	1	0
464	F10	horn	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	0	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1
482	F10	horn	0	0	0	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	1	0	1	1	1	1	0	0	1	0	0
484	F10	horn	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	1	0	1	0	1	0	1	0	0	0	0	1	1	1	0	1	1	1	1
492	F10	horn	0	0	1	0	0	1	1	0	1	0	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	1	1	-	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	1	0	1
494	F10	horn	0	0	1	0	0	1	1	0	1	1	1	0	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	0	-	1	1	1	0	0	1	0	0	1	1	1	0	1	0	1	1	1	1	0	1
496	F10	horn	0	0	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	1	1	0	1	0	0	0	1	0	1	0	-	1	0	0	0	0	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1
498	F10	kapia	0	0	0	0	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	0	1	0	1	0	1	1	1	-	0	-	-	0	1	1	1	0	1	0	1	0	1	-	1	-	0	-	0	0
503	F10	kapia	1	0	0	1	0	1	1	0	1	0	1	0	0	0	1	1	1	1	0	0	-	0	1	1	0	0	1	0	0	1	1	0	1	1	1	0	1	1	1	0	0	0	1	1	0	1	1	1	1
1251	F10	sivri	0	1	1	0	1	1	1	1	1	1	0	1	0	-	0	1	-	1	1	1	0	0	0	0	1	1	1	0	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0
1252	F10	lamuyo	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	0	1	1	0	0	1	0	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	1	0	0	0	1	0
1253	F10	kapia	1	0	0	1	1	1	1	1	0	1	1	0	0	1	-	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	0	0	0	1	0	1	1	1	1	1	0	0	1	1	0	0	1
1383	F10	horn	0	0	0	1	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	0	-	1	0	1	0	1	0	1	0	0	1	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0	1	0	1
1384	F10	kapia	0	0	0	1	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	0	1	1	0	1	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	0	0	1	0	1	1	0	1	0	-
1385	F10	horn	0	0	0	0	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	0	1	1	0	1	0	1	0	1	0	0	0	0	0	1	1	1	0	1	0	0	1	1	0	1	1	0	1	0	1
2000/1	F10	cayenna	0	1	1	1	1	1	2	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	2	0	1	1	1	1	1
1255	F11	kapia	1	0	0	0	1	1	1	0	1	1	0	0	1	-	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	1	0	-	1	1	0	-	1	0	1	0	0	1	1	0	0	-	0	1	1
1257	F11	kapia	1	0	0	0	1	1	1	0	1	1	0	0	1	-	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1
1258	F11	conical white	0	0	1	1	1	0	1	0	1	1	0	1	1	-	1	0	0	1	0	0	-	0	1	0	0	0	1	0	-	1	1	1	1	1	0	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1
1259	F12	sivri	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	0	1	0	0	1	1	0	1	0	0	1
1260	F12	sivri	0	1	1	1	1	1	1	1	1	1	0	0	0	-	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	1	0	0	1

Table 6. Genetic characterization through Kasp-48 SNP analysis. Allelic variants are highlighted by different color: allele A (id-0, blue color), allele B (id-1, green color), heterozygous (id- 2, yellow color), missing data (deep blue)

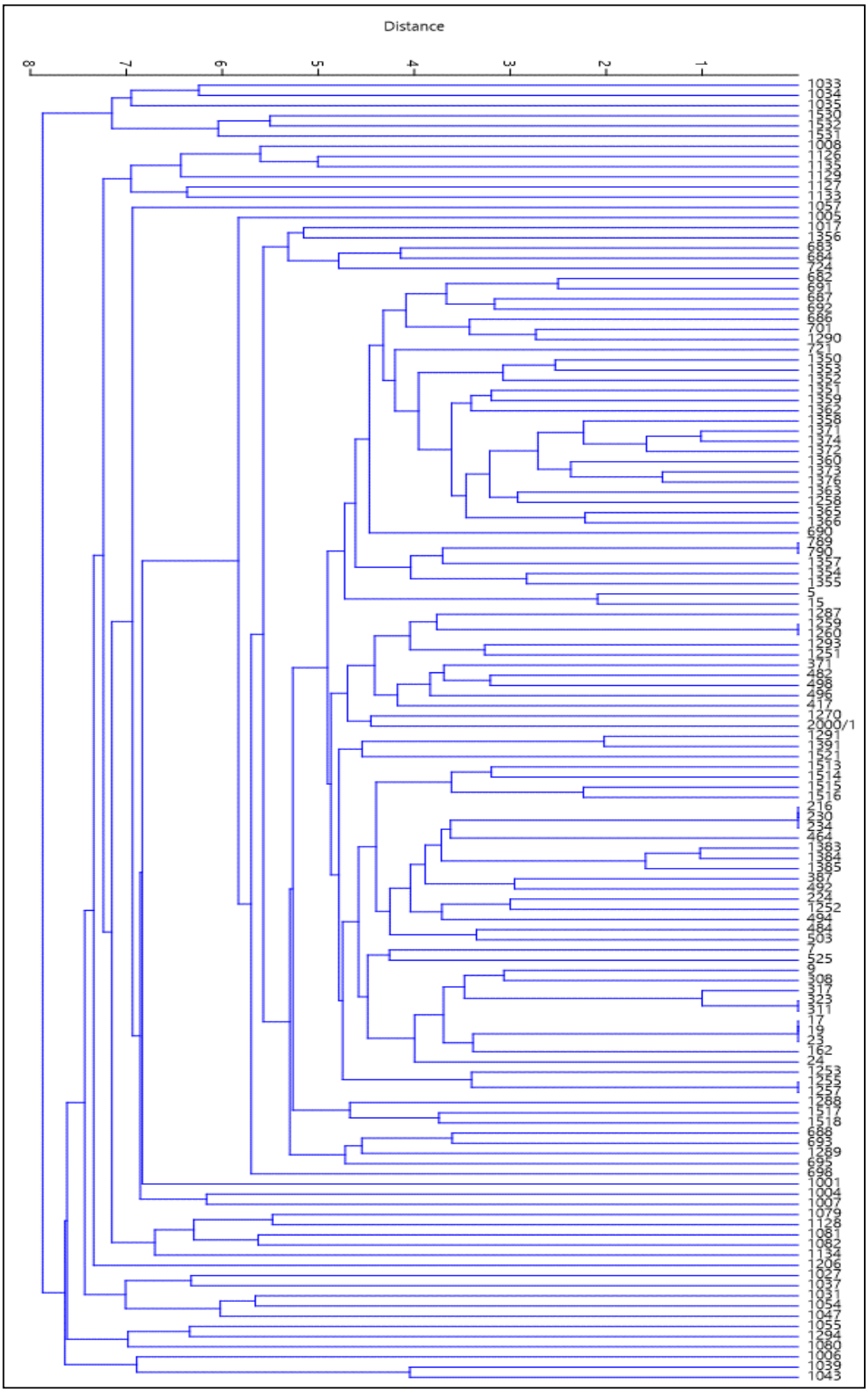


Figure 32. Genetic similarity tree of the pepper collection

After grouping genotypes for generation, molecular characterization allowed to analyze the fixity level of the collection. Confirming the inversely proportional trend between selection generation and locus heterozygosity, the following Tab. 7 shows the detected percentage.

Selection generation	Heterozygous locus (%)
F1	35,9
F4	6,8
F5	4,1
F6	3,9
F7	0,9
F10	0,4
F11	0,0
F12	0,0

Table 7. Ratio between selection generation and level of heterozygosity

Further investigation was done on the markers representativeness. The frequency of each allele/SNP used considering the mean of every genotype analyzed is reported below (tables 8-9).

	SPN_011	SPN_012	SPN_013	SPN_014	SPN_021	SPN_022	SPN_023	SPN_024	SPN_025	SPN_031	SPN_032	SPN_033	SPN_041	SPN_042	SPN_043	SPN_044	SPN_051	SPN_052	SPN_053	SPN_061	SPN_062	SPN_063	SPN_071	SPN_072
% allele 0	46	63	37	41	29	35	27	42	4	44	50	52	80	72	19	62	51	7	60	59	22	48	28	46
% allele 1	41	21	58	49	56	51	56	34	91	43	36	34	13	15	71	30	37	87	28	24	68	37	60	35
% allele 2	13	16	5	9	15	14	17	24	5	13	14	14	7	13	10	8	12	6	12	17	10	14	12	20

	SPN_073	SPN_074	SPN_075	SPN_076	SPN_081	SPN_082	SPN_091	SPN_092	SPN_101	SPN_102	SPN_103	SPN_104	SPN_111	SPN_112	SPN_113	SPN_114	SPN_121	SPN_122	SPN_123	SPN_124	SPN_125	SPN_126	SPN_127	SPN_128
% allele 0	28	60	4	67	100	21	28	31	44	15	42	75	10	43	26	44	34	37	27	17	78	22	51	17
% allele 1	60	30	90	16	0	58	54	59	40	71	50	15	83	41	56	40	55	49	65	73	16	72	37	78
% allele 2	12	9	6	17	0	22	19	10	15	14	8	10	7	16	18	16	10	14	9	10	7	6	13	5

Table 8-9. Markers representativeness (%) in relation to each allele

The investigation allowed furthermore the analysis of markers informativeness level by evaluating the allele frequency for each SNP. For instance, SNP 081 was never polymorphic while SNP 031 was perfectly distributed with both allele frequency at 50%. Tab. 10 shows the

identified coefficient (%) where low scale is referred to absence of polymorphisms while opposite pole shows equally allele distribution

SNP																																															
081	025	075	052	111	041	125	128	104	042	124	102	043	076	126	062	012	123	082	061	044	053	071	073	113	074	092	023	021	091	013	121	033	022	032	051	127	122	072	063	024	014	103	011	101	114	112	031
0	13	14	20	28	34	38	39	40	43	44	44	48	49	50	54	58	62	63	65	68	69	69	69	70	70	71	71	73	74	79	79	83	84	86	86	86	88	89	90	92	92	93	95	96	96	98	99

Table 10. SNP informativeness level (%)

3.3.3 Clustering and derivative correlation

Clustering investigation was conducted by relating the SNP profiles of the genotypes and several factors characterizing the collection. Correlation studies and PCoA analysis were focused on the following relationships:

- **resistance to pathogens (breeding lines F4 to F10)**

Every genotype of each resistance group was investigated through comparison of molecular pattern (48 SNP) and resistant phenotype in order to identify significant correlated polymorphisms. No positive correlation between SNP polymorphisms and resistant or susceptible phenotype was found for any class. Dispersion reported in Fig. 33 is an example of negative clustering analysis for *L. taurica* genotypes group.

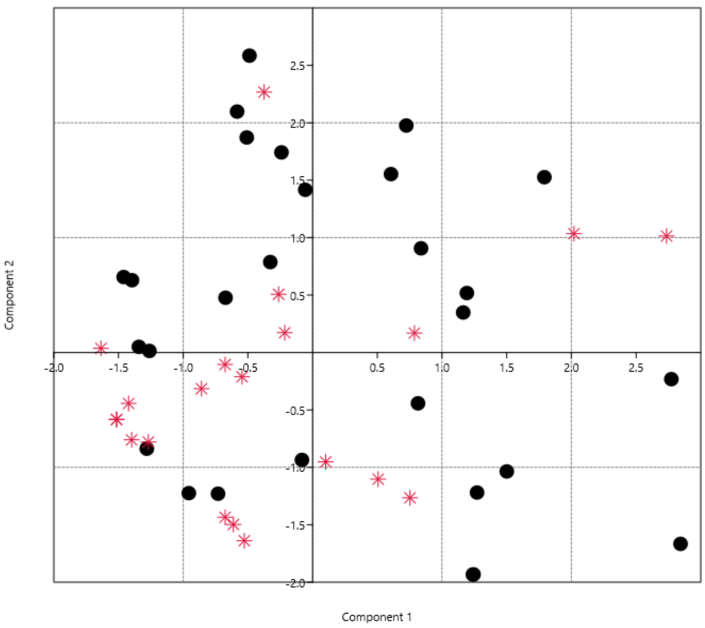


Figure 33. Clustering for *Leveillula taurica* genotypes group. Black point (resistant), star red (susceptible)

- **pepper segment type (breeding lines F4 to F10)**

Cluster data of the entire collection showed the identification of 3 macro-families with a positive correlation with fruit shape.

Results highlighted on graph below (Fig. 34) show the presence of main groups.

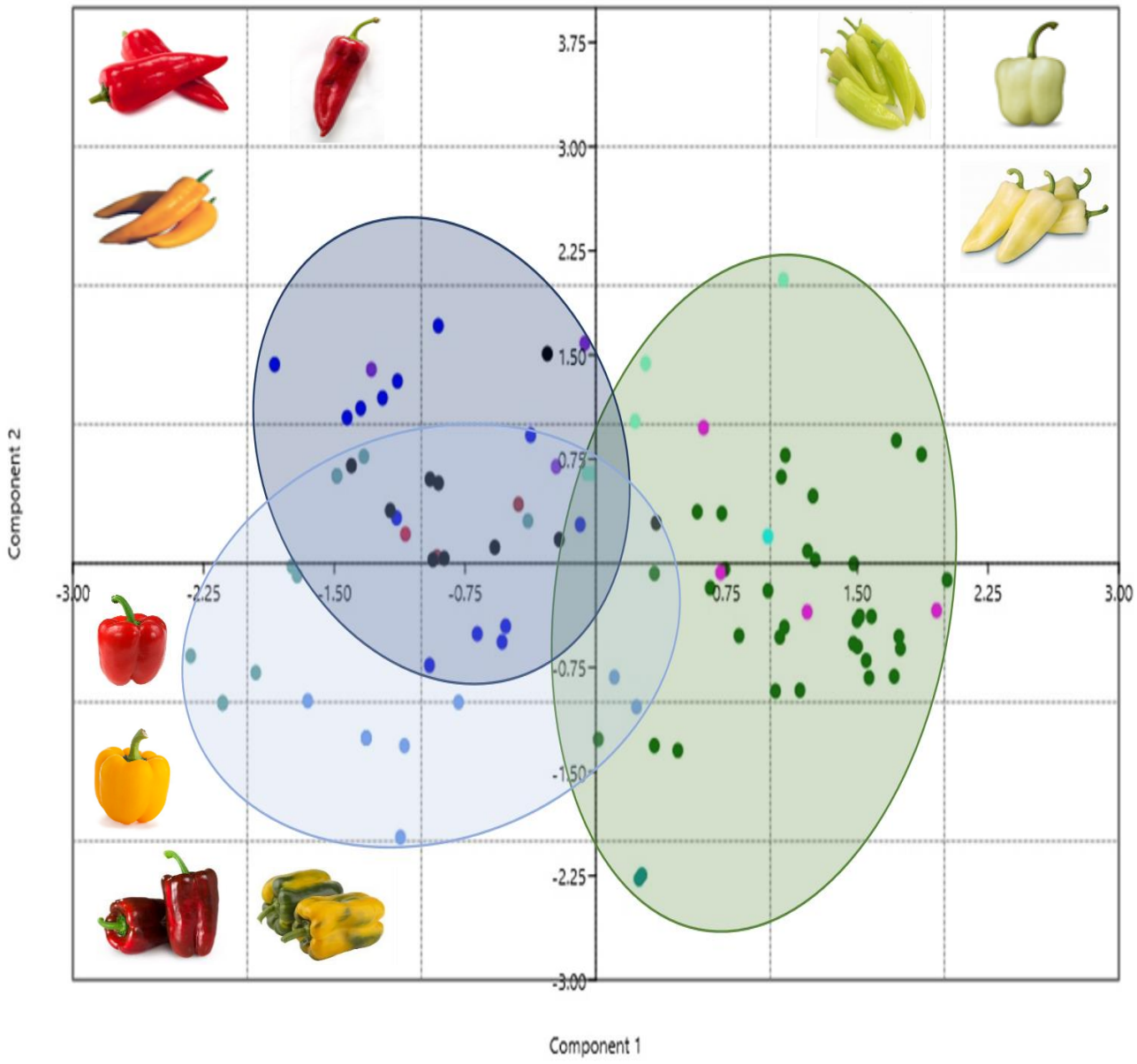


Figure 34. Positive clustering for fruit shape. Greenish/white type pepper, green and violet point (right side); conical red pepper (bull horn and kapia type), blue and black points (left up); square pepper (blocky and lamuyo), light blue (left down).

- **provenience area (breeding lines F4 to F10)**

Further analysis has been directed to the investigation between genetic profile and provenience area (market focus). Positive correlation has been found as in the previous clustering. Main families identified are molecularly distinguishable following origin market area (Fig. 35).

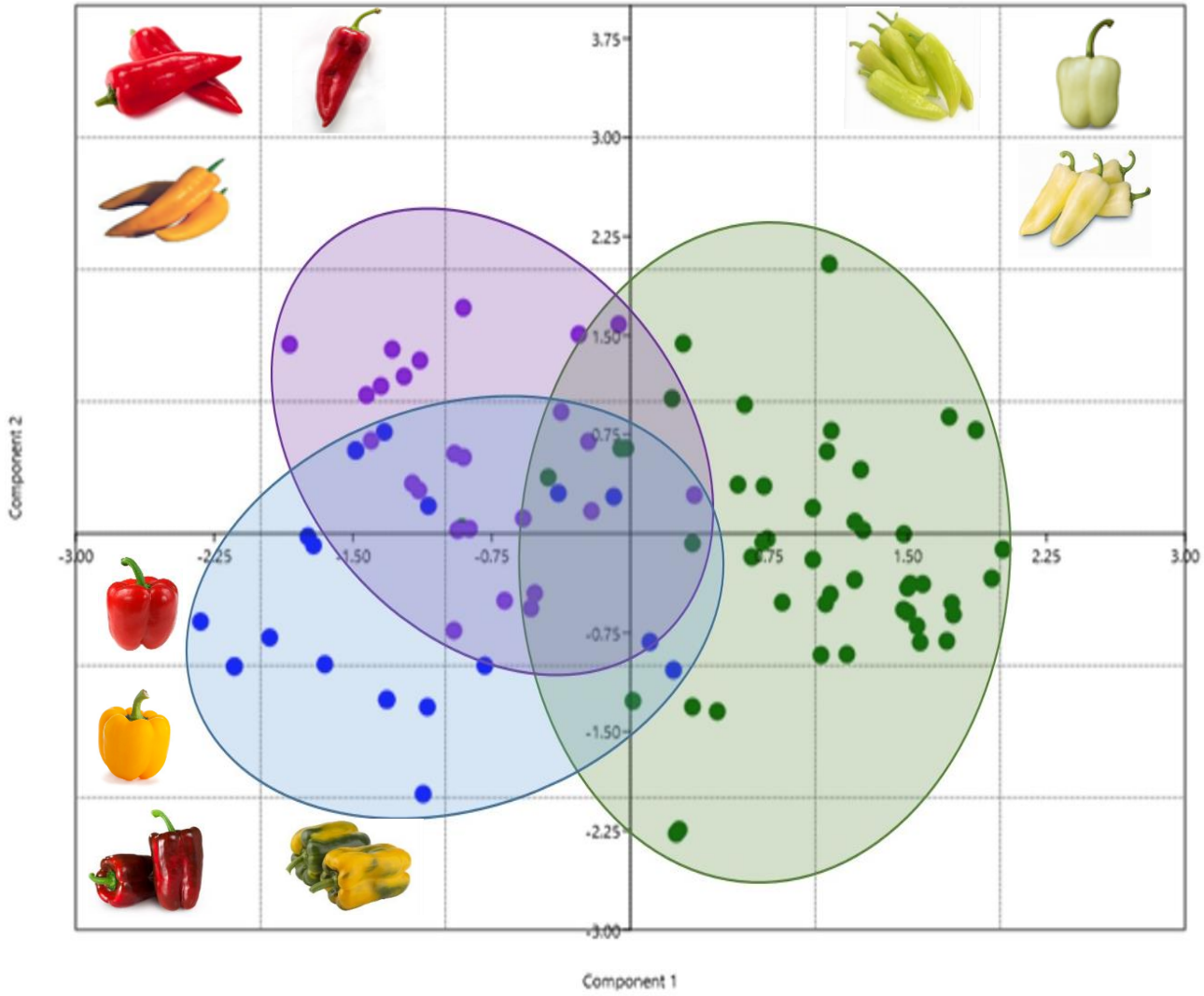


Figure 35. Positive clustering for provenience area**.** East Europe material (white type-charly-lastochka), green area (right side); west Europe material (blocky-lamuyo, square type), blue area (left down); both Europe area (bull horn-kapia, conical red type), violet area (left up)

Genotyping results allow to obtain a reading key about the possibility to recognize pepper type using the 48 SNP assay. In addition to this, about half (25 on 48) of the SNPs are not discriminatory for this determination. The rest show us a wider genetic similarity between the square and conical type. Key profiles for each main groups are reported below in Tab. 11.

	PEPPER TYPE AND PROVENIENCE AREA		
	SQUARE (WEST EU)	**CONICAL (EMEA)**	**WHITE (EAST EU)**
SNP-011	A	A	B
SNP-012	A	A	A
SNP-013	A	A	B
SNP-014	A/B	A	B
SNP-021	A	B	B
SNP-022	B	B	A
SNP-023	B	B	B
SNP-024	B	A	B
SNP-025	B	B	B
SNP-031	A	A	B
SNP-032	A/B	A/B	A
SNP-033	B	A	B
SNP-041	A	A	A
SNP-042	A/B	A	A
SNP-043	B	B	B
SNP-044	A	B	A
SNP-051	A	B	A
SNP-052	B	B	B
SNP-053	A	A/B	A/B
SNP-061	A	A/B	A
SNP-062	B	B	B
SNP-063	A/B	A	B
SNP-071	B	B	B
SNP-072	A/B	A	B
SNP-073	B	B	B
SNP-074	A	A	B
SNP-075	B	B	B
SNP-076	A	A	A
SNP-081	A	A	A
SNP-082	B	B	B
SNP-091	A/B	B	B
SNP-092	A	B	B
SNP-101	A/B	A	B
SNP-102	B	B	B
SNP-103	B	B	A
SNP-104	B	A	A
SNP-111	B	B	B
SNP-112	B	B	A
SNP-113	B	B	B
SNP-114	A	A/B	B
SNP-121	A	B	B
SNP-122	A	B	B
SNP-123	A	B	B
SNP-124	A/B	B	B
SNP-125	A	A	A
SNP-126	A	B	B
SNP-127	A	A	B
SNP-128	B	B	B

Table 11. Molecular pattern derivatives useful for pepper typology recognition

- **provenience area (hybrids F1)**

As for the previous analysis, investigation between genetic profile and provenience area (market focus) was performed also for hybrids genotypes (F1).
Confirming the positive correlation, families were identified and described in Fig. 36.

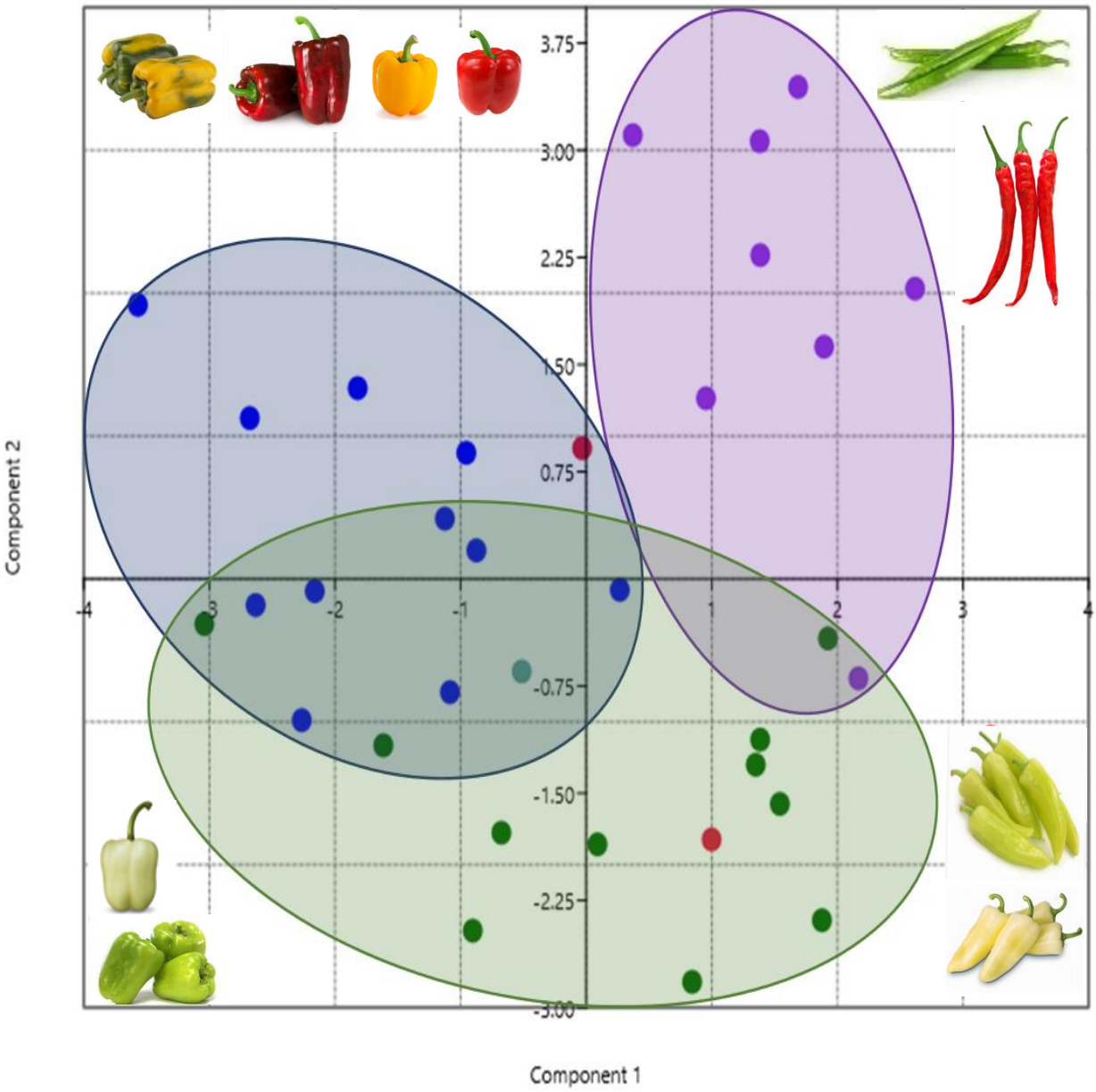


Figure 36. Positive clustering for provenience area (hybrids). East Europe material (white type-charly-lastochka), green area (down side); west Europe material (blocky-lamuyo, square type), blue area (left up); extra Europe, Asia, Africa (cayenna type), violet area (right up).

3.4 Conclusions

The constitution of a pepper germplasm collection was carried out via company experimental platform developed overall the EMEA area. In order to improve the efficiency of the genotypes choice, accurate panel set-up concerned further research activities related to evaluation of pathogen resistance by in vivo bioassay.

In addition to this, new pre-breeding cross combinations were obtained for the introgression of resistance gene and for broadening the genetic basis of the collection. This will be useful for the whole research activity of the Company.

Subsequent genetic characterization of the collection was determined for the understanding of the population structure. Genotyping was performed using Kasp 48 SNP assay. Evaluation of genetic diversity between the accessions was performed in order to determine the pepper collection genetic similarity. Derived SNP informativeness level was investigated considering the distribution of alleles within the population. This highlighted the efficiency of each SNP in term of discriminatory capability for polymorphisms detection.

Clustering analysis was conducted for the association studies relating SNP profiles and morpho-agronomic factor characterizing the collection.

Generally, no correlation was found between SNP and genotypes group (F4-F12) subdivided for resistance genes presence.

Using both breeding lines and hybrids, on the other hand, results show positive response through the highlight of molecularly distinguishable families for fruit characteristics (white type, conical red, square shape) as well as origin market area (east/west EU or present in both regions). Furthermore, polymorphisms analysis, allowed to define a molecular pattern useful as reading key about the pepper typology recognition related fruit structure and consequently provenience area.

In addition to this, data derived from the obtained genotypic profiles and the detected genetic variation contributed to the identification of interesting elite materials. These accessions are useful for the successive step of research activity, representing at the same time an important tool for the knowledge and use of breeding material within Company's program.

.

Edward Henry Weston (1886-1958)

Chapter 4

Screening of the proprietary collection with molecular markers for resistance to major pepper diseases

4.1 Introduction

One of the most important problems affecting pepper production is the occurrence of common pathogens causing severe diseases, epidemics resulting and significant yield losses. Managing pepper diseases by the use of genetic resistances, if such resistance is available, is usually the simplest and most cost-effective method of disease control and for this reason, in the last decades, most of the pepper breeding programs have been addressed to the constitution of resistant hybrids against a wide range of pest and pathogens.
A great advance in Capsicum molecular breeding for diseases resistances has been obtained during last 30 years (for a thorough review, see e.g. Parisi *et al.*, 2020).

Heritable resistance to *Tomato spotted wilt orthotospovirus* (TSWV) due to a single dominant gene (*Tsw*) responsible for hypersensitive reaction was identified in several accessions of *C. chinense* (PI159236, PI159234, PI152225) (Boiteux *et al*., 1995; Moury *et al*., 1997). *Tsw* gene was mapped in the distal portion of chromosome 10 using a segregant F2 population *C. chinense* (PI152225) x *C. frutescens* (PI195301) and a CAPS marker (SCAC 568) was identified (Moury *et al*., 2000). The position of *Tsw* has been defined more precisely in the area of NLR genes, 295-kb candidate region on chromosome 10 (Kim *et al*., 2017).
The resistance conferred by the *Tsw* gene is overcome by early plant virus inoculations as well as high temperatures (28-33 C°) (Moury *et al*., 1998). Nowadays the *Tsw* gene is widely used in the most pepper breeding program as unique source of TSWV resistance however, the rapid emergence of resistant-breaking (RB) isolates has been reported in Australia, Hungary, Argentina, Italy, Spain and more recently in California and China (Turina *et al*., 2016; Macedo *et al*., 2019). A new resistance possibly controlled by a different gene tightly linked to *Tsw* was identified in *C. chinense* AC09-207 (Hoang *et al*., 2013), moreover, other domesticated *Capsicum* species such as *C. frutescens*, *C. chacoense*, *C. pubescens*, *C. galapagoense*, *C. baccatum* var *pendulum* and var *baccatum* are known to carry the *Tsw* gene and can be thus used as source of resistance (Di Dato *et al*., 2015; Cebolla *et al*., 2003). In progress are studies on *Capsicum* germplasm resistant to RB-TSWV (Parisi *et al*., 2015). Among the few cases, *C. baccatum* accession PIM26-1 showed good level of tolerance to WT- and RB-TSWV isolates (Soler *et al*., 2015).

As regard resistance to *Potyvirus* in pepper, several resistance genes have been reported. The *pvr1* locus, identified in *C. chinense* PI159236 and PI152225 presented alleles with different

resistance levels (Yeam *et al*., 2005). The *pvr2* resistance alleles are found in the *C. annuum* accessions Yolo Y, PI264281, SC46252, Florida VR2 (Ruffel *et al.*, 2002).

Pvr2 [1] and *Pvr2* [2] alleles, have been used extensively for more than 50 years towards breed *Potyvirus* resistant pepper cultivars. Both alleles confer efficient resistance, while only pvr2 [2] is effective against *Tobacco etch virus* (TEV). The *pvr3* gene, reported in *C. annuum* cv. Avelar, confers resistance to PepMoV. The *pvr4* derived from *C. annuum* CM334 confers resistance to PVY-0, PVY-1,2 and PepMoV. Other sources of this gene were found *C. galapagoense*, *C. praetermissum*, *C. baccatum* var. *pendulum*, *C. frutescens* and *C. chinense* (Di Dato *et al*., 2015; Caranta *et al*., 1999). Resistance to PVY-0 and PVY-1 (isolate P-62-81) derived from the respectively recessive loci *pvr5* and *pvr8* is present in *C. annuum* CM334 (Andrés *et al*., 2004). The *pvr6* gene from *C. annuum* cv. Perennial and dominant gene *Pvr7* from C. *chinense*, confer resistance to ChiVMV and PepMoV Florida (V1182) strain, respectively. *Pvr4* and *Pvr7* are reported then to be the same locus that map tightly linked to *Tsw* on chromosome 10 (Venkatesh *et al*., 2018). Many markers have thus ben developed associated to genes identified. Scar SCUBC19, linked to *Pvr4* and mapped on chromosome 10 at distance variable from 5 to 10 cM, was developed by BSA-RAPD (Arnedo-Andres *et al*., 2002), a CAPS marker tightly linked to *Pvr4* was developed by BSA-AFLP (Caranta *et al*., 1999). Three CAPS markers for three recessive viral resistance alleles *pvr1*, *pvr11*, and *pvr12* and a functional SNP marker at the *pvr2-eIF4E* locus were developed (Rubio *et al*, 2008). KASP_pvr, a KASP marker for *pvr1* was validated using a *C. chinense* F2 population (Kang *et al*., 2005). Seven SNP highly associated with the resistance were identified on chromosomes 4, 6, 9 and 12 by GBS, in particular two of them were closely linked to *pvr2* on chromosome 4 (Tamisier *et al*., 2020).

As far as resistance to *Tobamoviruses* is regarded, different *L* genes were identified: *L1*, *L2*, *L3* and *L4* (Rast, A.T.B, 1988). *L1* gene confers resistance to P0 strains, *L2* gene confers resistance to P0 and P1, *L3* to P0, P1 and P1,2 while *L4* is responsible for resistance to all strains (P0, P1, P1,2 and P1,2,3) (Boukema, I.W, 1980). The *L* locus was mapped on the sub-telomeric region of pepper chromosome 11 (a syntenic region to the tomato chromosome 11 which carries genes for *F. oxysporum*) (Lefebvre *et al*., 2002). *L4* gene introgressed from *C. chacoense* confers resistance to the most aggressive and common tobamovirus pathotypes P1.2.3 (Matsunaga *et al*., 2003; Yang *et al*., 2009). Using a linked dominant marker 060I2END, several accessions of *C. baccatum* var. *pendulum* as well as in germplasm belonging to *C. frutescens*, *C. praetermissum*, *C. pubescens*, *C. chinense* and *C.*

praetermissum were identified as good sources of resistance to pathotypes P1.2.3 (Di Dato *et al*., 2015). SCAR WA31-1500S, linked to *L4* genes and several further tightly linked markers such as 189D23M (located within 0.1 cM of the *L3* gene) were identified (Matsunaga *et al*., 2003). L-linked markers (087H3T7, 060I2END and 158K24) related to the syntenic tomato I2 (resistance to *F. oxysporum* f. sp. *lycopersici*) and potato R3 (resistance to *P. infestans*) loci were developed. Regarding *L3* and *L4* gene, through further mapping of the previously identified marker 189D23M, different linkages were demonstrated, suggesting the possible existence of different genes closely linked instead that different alleles at the same locus (Yang *et al*., 2012).

A single dominant resistance gene *Cmr1* for *Cucumber Mosaic Virus*, was identified in the centromeric region of pepper chromosome 2 of the *C. annuum* cv. Bukang (Kang *et al*., 2010). However, a new isolate (CMV-P1), able to break down the resistance conferred by *Cmr1*, emerged in Korea (Lee *et al*., 2006). Recently, a new single recessive gene, named *Cmr2* with a single associated AFLP markers has been identified (Choi *et al*., 2018), as a gene that confers resistance to several CMV strains including to CMV-P1, CMV_{Korean} and CMV_{FNY}. In *Capsicum* spp. almost all the CMV resistance sources show a partial resistance controlled by multiple genes (Choi *et al*., 2018). QTLs on chromosome 11 associated with TMV genes were identified together with a major QTL on chromosome 12 explaining between the 45% and 63.6% of the phenotypic variation. Other two major QTLs associated with the tolerance to $CMV_{HB\text{-}jz}$ strain were identified on chromosome 5 and 11 (Caranta *et al*., 1999). Recently, NGS allowed the identification of two novel major QTLs responsible for the resistance to CMV-P1 on chromosomes 5 (52.7-58.1 cM) and 10 (21.9-32.5 cM) (Ben Chaim *et al*., 2001; Caranta *et al*., 2002; Eun *et al*., 2016). *CA02g19570*, a single gene located on chromosome 2 and conferring resistance to CMV_{FNY} was reported as the candidate for the QTL *qCmr2.1* (Guo *et al*., 2017). NGS enabled identification for other three QTLs for resistance to the $CMV_{HB\text{-}jz}$ and a major QTL on chromosome 11 that explain about 20% of the phenotypic variation (Li *et al*., 2018).

As regards resistance to nematodes, a single dominant gene (*N* gene) was reported as responsible for resistance mechanisms to *M. arenaria* (races 1 and 2), *M. incognita* and *M. javanica* identified in *C. chacoense C. annuum, C. chinense and C. frutescens* (Sarath Babu *et al*., 2011). Resistance to root-knot nematodes (RKN) is also associated with several dominant *Me* genes that act in an independent way (Djian-Caporalino *et al*., 2007). While *Me4*, *Mech1*

and *Mech2* are specific to some *Meloidogyne* species, others, like Me1, *Me3*, and *Me7* are able to control a wide range of nematodes species, including *M. arenaria*, *M. javanica*, and *M. incognita*. Genes *Me3* e *Me4* were found to be linked, separated by 10 cM and together with *Mech1*, *Mech2*, *Me1* and *Me7* represent the main cluster of 28 cM on chromosome 9. Demonstrating the existence of orthologous regions for nematode resistance in Solanaceae for both tomato and potato comparative mapping evidenced a colinearity with chromosome 12 (Djian-Caporalino *et al*., 2001). A subsequent study reported the *N* gene allelic to *Me7* and located 7 cM apart from *Me1*, and 2 cM from *Me3* on chromosome 9 together with markers tightly linked to *Me1*, *Me3*, *Me7* and *N* genes for assisted breeding (Fazari *et al*., 2012). Moreover, a QTL was mapped on chromosome 9 providing specific resistance to *M. javanica* (Barbary *et al*., 2016). Additional markers were identified such as a set of microsatellites tightly linked to *N* gene and a codominant CAPS marker for *Me1* gene (Uncu *et al*., 2015; Celik *et al*., 2016).

As regards *Phytophtora capsici* several *C. annum* resistant accessions have been identified for carrying a single dominant gene or multiple genes with additive and epistatic effects (Lefebvre *et al*., 1996). Among them, Serrano Criollo de Morelos (CM334) has the highest resistance level (Quirin *et al*., 2005). Through a comparative mapping strategy involving three intraspecific *C. annuum* populations, two main chromosomal regions were identified as conferring resistance to *P. capsici* (Thabuis *et al*., 2003). Common major QTL was positioned on chromosome 5 from the alignment of the relative genetic maps while, using a RIL mapping population, 16 chromosomal intervals containing single or clusters of resistance QTLs were identified. Moreover, five QTLs with an effect to the resistance to root rot were reported (Ogundiwin *et al*. 2005). A Random Amplification of Polymorphic DNA (RAPD) marker (OpD04.717) linked to the major QTL *Phyto.5.2* has been identified (Quirin *et al*., 2005). A major QTL (*Phyt-1*) on chromosome 5 and two minor QTLs on chromosome 1 and 11 explaining over 80% of the phenotypic variance, were identified. *Phyt-1* was described in the same chromosomal region of other major QTLs (*Phyto-P*; *Phyt.5.1*; *Phyt.5.2*) (Sugita *et al*., 2006). Seven QTLs, four of which related to the root rot (66.3% of the phenotypic variation) and other three responsible for switching off the resistance (45% of variation), were identified with closely linked RFLP markers (Kim *et al*., 2008). QTLs for resistance against different *P. capsici* isolates were mapped used in two separate studies using intraspecific RIL populations. The first one, identified seven QTLs located on chromosome 5 (among total of 15) and phenotypic variation explained from ~5% to ~50%. The second, allowed to detect 4

QTLs linked to the three main-effect loci related to *P. capsici* resistance (Lu *et al*., 2012). As far as associated molecular markers are regarded, SA133_4 SCAR marker, linked to root rot resistance in the region of QTLs on chromosome 5, was developed via bulked segregant analysis (BSA) combined with RAPD markers (Truong *et al*., 2012), while SNP markers tightly linked to the major QTL on chromosome 5 were identified combining BSA and microarray analysis (Affymetrix GeneChips) (Liu *et al*., 2014). On the latter *Phyto5SAR* was identified in a region containing clusters of resistance genes (NBS-LRR) associated with plant defense responses, while *Phyto5NBS1* was reported to have discrimination between susceptible and resistant lines accuracy over 90%. A single dominant gene, *CaPhyto*, on chromosome 5 and two candidate genes, *Capana05g000764* and *Capana05g000769* were reported to confer resistance to race 2 and a microsatellite marker ZL6726 (1.5 cM from *CaPhyto*) were reported (Wang *et al*., 2016). Several other molecular markers linked to resistance to *P. capsici* have been reported in pepper for more rapid selection (Quirin *et al*., 2005; Kim *et al*., 2008; Truong *et al*., 2012; Liu *et al*., 2014; Xu *et al*., 2016). Furthermore, was identified *Pc5.1*, a key QTL cluster on chromosome 5 conferring resistance against at least 12 *P. capsici* isolates (Mallard *et al*., 2013). Chromosome 5 was confirmed as the main region responsible for the resistance responses to the pathogen (Minamiyama *et al*., 2007). A single dominant gene (*PhR10*) positioned on the long arm of chromosome 10 and responsible for the resistance to race 3 was identified by means of BSA combined with sequencing of Specific Locus Amplified Fragment (SLAF-seq) (Xu *et al*., 2016).

As regards Bacterial spot, nine races (P0-P8) have been reported among worldwide *Xanthomonas* strains (Sahin *et al*., 1998). Five non-allelic dominant genes named *Bs1, Bs2, Bs3, Bs4,* and *Bs7* were reported to control hypersensitive reaction to *Xs* in different pepper accessions. *Bs7* in UNEF1556 (*C. baccatum* var. *pendulum*), Bs4 in PI235047 (*C. pubescens*), *Bs2* in PI260535 (*C. chacoense*), *Bs3* in PI271322 (*C. annuum*) and *Bs1* in PI163192 (*C. annuum*) (Wai *et al*., 2015). Two recessive genes, *bs5* and *bs6*, were moreover identified in PI271322, Pep13 and PI163192 (*C. annuum*) (Jones *et al*., 2004). *Bs1, Bs2* and *Bs3* have been widely introgressed into several commercial pepper cultivars. First high-resolution genetic mapping of *Bs2* (Tai *et al*., 1999) identified tightly associated molecular markers in *C. annuum* near-isogenic lines donor PI260435 (*C. chacoense*). AFLP markers linked to the *Bs3* were identified (Pierre *et al*., 2000) and a marker able to detect InDel polymorphism in the *Bs3* promoter (PR-Bs3) was reported (Römer *et al*., 2010). Additive effect, leading to complete resistance against P6 was demonstrated for the combination of *Bs5* and *Bs6* genes

(Vallejos *et al*., 2010). Kompetitive Allele-Specific PCR (KASP) has been used to develop markers linked to the Bs3 locus (Holdsworth *et al*., 2015). Functional studies showed role of *CaPO2* (*C. annuum* peroxidase) gene (Choi *et al*., 2007), while silencing of *CaMLO2* was shown to enhance the resistance against virulent *Xanthomonas* (Kim *et al*., 2012).

Immune or highly resistant genotypes to *Leveillula taurica* derived from different *Capsicum* species have been reported (Anand *et al*., 1987; Lee *et al*., 2001). In H3 cultivar at least three genes could be involved towards resistance to this fungus (Daubèze *et al*., 1995). Lefebvre and colleagues mapped the first gene and reported epistatic interactions explaining more than 50% of the genotypic variance and describe quantitative nature of the resistance identifying a QTLs on chromosome 6 (*Lt 6.1*) (Lefebvre *et al*., 2003). A single dominant locus (*pmr1*) was reported as responsible for the powdery mildew resistance, located in a syntenic region of 4 Mb of pepper chromosome 4. Two genes across the 622 predicted ones, within this locus, were found for sequence similarity to the nucleotide-binding site leucine-rich repeat domain containing R proteins (NBS-LRR). Six molecular markers (one SCAR and five SNP) were identified by the authors as tightly linked to *pmr1* and the regions showed close relatedness between *C. baccatum* and *C. annuum* (Jo *et al*., 2017). Moreover, two recessive S-genes *CaMlo1* and *CaMlo2* responsible for pepper resistance to powdery mildew were reported in the frame of S-genes. Loss of function of these genes is responsible for reduction of disease susceptibility (Quirin *et al*., 2005).

The exploitation of *Capsicum* germplasm (close related species, wild relatives, landraces and pre-breeding materials) for biotic stresses and its use within breeding program, despite the effort made, still represent a challenging task (Parisi *et al.*, 2020). Indeed, the risk of a resistance breakdown together the climate change, inevitably affect the durability of diseases resistances making gene pyramiding as the main strategy to create more durable and broad-spectrum mechanisms through the combining of one or more alleles of main genes (Özkaynak *et al*., 2014).

In this chapter, we describe the optimization of user-friendly protocols for known markers associated to resistance genes and QTLs for most relevant pathologies. Such protocols/markers will be applied to routine screening of the company proprietary collection. Further complementary information generated in this manner is required for the identification of elite materials analyzed on subsequent pre-breeding activities.

4.2 Materials and methods

4.2.1 Genes/QTLs identification towards markers assisted selection

Bibliographic studies were conducted for the identification of the most relevant pepper diseases (viruses, nematodes, fungi, bacteria) present in the Mediterranean area and related genes and QTLs involved in resistance mechanisms. Associated molecular markers have been investigate through scientific literature using Scopus databases and NCBI PubMed.

Among the major pathogens of pepper *Phytophthora capsici*, Powdery mildew, *Cucumber mosaic virus*, *Tobamoviruses*, *Tomato spotted wilt virus*, *Potyviruses*, Bacterial Spot and nematodes were chosen as those of interest. Related genes and QTLs implied in resistance mechanisms were selected from bibliographic research and associated molecular with relative primer sequences were identified and retrieved from literature.

4.2.2 Verification of primer sequence for subsequent design

Using the NCBI and Sol Genomics Network (https://solgenomics.net) databases, homology search was performed for each primer pair retrieved from literature as major part of the markers and primers had been designed using mainly EST sequences before the availability of high-quality reference genomes.

Primer pairs comparison was conducted using BLASTn search at Sol Genomics Network web site on the pepper genomic sequences *C. annuum* cv. CM334, *C. annuum* Zunla, *C. annuum* var. *glabriusculum* and *C. annuum* UCD10X.

BLASTn function was crucial to identify, using all available genomic sequences of pepper, the position of primers and verify their sequences as well as expected length of PCR products.

For a given molecular marker (if available) the polymorphisms (SNPs, InDels) identified between different genomic sequences were compared with those described in literature. Selected SSR, InDel and SCAR markers primer pairs were purchased from Invitrogen, a branch of ThermoFisher Scientific (https://www.thermofisher.com).

For SSR markers M13 tailed-primer strategy provided to use three primers:

- a forward primer with a nucleotide extension at its 5' end, whose sequence was identical to the sequence of an M13 universal primer;
- a reverse primer with standard length;
- a fluorescent M13 universal primer, that was fluorescent dye-labeled at its 5' end.

4.2.3 Collection's DNA storage and procedures for working use

Germplasm collection, entirely composed of *Capsicum annum* species, was bred both on the company facilities in Italy and in Hungary. Seed of each accession was sown in trays (360 cell) and placed in the growth chamber for 10 days with a temperature range (22 °C night for 10 hours, 30°C day for 14 hours).

Once seedling were raised, plants were moved to the nursery (glass greenhouse) for 1 week. At the stage of 2 leaves, plants were removed and planted in 40 cell tray using specific sub-acid peat. Irrigation was carried out daily with two applications of foliar fertilizers (NPK 20-20-20).

Greenhouse transplant was carried out after 60 days from sowing.

Here below figures 37 to 40 show the step progress about plant development (sowing to transplant) and successive leaf sample collection for DNA extraction.

Figure 37 to 40. Germplasm collection step progress about plant development (Esasem facilities)

DNA Extraction with CTAB Method

Leaf tissue sampling from 3 plants for each genotype was carried out for the subsequent storage of DNA. A total of 5 plates (96 tubes) was collected and processed for the DNA extraction.

Extraction buffer CTAB (1.5 mL of 2X) was added to leaf material (100 mg).

In the presence of buffer CTAB, the tissues were grinded directly in tubes and then were incubated at 65°C for 45 minutes. The grinding was necessary for cell wall and plasmatic membrane break.

In the ratio 24:1, 650 µL of chloroform: isoamyl alcohol were added to 750 µL of each lysed sample and transferred to a new tube.

All samples were subsequently mixed by inverting the tubes (15 minutes) and centrifuged (15 minutes) at 10,000x g (13,000 rpm) at 10 °C.

Supernatant were taken (750 µL) and transferred into new 1.5 mL eppendorf tube.

For each sample 5 µL of RNase were added and then mixed by inverting the tubes (1 minutes) and incubated (15 minutes at 37 °C).

In order to DNA precipitation were added to samples 560 µL of cold isopropanol and then they were mixed via tubes inverting and centrifuged at 4 °C for 10 minutes at 10,000 x g (13000 rpm). Supernatant was removed after the incubation and added 1 mL of Wash solution II.

Solution was removed following another incubation (5 minutes) and DNA pellet was dried. Finally, within new 1.5 mL eppendorf tubes, DNA was resuspended in 100 µL of water.

Reagents and solutions:

CTAB 2X (Vf = 1L)

- CTAB 2%: 20 g;
- Tris/HCl (pH 8.0) 200 mM: 200 mL by stock 1M;
- EDTA 20 mM: 40 mL by stock 0.5 M;
- NaCl 1.4 M: 81 g;
- PVP K 30 1.0%: 10 g;
- b-Mercaptoethanol (BME) 0.28 M: 20 mL;
- dd H2O: at volume of 1 L.

RNase (10 mg/mL)

- RNase A: 200 mg;
- Tris/HCl (pH 8.0) 10 mM: 200 µL by stock 1M;
- NaCl 150 mM: 600 µL by stock 5 M;
- dd H2O: at volume of 1 L. Wash solution I;
- Ethanol 76%: 760 mL from Ethanol 100%;
- Sodium acetate 200 mM: 66.7 mL from NaAc 3 M;
- ddH2O: at volume of 1 L.

Wash solution II

- Ethanol 76%: 760 mL from Ethanol 100%;
- Ammoniumacetate 10 mM: 771 mg of NH4Ac;
- ddH2O: at volume of 1 L.

DNA quantification and diluitions to working concentration

DNA quantity and quality were evaluated using the spectrophotometer NanoDrop® ND-1000 and by agarose gel electrophoresis (0.9% w/v ultrapure agarose stained with SYBR Safe DNA gel stain Invitrogen).
DNA concentration established for PCR analysis was 50 ng/µL therefore all the samples composed the collection are respectively diluited reaching the working diluition in a 96 strip plate.

4.2.4 PCR protocols optimization

Each obtained molecular marker has been submitted to amplification protocol optimization in order to successive validation. Basic PCR mixes were composed with listed components:

- ddH_2O, double distilled water;
- green buffer PCR, (15 mM-$MgCl_2$) or buffer PCR $MgCl_2$ free (Thermo Scientific);
- dNTP (Thermo Scientific);
- primers for each molecular marker and M13 primer 700 labelled FAM for SSR markers;
- Dream Taq DNA polymerase (Thermo Scientific);
- 50 ng of genomic DNA.

dNTPs (dATP, dTTP, dCTP, dGTP) with concentration of 10 mM, obtained using 10 µL of each dNTP and adding 360 µL of water were used.
Applied biosystems 2720 thermal cycler was used for PCR reaction.
Below, reaction PCR mixes used for each molecular markers associated to pathogens resistances.

For the first stage of the protocol set-up, 1-2 resistant homozygous and 1-2 susceptible homozygous (depending on the availability) genotypes were used to identify promising markers associated with resistance genes.
The $MgCl_2$ concentration and temperature of annealing were adjusted to obtain single, clear bands.
Subsequently markers that passed the first stage were tested on a larger group of genotypes for each pathology (including F1 where available).

Enzymatic Digestion

For Caps markers the following restriction enzymes were used:

- restriction enzymes and relative recognition site with related restriction enzyme, buffer type and digestion temperature;
- mix for digestion of amplified fragments derived from PCR previously described.

Restriction enzymes	Recognition sites	
Cail (AlwNI)	5' C A G N N N ↓C T G 3'	
	3' G T C ↑N N N G A C 5'	
XbaI	5' T ↓C T A G A 3'	
	3' A G A T C ↑T 5'	
TaqI	5' T ↓C G A 3'	
	3' A G C ↑T 5'	
HaeIII	5' G G ↓C C 3'	
	3' C C ↑G G 5'	
HinfI	5' G ↓A N T C 3'	
	3' C T N A ↑G 5'	
XmnI	5' G A A N N ↓N N T T C 3'	
	3' C T T N N ↑N N A A G 5'	

Pathogens	Marker	Restriction enzyme	Buffer	Digestion Temperature
Potyvirus	CSO	Cail (AlwNI)	Tango	37 °C
Tomato spotted wilt virus	scac568F	XbaI	Tango	37 °C
		TaqI	TaqI buffer	65°C
		HaeIII	HaeIII buffer	37 °C
Leveillula taurica	CZ2_11628	TaqI	TaqI buffer	65°C
Cucumber mosaic virus	caTm-int1	HinfI	HinfI buffer	37 °C
	CAPS A	XbaI	Tango	37 °C
	CAPS B	XbaI	Tango	37 °C
Phytophthora capsici	NBS1-CAPS	XmnI	XmnI buffer	37 °C

Reagents	MIX 1X (µL)
DNA sample amplified	10,0
ddH_2O	7,0
Restriction enzyme	1,0
Buffer PCR (10X)	2,0

Tables 12-13-14. Caps marker and related restriction enzymes details for digestion process

All enzymes and buffer were provided by Thermo Scientific.

Gel electrophoresis

All samples were separated using agarose gel:

- CAPS markers: 2% agarose gel in (TBE 0.5X) at the concentration of 0.8% additioned with 2.5 µL of SYBR Safe DNA gel stain (Invitrogen);

- SCAR and SSR markers: 1.5% agarose gel in TBE 0.5X at the concentration of 0.8% additioned with 2.5 µL of SYBR Safe DNA gel stain (Invitrogen).

Molecular size marker GeneRuler 50 bp DNA ladder (Thermofisher) was used.

4.3 Results

4.3.1 Defining of associated molecular markers pool

Taking into account the company strategy regarding breeding objectives for diseases resistances, deep bibliographic research was conducted towards the identification of resistance genes and Quantitative trait loci (QTLs) for the following main pepper pathogens present in Mediterranean area: nematodes, *Tobamovirus*, *Cucumber mosaic viruses*, *Potyviruses*, *Tomato spotted wilt virus*, Bacterial spot, Powdery mildew and *Phytophthora capsici.*

After that, research was extended to retrieve associated molecular markers for detection of positive allelic variants useful for the marker assisted selection activities.

A total of 43 resistance genes/QTLs associated with identified pepper diseases and relative 121 molecular markers were identified. For further bioinformatics analysis, overall 73 molecular PCR markers (SCAR, SSR, InDel, SNP based as CAPS allele-specific and HRM) were chosen as most promising for user-friendly marker assisted selection protocols.

Global list of identified resistance genes/QTLs and associated molecular markers with related primer pair sequences is listed on the Tab. 15.

Schematic representation of disease resistance gene/QTLs position on pepper genome is reported in Fig. 41.

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Leveillula Taurica	pmr 1	dominant	4	ZL1_10691	scar	TCCTGTTTTCTCCCCCTTTT	CTTTGGCAATATCCCGTTCA	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HZ2_11079B	hrm	CTCTTTCGTTTGTTTTGCTTCA	CTTTCAGCTCCTCTCCCAGC	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	CZ2_11628	caps	GCTAGGATCCTGCTCGTGAGA	GTTGCTCTTGCTTCTGCTGC	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HZI_11658	hrm	TGCAAAATTTGATTCTTATAGTGGG	CCTGTCGAAACTACGAGTCAAAA	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	ZL1_1826	scar	CGAAGTCATTAAAGTTCATTGGG	GCAATAAATGCCCTTCCACA	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HPGV_1313	hrm	GGGTTTTCACTCCTCTTTTGC	TCCACCATGAAGGTGTAACG	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HPGV 1344	hrm	AAAAGGCAAGAGCATTACATGA	TTGTTGTTGTCGTTGTTGTTGA	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HPGV 1412	hrm	TCTCGGAGGGAAAACTGAAA	AAGCATAAGGGCATGTTTGG	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HRM4.1.6	hrm	AATTAAAGGACTTAAGTTTGACAGTT	GAAATTGTCGATGAACATCCGT	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	HRM2_A4	hrm	TTCAGCCAGTGATCTGGAGC	TCAAATTCCTTGCACAAATCAT	Jo *et al.*, 2017
Leveillula Taurica	pmr 1	dominant	4	CA04g00360	hrm	GCACGCCCATACCTTGTCAA	ATCAATGCCAAGCCCATCCA	Ahn *et al.*, 2018
Leveillula Taurica	pmr 1	dominant	4	CA04g00830	hrm	GGGGGCTAGTCTTCTCTTCT	GGCAACAAGGTGGAAAGACG	Ahn *et al.*, 2018
Leveillula Taurica	pmr 1	dominant	4	CA04g00250	hrm	CGGATCATCCCGGCATTGAT	TCACCTCCGATTCACAACTCA	Ahn *et al.*, 2018
Leveillula Taurica	pmr 1	dominant	4	opa15	scar	GATTTAGTCGAGGTGCATGAAAGT	TAYSARGCAGARYTASWRWTCCAAGT	Rajes *et al.*, 2018

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Cucumber Mosaic Virus	cmr 1	dominant	2	cmr1	snp	N.A.	N.A.	Kim *et al.*, 2017
Cucumber Mosaic Virus	cmr 1	dominant	2	CAPS - A	caps	GTAGTAGGGTACGGACTCATA	GTCCCGACGATAGCCCAAAAG	Kang *et al.*, 2010
Cucumber Mosaic Virus	cmr 1	dominant	2	CAPS - B	caps	GTAGTAGGGTACGGACTCATA	GGAGTTTCATCATATGAAGCC	Kang *et al.*, 2010
Cucumber Mosaic Virus	cmr 1	dominant	2	CaTm-int1	caps	TCAGCAAAGAAAGATTCACGAAC	ACGTACACTTGATGATGCCTTGT	Kang *et al.*, 2010
Cucumber Mosaic Virus	cmr 1	dominant	2	CaTm-int3HRM	hrm	TGGTGTTTTTATCAGCCTTAGC	GAAGGACAAGAATTCATGATATGG	Kang *et al.*, 2010
Cucumber Mosaic Virus	cmr 1	dominant	2	CaT1616BAC	snp	AATTGGTTCTGGATCACTGCC	CACCTTGATCTGCTCCTTTCTG	Kang *et al.*, 2010
Cucumber Mosaic Virus	cmr 1	dominant	2	240H02sp6	snp	TTGGTTGAGGCAAGTTTTCA	TCATTTTCCTATGTCATTCATGG	Kang *et al.*, 2010
Cucumber Mosaic Virus	qtl qcmv11.1	dominant	11	6201026	slaf	N.A.	N.A.	Li *et al.*, 2018
Cucumber Mosaic Virus	qtl qcmv11.2	dominant	11	5409028	slaf	N.A.	N.A.	Li *et al.*, 2018
Cucumber Mosaic Virus	qtl qcmv12.1	dominant	12	17652010	slaf	N.A.	N.A.	Li *et al.*, 2018
Cucumber Mosaic Virus	cmr 2	recessive	8	affy4	kasp	Fam:CCGACTTCGAGCAAGCCTACAT Common:CGTCCTGACCCGCCTGCCAT	Hex:CGACTTCGAGCAAGCCTACAG Common: CGTCCTGACCCGCCTGCCAT	Choi *et al.*, 2018
Cucumber Mosaic Virus	cmr 2	recessive	8	IBP160	hrm	CTTGACGTTGGACCCATCAA	TGGACGTTCCCAATCAGAGA	Choi *et al.*, 2018
Cucumber Mosaic Virus	cmr 2	recessive	8	cmvAFLP	hrm	TGCAGTTGGAGCAGAAGATG	CATGGAAAGACTCCCAAGGAAC	Choi *et al.*, 2018

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Nematodes (Meloidogyne species)	Me1	dominant	9	16830-H-V2	hrm	CATCTTGTAATCGGGATGCC	GGCGATTGACCCATCTCTTT	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	16830-CAPS	caps	ATGACCATCTGGAAATCAGGTGT	TCATGACGGTTAATAGGATTTGCA	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	16880-1-V2	scar	TGACCCCTCAGACTGAACAG	CTCCTTCGCTGCTACCTTCT	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	K32	kasp	A1:GAAGGTGACCAAGTTCATGCTACTTGCACCGATTTTTAAAGCTAGAC A2:GAAGGTCGGAGTCAACGGATTGTACTTGCACCGATTTTTAAAGCTAGAT	C:TCAAGCGAGGAAAACCTTCATCAGCT	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	K77	kasp	A1:GAAGGTGACCAAGTTCATGCTCCGAGTCCCTTAAACTATCATATAATGCCTGCG A2:GAAGGTCGGAGTCAACGGATTCGAGTCCCTTAAACTATCATATAATGCCTACA	C:TCAAGCGAGGAAAACCTTCATCAGCT	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	S24	ssr	AGCAAAACGGTAGCACAACA	TACTATCTCCTGAGCCGGGG	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	ZY6-13	ssr	CAAGATGGCTAAATCACAAGTCA	ACCGCGTTTTCTTTTCTTTTGT	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	24-20	ssr	ACCAATACCATGTCAAGAACCAC	CCAGAGAGCTTAGACACCCC	Wang *et al.*, 2018
Nematodes (Meloidogyne species)	Me1	dominant	9	C2At2g06530	ssr	AAGGTGTCTCCCTCAGAATTCAG	ATCTGTCCCATTGCCTTTGTAAC	Uncu *et al.*, 2015
Nematodes (Meloidogyne species)	Me1	dominant	9	C2At2g06530	caps	TTGGTGCTGTAAGGGACTAAA	TCTTAATCAATCATTCACACAGCA	Uncu *et al.*, 2015
Nematodes (Meloidogyne species)	Me1	dominant	9	SCAR_HM6	scar	GCTTATCATGGCTAGTAGGG	CGGACCATACTGGGACGATC	Fazari *et al.*, 2012
Nematodes (Meloidogyne species)	Me1	dominant	9	SCAR_HM60	scar	TATCCGTGGTCATCCTAGCC	TGTGGTTCATCGGGACTGTA	Fazari *et al.*, 2012
Nematodes (Meloidogyne species)	Me1-N	dominant	9	SSCP_PM5	sscp	GTTCTATTTTCTCGTTTCATCTTTCATTTT	ATAACCTACAAATTTCTGGAGCTTTGTCTA	Fazari *et al.*, 2012
Nematodes (Meloidogyne species)	Me1-Me7	dominant	9	scar_CD	scar	GAAGCTTATGTGGTAMCC	GCAAAGTAATTATATGCAAGAGT	Ponnam *et al.*, 2019
Nematodes (Meloidogyne species)	Me1-Me7	dominant	9	SCAR_PM54	scar	CTGCAGGGTAGCAAAGTAATTATAT	CCAAAATTAGTCATGTTCTTATGTTCTTAC	Fazari *et al.*, 2012
Nematodes (Meloidogyne species)	Me3-Me4	dominant	9	scar_B94	scar	GCTTATCATGGCTAGTAGGG	CGGACCATACTGGGACGATC	Ponnam *et al.*, 2019
Nematodes (Meloidogyne species)	Me3-Me4-Me7	dominant	9	SSCP_B54	sscp	CGGTGGCTGTTACGCTC	GCATGTCTTTCTTTACC	Ponnam *et al.*, 2019

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Nematodes (Meloidogyne species)	Me3-Me4-Me7	dominant	9	SSCP_B322	sscp	GATTCCATAACCTGGAAATTTCTGG	CGAACCCGGTCTATTTTC	Ponnam *et al.*, 2019
Nematodes (Meloidogyne species)	Me3-Me4-Me7	dominant	9	SCAR_PM6a	scar	TTCTTACCCCTGTACATCACATCCT	AACCTGGAAATTTCTGGAGGTATG	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me3-Me4-Me7	dominant	9	SCAR_PM6b	scar	CCCGGTCTATTTTCCCTTTT	TGTCTAAATTCTCATGGCAGTG	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me3-Me7-N	dominant	9	SCAR_N	scar	AATTCAGAAAAAGACTTGGAAGG	TAAAGGGATTCATTTTATGCATAC	Fazari *et al.*, 2012
Nematodes (Meloidogyne species)	Me3-Me7-N	dominant	9	SCAR_315	scar	N.A.	N.A.	Zhang *et al.*, 2012
Nematodes (Meloidogyne species)	Me7	dominant	9	SCAR	scar	AATTCAGAAAAAGACTTGGAAGG	TAAAGGGATTCATTTTATGCATAC	Ponnam *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	G21U3	hrm	ACAAATGACAACTTTCCTGC	ACATGGACAGGAGATACGAA	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	2111b1	hrm	TCTGGGCCAAAAATGCTACCA	TGGTTTGACTAACACTCCTGCA	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	SF164024	hrm	GCAGCATCAGCATCCGAATCTT	TCCCTTATTATGAGTCGTTGCCT	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	SF16406	hrm	CAGTGGCTGATCGAGAGCAC	CAGACATCAGAGGCATGCCA	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	SF164076	hrm	GGGAGGATTGCCCAAGACAA	GGGATGATTAGTAGTATCCAGTGGC	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	611109646	hrm	AGAAGCATTGGGTGGAAC	TCAACCTGCTTCTCCTTAAC	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	CA1-1b	hrm	GACGAAATTTGCGCTATTCA	GATCCATGAAGTCCATCTGC	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	G24U5	hrm	ACAACCAAGAACATATGGCT	AGTTCATTGTTTCCGATGTTG	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7	dominant	9	G43U3	hrm	AAGACGATCCTGTTAGAGTG	TGGGACTTTTACCCACTCT	Changkwian *et al.*, 2019
Nematodes (Meloidogyne species)	Me7-Mech 1	dominant	9	F4-R4	caps	AGAACAATAGAATCTCTCTTG	CTTCAGGAACCCCTCAGC	Djian *et al.*, 2007
Nematodes (Meloidogyne species)	N	dominant	9	CASSR37	ssr	ACATACCCAAAAACTCTCTCAC	GATTGACCATGTTTCCGTAT	Changkwian *et al.*, 2019

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Tobamovirus	L1	dominant	11	L-V0-6	scar	ATTGCAGGTACACCCAACCAATCT	CCCGCAAAGATATGTGTACCAACA	Lee *et al.*, 2012
Tobamovirus	L1	dominant	11	L1-SCAR	scar	AATAGGCAAGCAAAGGCTAAGTTG	TCGGACACTTCAAAAATGTCTACG	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	189D23M-NK	scar	ATTGTCAGAGTCGGGAAGCA	TACTATGCACAGGGTCTAGG	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	189D23M-YB	scar	ATTGTCAGAGTCGGGAAGCA	AACGACAAGGGTTTATTGTATGC	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	A339-NK	scar	TCTCGGTAGGCCATTTTGCT	GTAAGTTGCTATGCCCACCA	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	A339-YB	scar	GTTTTACATGAAACGCGTTC	GAAGATAGTGGTGGAGAAAA	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	253A1R-YB	scar	GCTTCTCCCCAAATGTAGCA	CTCAAACGAGTCATGGCTGA	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	L3-SCAR	scar	AACAATTTACAAATAATACACAAGGC	TTGGGAAGGAAAGACATCAT	Lee *et al.*, 2012
Tobamovirus	L3	dominant	11	PMFR11-269 - PMFR11-283	scar	CTGCAGAACAACAATGGCACG	GGACTGCAGAGGAGGAAGC	Sugita *et al.*, 2004
Tobamovirus	L3	dominant	11	PMFR21-200	scar	GCCAAAATGGTAATTGGACATTTTTAACACAAACT	GGACTGCAGAGGAGGAAGC	Sugita *et al.*, 2004
Tobamovirus	L3-L4	dominant	11	L4SC340	scar	AAGGGGCGTTCTTGAGCCAA	TCCATGGAGTTGTTCTGCAT	Kim *et al.*, 2008
Tobamovirus	L4	dominant	11	L4	snp	N.A.	N.A.	Kim *et al.*, 2017
Tobamovirus	L4	dominant	11	087H3T7-HRM	snp	CATGATTACATTTTATGTTGC	AAAAGGAAGGTTCTCATTGTT	Yang *et al.*, 2009
Tobamovirus	L4	dominant	11	087H3T7	caps	CCTTTGCCTGCATTATTCTTG	GCCCAAATTTATTCCCAAATGC	Yang *et al.*, 2009
Tobamovirus	L4	dominant	11	060I2END	scar	GCACATCAGCAGGTTTAGTACG	CCAACTGTCAAACCTCGG	Yang *et al.*, 2009

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Tobamovirus	L4	dominant	11	158K24HRM	snp	CAGATTAAGTGTTCAAAATGAGTGATG	TGATTCCATGAAAATAAATTGTAAAGA	Yang *et al.*, 2009
Tobamovirus	L4	dominant	11	L4RP-3F/L4RP-3R	N.A.	TCTTCAGCACCTCAATTCGGTTC	GAAGAGGGCATCCCTTTTACT	Yang *et al.*, 2011
Tobamovirus	L4	dominant	11	3'endR	N.A.	N.A.	TCACAGGCATTCACAGTCAAACATAGTGCGACC	Yang *et al.*, 2011
Tobamovirus	L4	dominant	11	WA31-1500S	scar	CGTACTGTGGCTCAAAACTC	ATTCGCACCGTT TAGCCCGT	Hiroshi *et al.*, 2003
Tobamovirus	L4	dominant	11	L-V0-4	scar	TATCGATGCACCCCTCGTTTTAAT	ATCTACCACAATGGCAGTGACGAA	Lee *et al.*, 2012
Tobamovirus	L4	dominant	11	L4-SCAR	scar	ATCGATGCACCCCTCGTTTTAATC	GAGCAGTGTGGAGTGTCTATTGCTCA	Lee *et al.*, 2012
Tomato Spotted Wilt Virus	tsw	dominant	10	CAPS -SCAC 568	caps	GTGCCAGAGGAGGATTTAT	GCGAGGTGGACACTGATACT	Di Dato et al.,2015, Özkaynak et al.,2014, Moury *et al.*, 2000
Phythophtora Capsici	CaPhyto, qtl 5.1	dominant	5	CaNB-5480	N.A.	TCGAAATCAATACTCCTCCTTCCC	CTCACGCGTTGTTCTTAAAAGGTT	Kim *et al.* ,2019
Phythophtora Capsici	CaPhyto, qtl 5.1	dominant	5	5NBS1	snp	TTGATAGCCCCTGGTAAAGA	GTGGTGATATTCAAACGGG	Liu *et al.*, 2014
Phythophtora Capsici	CaPhyto, qtl 5.2	dominant	5	SNP M3-2	snp	N.A.	N.A.	Kim *et al.*, 2017
Phythophtora Capsici	CaPhyto, qtl 5.2	dominant	5	SNP M3-3	snp	N.A.	N.A.	Kim *et al.*, 2017
Phythophtora Capsici	CaPhyto, qtl 5.2	dominant	5	ZL6726	ssr	TCCAGCCATCCATTATTTCAT	ATCCCGAACTGCCAATAATTA	Wang *et al.*, 2016
Phythophtora Capsici	CaPhyto, qtl 5.2	dominant	5	SA133_4	scar	GAATCACAAGGAAAAGAAAACAAG	TGAAAGGAGTCTCTGAATCCATAA	Truong *et al.*, 2013
Phythophtora Capsici	meta qtl Pc 5.1	dominant	5	P5-SNAP	snap	F1:TGAGGTTGCTATTAAGATTGGTCCTGTTATATA F2:GAGGTTGCTATTAAGATTGGTCCTGTTATCCG	CATAGAAAGGGATATCATCTGGTACATGCAGAAA	Ponnam *et al.*, 2019
Phythophtora Capsici	meta qtl Pc 5.2	dominant	5	CAMS420	ssr	CAGCGTTCTATCGTCTCAAATG	TTGACAAACCAGAAATTGATCG	Minamiyama *et al.*, 2007
Phythophtora Capsici	meta qtl Pc 5.2	dominant	9	SSR-9	ssr	CAAGCACCTACAAATGCAAAAT	CCGGATGAGAAAACTTGCTACT	Ponnam *et al.*, 2019
Phythophtora Capsici	Pc qtl 5.1	dominant	5	Sn2	caps	TTCGATCCACACCATCATCT	TCCTTCAATGGCTTTCCATC	Mallard *et al.*, 2013

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Phythophtora Capsici	Pc qtl 5.1	dominant	5	X79231.1	caps	TCATGAGGTTGCTATTAAGAATGGTCCTGTTATATA	CATAGAAAGGGATATCATCTGGTACATGCAGAAA	Mallard *et al.*, 2013
Phythophtora Capsici	PhR10	dominant	10	P52-11-21	ssr	CAATCCAAACAAGTCCTAAG	GGTGCAATTGAAAATCTAAG	Xu *et al.*, 2016
Phythophtora Capsici	PhR10	dominant	10	P52-11-41	ssr	TTGATGAGATGGGAAGTAAA	CACCAACAATAATAGAACTACA	Xu *et al.*, 2016
Phythophtora Capsici	phyto 5	dominant	5	CaRP-5130	N.A.	GACACATTTTGCAGATTCATCAAC	CACCCAAAAAGGTAAAAAGAAACA	Kim *et al.* ,2019
Phythophtora Capsici	phyto 5	dominant	5	CaNB-5330	N.A.	AAGAAAGCCGTCACCTTCATAGAT	GCTAATTTGCAAGGATTACGCTCA	Kim *et al.* ,2019
Phythophtora Capsici	Phyto, qtl 5.2	dominant	5	OpD04.717	scar	CCATAAGGGTTGGTAAATTTACAAAG	TCGAGAGATAATTCAGATAGTATAATC	Quirin *et al.*, 2005
Phythophtora Capsici	N.A.	dominant	5	CA524065	ssr	TCTCTCTCTACATCTCTCCGTTG	TGTCGTTCGTCGACGTACTC	Wang *et al.*, 2016
Phythophtora Capsici	N.A.	dominant	5	NBS1-CAPS	caps	AGGACTTTGATAAGGTTTC	TGCAATATAGAGCTTCTGCTG	Wang *et al.*, 2016
Phythophtora Capsici	N.A.	dominant	5	Phyto5SAR	snp	GGCACAACAATAGTCACAACGG	GAGACTAAGAAAGTTGGACGCC	Liu *et al.*, 2014
Bacterial Spot	Bs2	dominant	9	Bs2	snp	N.A.	N.A.	Kim *et al.*, 2017
Bacterial Spot	Bs2	dominant	9	F4-R4	caps	AGAACAATAGAATCTCTCTTG	CTTCAGGAACCCCTCAGC	Djian *et al.*, 2007
Bacterial Spot	Bs2	dominant	9	14F/14R	snp	ACAAGTCACATTATTCAGATGCAGA	GGTCACATATCCAATGTGTTCATAA	Truong *et al.*, 2011
Bacterial Spot	Bs2	dominant	9	25-1	snp	INNER:AATCTTTCTGGCCTTGAACTCCTATC OUTER:CAATACAAAAGATATCACATCTCCTTGGA	INNER:AGAGTTTTTACGATTCAGATGAATATTTGC OUTER:CTTGTATCTGTCATTGTGTGTGTTCTCA	Truong *et al.*, 2011
Bacterial Spot	Bs2	dominant	9	25-2	snp	INNER:AATCTTTCTGGCCTTGAACTCCTATC OUTER:GAGTTTTTACGATTCAGATGAATATTTGC	INNER: ATAATCATCAAACTCGATCACAATCAAC OUTER: CTTGTATCTGTCATTGTGTGTGTTCTCA	Truong *et al.*, 2011
Bacterial Spot	Bs3	dominant	2	PR-Bs3	InDel	GCACACCCTGGTTAAACAATGAACACG	GATGATAACTTGAAGTTGTGAGGATGG	Römer *et al.*, 2010
Bacterial Spot	Bs3	dominant	2	kasp Bs3	kasp	FAM:GATAACTTGAAGTTGTGAGGATGGTTT HEX:GATAACTTGAAGTTGTGAGGATGGTTA-	AACAATGAACACGTTTGCCTGACCAATTT	Holdsworth *et al.*, 2015

Pathogen	Gene/Qtls	Resistance	Chromosome	Associated marker	Marker type	Forward primer	Reverse primer	References
Potyvirus	Pvr4/Pvr7	dominant	10	Pvr4-EI	caps	CCTGGAAAAGCTATATATCTATTCGGACAT	TGGGGGAAGGTAAAAGGAAGCTAA	Venkatesh *et al.*, 2018
Potyvirus	Pvr4/Pvr7	dominant	10	SNP-H2.3-F	snp	GGTCCGCTGATCAATTTTGG	GTTGTCCCAAGGAGGTTGGT	Venkatesh *et al.*, 2018
Potyvirus	Pvr4/Pvr7	dominant	10	SNP-H1.7	snp	TTGGCCCAAAGTCCATTTCTT	TCAGTGGAACACGTCAGGCA	Venkatesh *et al.*, 2018
Potyvirus	Pvr4/Pvr7	dominant	10	CSO	caps	CGAAGAGAGAAGGTC	TCAGGGTAGGTTATT	Caranta et al.,1999, Özkaynak *et al.*, 2014
Potyvirus	Pvr4/Pvr7	dominant	10	SCUBC191423	scar	GCCCCGGTTTATATATTACGAAAAGA	AATGGAGAAGCATAATGACGGAGA	Arnedo *et al.*, 2002
Potyvirus	Pvr1/Pvr2	recessive	4	pvr1	snp	N.A.	N.A.	Kim *et al.*, 2017
Potyvirus	Pvr1/Pvr2	recessive	4	kasp pvr1	kasp	FAM:TGAAACAATGTAAGTCTGCTCT HEX:CATTCATGGACTTTCTGGTTTGATAATC	ATAATATCCACCACCCAAGCAAGTTAGTT	Holdsworth *et al.*, 2015
Potyvirus	Pvr1/Pvr2	recessive	4	caps pvr1	caps	ACGTTTGATGAAGCTGA GAAGGTGA	AACTTTGGACGTGCACAAG CAGAC	Holdsworth *et al.*, 2015
Potyvirus	Pvr1/Pvr2	recessive	4	Pvr1-S/pvr1-R1	caps	GCTAATGAGGCAGATGATGAAGTTG	CAACCATAAATATACCCCGAGAAT	Yeam *et al.*, 2005
Potyvirus	Pvr1/Pvr2	recessive	4	pvr1-R2	caps	GGGCTAAAATACGCTCATCTCCCTTC	GGCTCAATTTTATGCTTGAAACAATGTAAGC	Yeam *et al.*, 2005
Potyvirus	Pvr1/Pvr2	recessive	4	caps PVR2	caps	AAAAGCACACAGCACCAACA	TATTCCGACATTGCATCAAGAA	Ruffel *et al.*, 2006
Potyvirus	pvr6	recessive	3	caps PVR6	caps	ATGGCCACCGAAGCACCACCACCGG	ACACGGTGTATCGGCTCTTAGCT	Ruffel *et al.*, 2006

Table 15. Pathogens, resistance genes/QTLs, molecular markers and primer pair identified

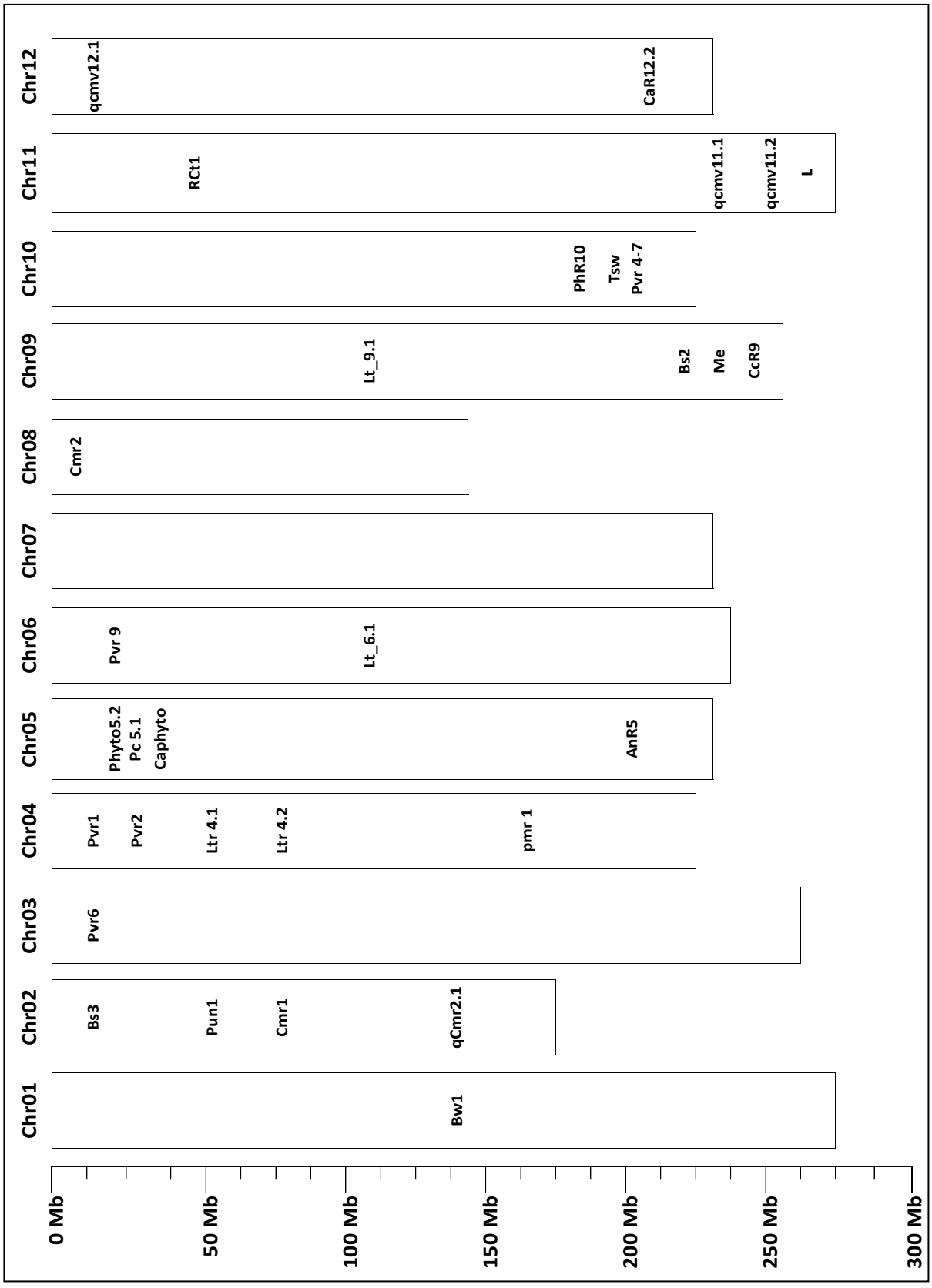


Figure 41. Position of disease resistance genes or QTLs on pepper genome

4.3.2 Homology search: primer sequences and polymorphism verification

In order to verify the retrieved primer sequences, homology search was performed against available genome sequences, through BLASTn tool in NCBI and SGN databases.

For 32 molecular markers, BLASTn homology searching, did not provided a perfect match between identified sequences of primer pairs and none of the available genome sequences therefore excluded from successive analysis as less reliable.

Perfect homology with 100% of alignments was found for 44 selected molecular markers.

In silico confirmation allowed the definition of 42 molecular markers (SSR, SCAR, InDel and CAPS) to be used directly in laboratory for the verification of polymorphic variants on genotypes which constitute the pepper collection (Tab. 16).

Gene/Qtls	Molecular marker	Marker type	Primer name	Labeled name	Primer sequence	Notes
Me1	16880-1-V2	scar	N16880-1-V2F	N-1-V2F	TGACCCCTCAGACTGAACAG	Expected band of 440 bp for resistant genotypes
			N16880-1-V2R	N-1-V2R	CTCCTTCGCTGCTACCTTCT	Expected band of 500 bp for susceptible genotypes
Me1	16830-CAPS	caps	N16830-CAPSF	NCAPSF	ATGACCATCTGGAAATCAGGTGT	Expected band of 300 bp for resistant genotypes
			N16830-CAPSR	NCAPSR	TCATGACGGTTAATAGGATTTGCA	Expected band of 260 bp for susceptible genotypes
Me1	S24	ssr	NS24F	NS24F	AGCAAAACGGTAGCACAACA	Expected two bands of 200 bp and 220 bp for resistant genotypes
			NS24R	NS24R	TACTATCTCCTGAGCCGGGG	Expected band of 210 bp for susceptible genotypes
N	SSR37	ssr	NSSR37F	NSSR37F	ACATACCCAAAAACTCTCTCAC	Expected band of 100 bp for resistant genotypes
			NSSR37R	NSSR37R	GATTGACCATGTTTCCGTAT	Expected band of 150 bp for susceptible genotypes
Me3, Me4, Me7	SCAR_PM6a	scar	NSCAR_PM6aF	NSCAR6aF	TTCTTACCCCTGTACATCACATCCT	Expected band of 146 bp for resistant genotypes
			NSCAR_PM6aR	NSCAR6aR	AACCTGGAAATTTCTGGAGGTATG	Expected band of 190 bp for susceptible genotypes
Me1	16839-H-V2	snp	N16839-H-V2F	N-H-V2F	CAAGGAAAACCAAACCAG	Expected sequence lenght 548 bp. To test with sequencer
			N16839-H-V2R	N-H-V2FR	CTCGCCTGTTATTTACCAAG	
Me1	K32	kasp	NK32F	NK32F	GCATTCAAGGCTTAGTTCTG	Expected sequence lenght 548 bp. To test with sequencer
			NK32R	NK32R	TGTGCTTCAAACATACCAAC	
Me3, Me7, N	SCARN	scar	NSCARNF	NSCARNF	AATTCAGAAAAAGACTTGGAAGG	Expected band of 315 bp for resistant genotypes
			NSCARNR	NSCARNR	TAAAGGGATTCATTTTATGCATAC	Expected band of 331 bp for susceptible genotypes
Pvr4	CSO	caps	PVR4CSOF	PVR4CSOF	CGAAGAGAGAAGGTC	Expected band of 444 bp for resistant genotypes
			PVR4CSOR	PVR4CSOR	TCAGGGTAGGTTATT	Expected band of 458 bp for susceptible genotypes
Pvr7	InDel (Pvr4-EI)	InDel	PVR7InDelF	PVR7InDelF	CCTGGAAAAGCTATATATCTATTCGGACAT	N.A.
			PVR7InDelR	PVR7InDelR	TGGGGGAAGGTAAAAGGAAGCTAA	
Pvr4	SCUBC19-1	scar	PVR4SCUBC19-1F	PVR4SCUF	GCCCGGTTTATATATTACGAAAAG	No band for resistant genotypes
			PVR4SCUBC19-1R	PVR4SCUR	AATGGAGAAGCATAATGACGGAGA	Expected band of 1423 bp for susceptible and heterozygous genotypes
Tsw	scac568F	caps	tsvscac568F	tsvscacF	GTGCCAGAGGAGGATTTAT	Expected band of 568 bp for resistant genotypes
			tsvscac568R	tsvscacR	GCGAGGTGGACACTGATACT	Expected three bands of 568 bp and 320 bp and 220 bp for susceptible genotypes
Pmr1	ZL1_1826	scar	PMZL1_1826F	PMZLF	CGAAGTCATTAAAGTTCATTGGG	Expected band of 1259 bp for resistant genotypes
			PMZL1_1826R	PMZLR	GCAATAAATGCCCTTCCACA	Expected band of 1070 bp for susceptible genotypes
Pmr1	HRM4.1.6	snp	PMHRM4.1.6F	PMHRM4F	AATTAAAGGACTTAAGTTTGACAGTT	To test with sequencer primer HRM
			PMHRM4.1.6R	PMHRM4R	GAAATTGTCGATGAACATCCGT	
Pmr1	HRM2.A4	snp	PMHRM2.A4F	PMHRM2F	AAGGGAAAGCAAGAAGTG	Expected sequence lenght 569 bp. To test with sequencer primer HRM
			PMHRM2.A4R	PMHRM2R	TGTTTTTTGGATCTGAAGAG	
Pmr1	CZ2_11628	caps	PMCZ2_11628F	PMCZ2F	GCTAGGATCCTGCTCGTGAGA	Expected bands of 166 bp or 174 bp for resistant genotypes
			PMCZ2_11628R	PMCZ2R	GTTGCTCTTGCTTCTGCTGC	Expected band of 340 bp for susceptible genotypes
L4	L-V0-4	scar	TVL-V0-4F	TVL-V0F	TATCGATGCACCCCTCGTTTTAAT	Deletion of 34 bp
			TVL-V0-4R	TVL-V0R	ATCTACCACAATGGCAGTGACGAA	
L4	L4-SCAR	scar	TVL4-SCARF	TVL4F	ATCGATGCACCCCTCGTTTTAATC	Expected band of 102 bp for resistant genotypes
			TVL4-SCARR	TVL4R	GAGCAGTGTGGAGTGTCTATTGCTCA	Expected band of 136 bp for susceptible genotypes
L4	087H3T7	caps	TV087H3T7F	TV087HF	CCTTTGCCTGCATTATTCTTG	Expected band of 440 bp for resistant genotypes
			TV087H3T7R	TV087R	GCCCAAATTTATTCCCAAATGC	Expected bands of 140 bp and 300 bp for susceptible genotypes
L4	L4-3	snp	TVL4-3F	TVL4-3F	TCCCTCCTCTTTTTCTAAGC	Expected sequence lenght 625 bp. To test with sequencer
			TVL4-3R	TVL4-3R	ACAGGCATTCACAGTCAAAC	
L3	PMFR11	scar	TVPMFR11F	TVPMFRF	CTGCAGAACAACAATGGCACG	Expected band of 269 bp for resistant genotypes
			TVPMFR11R	TVPMFRR	GGACTGCAGAGGAGGAAGC	Expected band of 263 bp for susceptible genotypes

Gene/Qtls	Molecular marker	Marker type	Primer name	Labeled name	Primer sequence	Notes
L3	L3-SCAR	scar	TVL3-SCARF	TVL3F	AACAATTTACAAATAATACACAAGGC	N.A.
			TVL3-SCARR	TVL3R	TTGGGAAGGAAAGACATCAT	
L3	A339-NK	scar	TVA339-NKF	TVA339F	TCTCGGTAGGCCATTTTGCT	N.A.
			TVA339-NKR	TVA339R	GTAAGTTGCTATGCCCACCA	
Bs2	14F/14R	snp	Xs14F/14RF	Xs14RF	ACAAGTCACATTATTCAGATGCAGA	Expected band of 500 bp for resistant genotypes
			Xs14F/14RR	Xs14R	GGTCACATATCCAATGTGTTCATAA	Expected band of 600 bp for susceptible genotypes
Bs3	KASP_Bs3	kasp	XsKASP_Bs3F	XsKASPF	TTTCTGGTTCAGTTTATGGG	Expected sequence lenght 482 bp. To test with sequencer
			XsKASP_Bs3R	XsKASPR	TAGATTTAGCGGGTGACAAG	
Bs3	PR-Bs3	scar	XsPR-Bs3F	XsPRF	GCACACCCTGGTTAAACAATGAACACG	Expected band of 100 bp for resistant genotypes
			XsPR-Bs3R	XsPRR	GATGATAACTTGAAGTTGTGAGGATGG	Expected band of 115 bp for susceptible genotypes
cmr1	CAPS A	caps	CMV CAPS AF	CMVAF	GTAGTAGGGTACGGACTCATA	Expected band of 1000 bp
			CMV CAPS AR	CMVAR	GTCCCGACGATAGCCCAAAAG	
cmr1	CAPS B	caps	CMV CAPS BF	CMVBF	GTAGTAGGGTACGGACTCATA	Expected band of 1200 bp
			CMV CAPS BR	CMVBR	GGAGTTTCATCATATGAAGCC	
cmr1	caTm-int1	caps	CMVcaTm-int1F	CMVcaTmF	TCAGCAAAGAAAGATTCACGAAC	Expected band of 450 bp for resistant genotypes
			CMVcaTm-int1R	CMVcaTmR	ACGTACACTTGATGATGCCTTGT	Expected bands of 300 bp and 150 bp for susceptible genotypes
cmr1	CaTm-int3-HRM	snp	cmvCaTm-int3-HRMF	cmvCaHRMF	GAAGTTTTCTTCTTATTTTGCC	Expected sequence lenght 558 bp. To test with sequencer
			cmvCaTm-int3-HRMR	cmvCaHRMR	TGAGCCAGTATATTCCATTTC	
cmr1	CaT1616BAC	snp	cmvCaT1616BACF	cmvCaTF	GAACAACGATTTCTGACAAG	Expected sequence lenght 575 bp. To test with sequencer
			cmvCaT1616BACR	cmvCaTR	CTGAAGGTTTGAAGGAATG	
cmr2	IBP160	snp	cmvIBP160F	cmvIBPF	GAAGAGTTCAAGCAACATAGG	Expected sequence lenght 678 bp. To test with sequencer
			cmvIBP160R	cmvIBPR	GAAGCCAAGAATCTGATACAG	
cmr2	AFLP	snp	cmvAFLPF	cmvAFLPF	CTTCCAAGCCTTTTGTGC	Expected sequence lenght 632 bp. To test with sequencer
			cmvAFLPR	cmvAFLPR	TGACTTCCCTCAACCTCTTAC	
Phyto.5.1 (QTL)	NBS1-CAPS	caps	PhNBS1-CAPSF	PhNBSF	AGGACTTTGATAAGGTTTC	Expected band of 72 bp for resistant genotypes
			PhNBS1-CAPSR	PhNBSR	TGCAATATAGAGCTTCTGCTG	Expected band of 97 bp for susceptible genotypes
Phyto.5.2 (QTL)	CAMS420	ssr	PhCAMS420F	Ph420F	CACGACGTTGTAAAACGACCAGCGTTCTATCGTCTCAAATG	Ordered with M13. To test with sequencer
			PhCAMS420R	Ph420R	TTGACAAACCAGAAATTGATCG	
PhR10 (dominant gene, Phyto.5.2 QTL)	P52-11-21	ssr	PhP52-11-21F	PhP21F	CACGACGTTGTAAAACGACCAATCCAAACAAGTCCTAAG	Ordered with M13. To test with sequencer
			PhP52-11-21R	PhP21R	GGTGCAATTGAAAATCTAAG	
PhR10 (dominant gene, Phyto.5.2 QTL)	P52-11-41	ssr	PhP52-11-41F	PhP41F	CACGACGTTGTAAAACGACTTGATGAGATGGGAAGTAAA	Ordered with M13. To test with sequencer
			PhP52-11-41R	PhP41R	CACCAACAATAATAGAACTACA	
Phyto.5.2 (QTL)	ZL6726	ssr	PhZL6726F	PhZLF	CACGACGTTGTAAAACGACTCCAGCCATCCATTATTTCAT	Ordered with M13. To test with sequencer
			PhZL6726R	PhZLR	ATCCCGAACTGCCAATAATTA	
Phyto.5.2 (QTL)	CA524065	ssr	PhCA524065F	PhCAF	CACGACGTTGTAAAACGACTCTCTCTCTACATCTCTCCGTTG	Ordered with M13. To test with sequencer
			PhCA524065R	PhCAR	TGTCGTTCGTCGACGTACTC	
Phyto.5.1 (QTL)	Phyto5NBS1	snp	Phyto5NBS1F	PhNBS1F	CATACAGCCAAAGTTAGAGC	Expected sequence lenght 489 bp. To test with sequencer
			Phyto5NBS1R	PhNBS1R	GCATAGAGTTCTCCCATTTC	
Phyto.5.1 (QTL)	Phyto5SAR	snp	Phyto5SARF	PhSARF	GGGCAGAAAGATTACAATGTC	Expected sequence lenght 515 bp. To test with sequencer
			Phyto5SARR	PhSARR	TTTTATTCTCACACCATACACG	
Phyto.5.1 (QTL)	P5-SNAP	scar allele specifico	PhP5-SNAPF1	PhP5F1	TCATGAGGTTGCTATTAAGATTGGTCCTGTTATATA	Specific allele three primers
			PhP5-SNAPF2	PhP5F2	GAGGTTGCTATTAAGATTGGTCCTGTTATCCG	
			PhP5-SNAPR	PhP5R	CATAGAAAGGGATATCATCTGGTACATGCAGAAA	

Table 16. Chosen molecular marker list and ordered primer

Evaluation of genomic DNA

Extracted genomic DNA of samples composing pepper collection were checked for quantity (concentration, ng/µl) and quality values (purity, absorbance ratio 260/280nm). For all samples were obtained a good DNA concentration as well as quality values.

4.3.3 PCR protocols fine tuning for molecular markers validation

Each molecular marker was tested on reference genotypes with known phenotype behaviour as well as allelic resistance composition (homozygous and heterozygous condition for resistance/susceptibility) in order to evaluate the potential use within the company breeding program. To confirm the marker discrimination capacity after initial step, when only few genotypes were used to verify the presence of amplified fragments, their weight, readable pattern and optimize PCR conditions, large groups of resistant vs susceptible genotypes were tested. Optimized protocols are presented below:

- ***Nematodes* molecular markers**

Markers 16880-1-V2, 16830-CAPS, S24, SSR37, SCAR_PM6a and 16839-H-V2 were initially tested. From the results obtained, it was possible to observe for the marker SSR37, SCAR_PM6a and 16880-1-V2, polymorphic profiles between resistant genotypes therefore they were used for further analysis using a wide genotypes panel consisting of ten resistant genotypes (NR1 to NR10 homozygous resistant except NR2 and NR3 as heterozygous) and two susceptible genotypes (NS1-NS2). Selected markers were amplified by PCR, separated using 2% agarose gel electrophoresis at 80V for about 2 hours.

Optimized amplification protocol:

Reaction mix (µL)	
MIX 1X	
ddH_2O	17,3
Buffer PCR (10X)	2,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24,0

PCR conditions			
Cycle	**Steps**	**Temp.**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	59 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 17. Reaction mix for *Nematodes* markers

Amplification results are reported below.

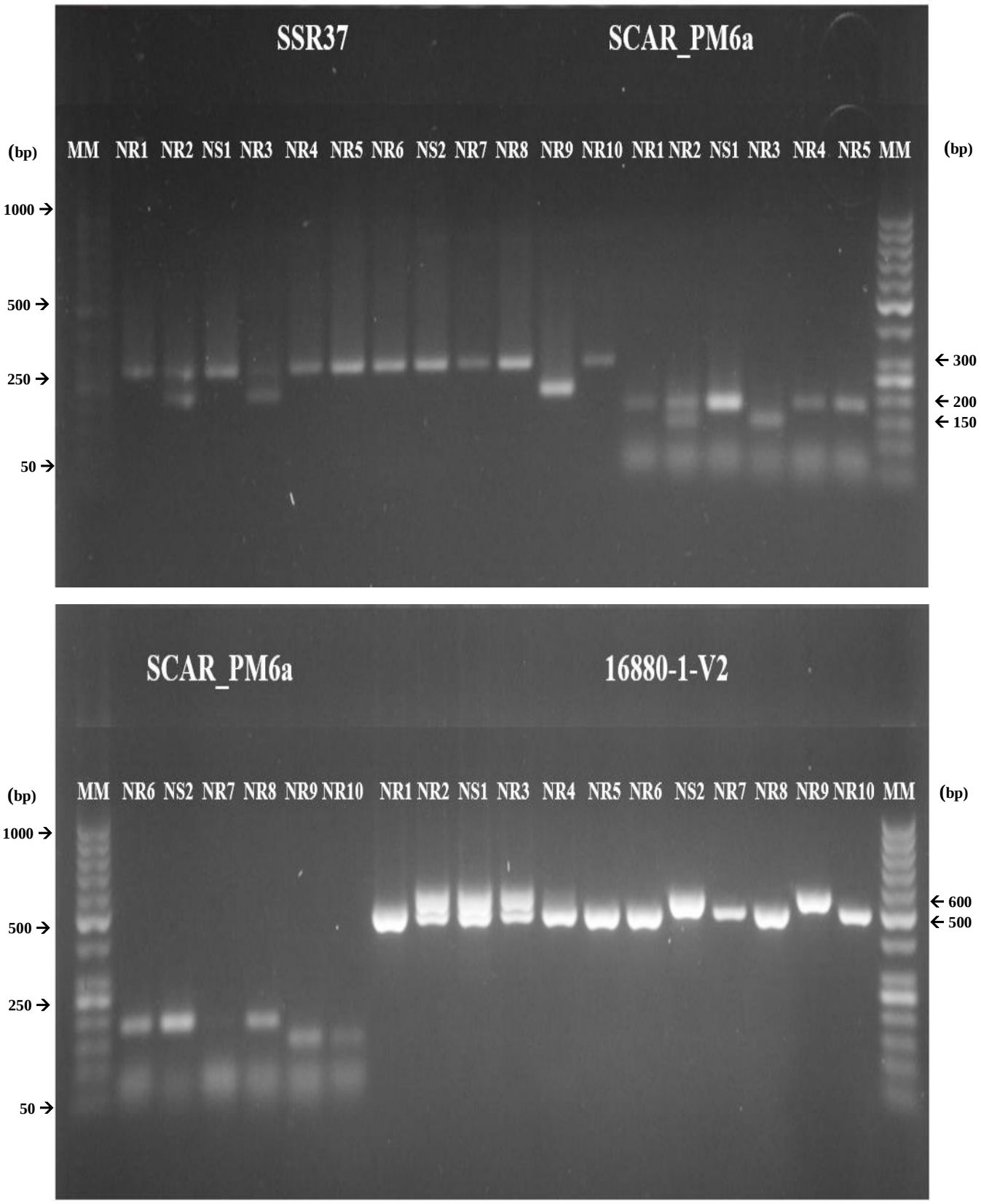


Figure 42. Amplification of SSR37, SCAR_PM6a 16880-1-V2 using wide panel for *Nematodes* genes. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder)

Three molecular markers are associated to different resistance genes to nematodes: SSR37 to *N*-gene, SCAR_PM6a to *Me3/Me4/Me7* and 16880-1-V2 to *Me1*.

Results, summarized in Tab. 18, show that the phenotypic field resistant could be due to different resistance genes or their combination such as resistance of NR9 (seem determines by N-gene and *Me/3Me4/Me7* genes) as well as NR10 (linked to cluster *Me/3Me4/Me7* and *Me1* genes). The other 5 genotypes shown to carry only one resistant genes at homozygous state.

Table below show identified polymorphisms related resistance gene associated to root knot nematodes.

ID	**Phenotype**	**Allele**	***N*-gene (SSR37)**	***Me3/Me4/Me/7* (Scar_PM6a)**	***Me1* (16880-1-V2)**
NR1	resistant	RR	0	0	1
NR2	resistant	H	2	2	2
NR3	resistant	H	2	1	2
NR4	resistant	RR	0	0	1
NR5	resistant	RR	0	0	1
NR6	resistant	RR	0	0	1
NR7	resistant	RR	0	0	1
NR8	resistant	RR	0	0	1
NR9	resistant	RR	1	1	0
NR10	resistant	RR	0	1	1
NS1	susceptible	SS	0	0	2
NS2	susceptible	SS	0	0	0

Table 18. Polymorphism result for SSR37, SCAR_PM6a and 16880-1-V2 associated to root knot nematodes (0= susceptible allele; 1= resistant allele; 2= heterozygous)

Overall, scar marker 16880-1-V2 was identified as most promising marker usable for marker assisted selection.

Comparing expected band from literature (resistant 440 bp, susceptible 500 bp) a difference has been found in amplified fragment size with 500bp band for resistant genotypes and 600 bp for susceptible genotypes.

- ***Potyvirus* molecular markers**

After a first PCR amplification for markers SCUBC19-1 and InDel (Pvr4-EI) with 3 known genotypes (two resistant PR1 - PR2 and one susceptible - PS1) showed no polymorphic bands for those markers.

Using the same PCR condition and 3 genotypes (two susceptible - PS1/PS2 and one resistant PR3) also marker caps CSO was tested and after a first selective stage the marker discriminative efficiency was checked on a larger group of genotypes.

In order to evaluate the marker effectiveness (*Pvr4* gene confer complete resistance to the different *Potyvirus* pathotype Pvy 0,1,2), samples list was composed with susceptible and resistant genotype (Pvy 0,1,2) including also materials with known resistance only for pathotype pvy:0.

A complete list of genotypes used for the validation of CSO marker and optimized PCR conditions are reported below:

ID	Phenotype Pvy:0,1,2 (Pvr4 allele)	Allele	Notes
PS1	susceptible	SS	Pvy:0 susceptible
PS2	susceptible	SS	Pvy:0 susceptible
PS3	susceptible	SS	Pvy:0 susceptible
PS4	susceptible	SS	Pvy:0 resistant
PS5	susceptible	SS	Pvy:0 resistant
PS6	susceptible	SS	Pvy:0 resistant
PR7	resistant	H	Pvy:0,1,2 resistant
PS8	susceptible	SS	Pvy:0 susceptible
PS9	susceptible	SS	Pvy:0 susceptible
PS10	susceptible	SS	Pvy:0 susceptible
PS11	susceptible	SS	Pvy:0 susceptible
PSR12	resistant	RR	Pvy:0,1,2 resistant

Table 19. Wide panel for CSO marker validation

Reaction mix (µL)	
MIX 1X	
ddH_2O	17,3
Buffer PCR (10X)	2,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24,0

PCR conditions			
Cycle	Steps	Temp.	Time
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	1’
	Annealing	53 °C	1’
	Extension	72 °C	1’30’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 20. Reaction mix for CSO marker

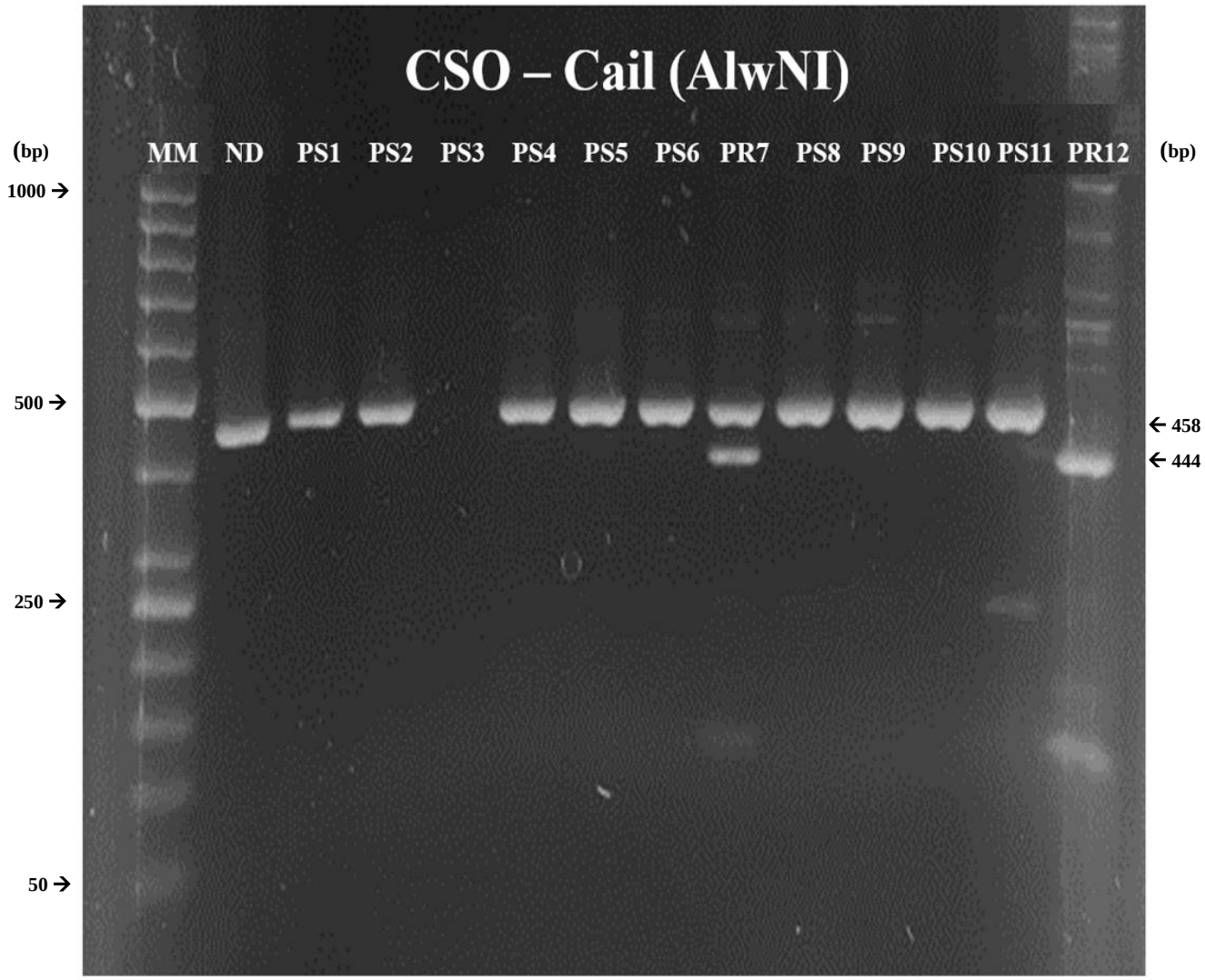


Figure 43. Amplification of CSO marker after digestion with restriction enzyme Cail (AlwNI). MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder), ND-amplified fragment non digested by restriction enzyme

As reported in literature, expected band for susceptible and resistant genotypes has been identified: 444 bp for resistant genotypes and 458 bp for susceptible.

From the results obtained, it was possible observe a perfect match between the phenotypic resistance and the marker response.

As codominant marker, CSO can distinguish homozygous and heterozygous genotypes associated to *Pvr4* gene while materials with resistance to only pathotype pvy:0 will be identifies with the same band of susceptible.

- ***Tomato spotted wilt virus*** **molecular markers**

Set up of PCR validation for molecular marker scac568F was conducted using 4 genotypes (one susceptible TS1, one resistant heterozygous TR2, two resistants homozygous TR3 - TR4) and three different restriction enzymes (XbaI, TaqI, HaeIII).

The optimized PCR protocol was as follows:

Reaction mix (µL)	
MIX 1X	
ddH_2O	17,3
Buffer PCR (10X)	2,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24,0

PCR conditions			
Cycle	**Steps**	**Temp.**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	53 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 21. Reaction mix for scac568F marker

Profiles obtained using this first pilot genotypes set, show different pattern generated from restriction enzymes activity: same profile was obtained for samples digested with XbaI and HaeIII while opposite cutting site related for TaqI.

Results show no clear correspondence between expected band and phenotype data but generated pattern indicate a cutting site on resistance genotypes using XbaI and HaeIII while opposite behaviour (cut susceptible) is reported for TaqI.

Using only XbaI as restriction enzyme, successive investigation concerned a wide panel composed by 12 genotypes:

- TS1 to TS4: susceptible, allele SS;
- TR5 to TR8 and TR11, TR12: resistant, allele RR;
- TR9 and TR10: resistant, allele H.

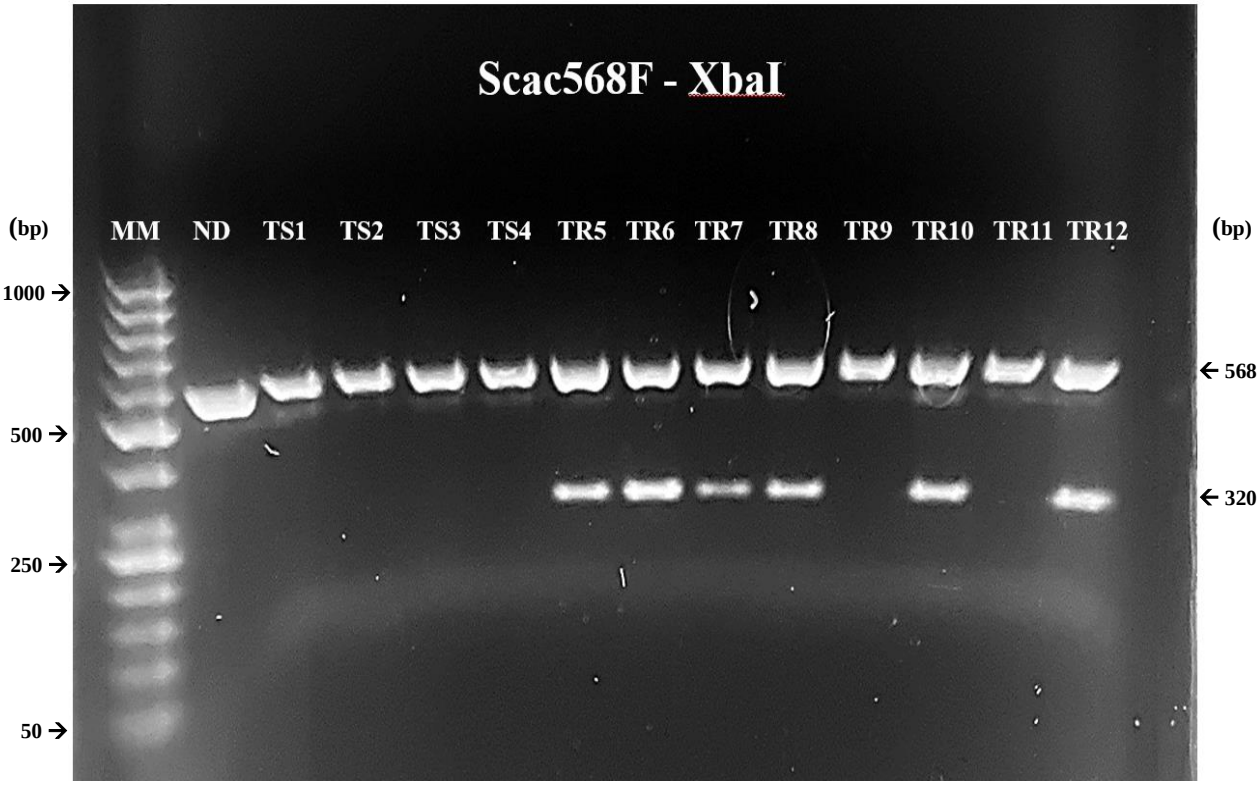
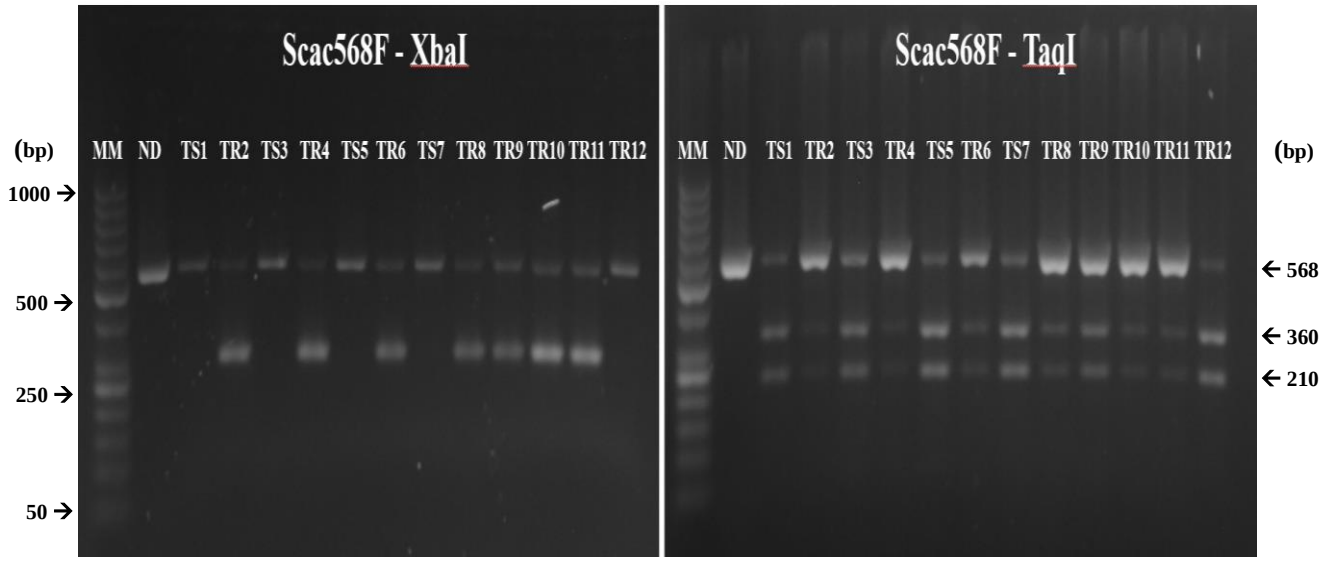


Figure 44. Amplification of Scac568F after digestion with restriction enzyme XbaI. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder), ND-amplified fragment non digested by restriction enzyme

Result show perfect match between phenotype data and generated bands except for samples TR9 and TR11. Generally it is possible hypothesize that primers link to aspecific fragment both on susceptible and resistant genotypes (common band on 560 bp) resulting a dominant marker. Using restriction enzyme XbaI and TaqI, further test was conducted with additional panel genotypes composed by another 12 genotypes set:

- TS1-TS3-TS5-TS7: susceptible, allele SS;
- TR2-TR4-TR6-TR8-TR10-TR11: resistant, allele RR;
- TR9-TR12: resistant, allele H.

Good amplification for 24 samples was obtained and subsequent digestion was performed as reported in figures 45 and 46.



Figures 45-46. Results of digested product with XbaI and TaqI. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder)

Confirming first results, restriction enzyme show a cutting site on resistant genotypes sequences for XbaI while on susceptible genotypes sequences for TaqI.

Generally marker result match the phenotype resistance data except for the sample TR12. Due to the amplification of common aspecific fragment for every genotype, marker it is not able to detect heterozygous condition.

Following identification bands are determined considering marker as dominant:

- XbaI susceptible: 568 bp;
- XbaI resistant: 568 bp and 320 bp;
- TaqI susceptible: aspecific bands 568 bp - 360 bp - 210 bp;
- TaqI resistant: main band 568 bp and faint aspecific band on 360 bp and 210 bp.

Further evaluation of data must consider origin of *tsw* gene and its introgression in cultivated variety. As reported in literature (Boiteux L.S, 1995) resistance dominant gene *tsw* derived from *C. chinense* and at least five different *C. chinense* possess a single dominant gene at *tsw* locus conferring hypersensitive resistance to *Tomato spotted wilt virus* (Moury *et al.*, 2000). For this reason, it is possible hypothesize that amplified sequence of marker scac568F may have been be introgressed from specific *C. chinense* accession explaining the susceptible profile of TR9 and TR11 as well as TR12 and assuming that these genotypes could have a different source of resistance.

- ***Leveillula taurica* molecular markers**

PCR amplification was carried out for ZL1_1826 using 12 genotypes:

- LTS1 to LTS3 and LTS9: susceptible, allele SS;
- LTR4: resistant, allele H;
- LTR5 to LTR8 and LTR10 to LTR12: resistant, allele RR.

The optimised PCR conditions were as follow:

Reaction mix (µL)	
MIX 1X	
ddH_2O	17,3
Buffer PCR (10X)	2,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24,0

PCR conditions			
Cycle	**Steps**	**Temperature**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	59 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 22. Reaction mix for *Leveillula T.* markers

Considering weight of expected bands (resistant 1,259 bp, susceptible 1,070 bp), amplified fragments were separated through gel electrophoresis (2% agarose). Resulting profiles are shown in the figure below.

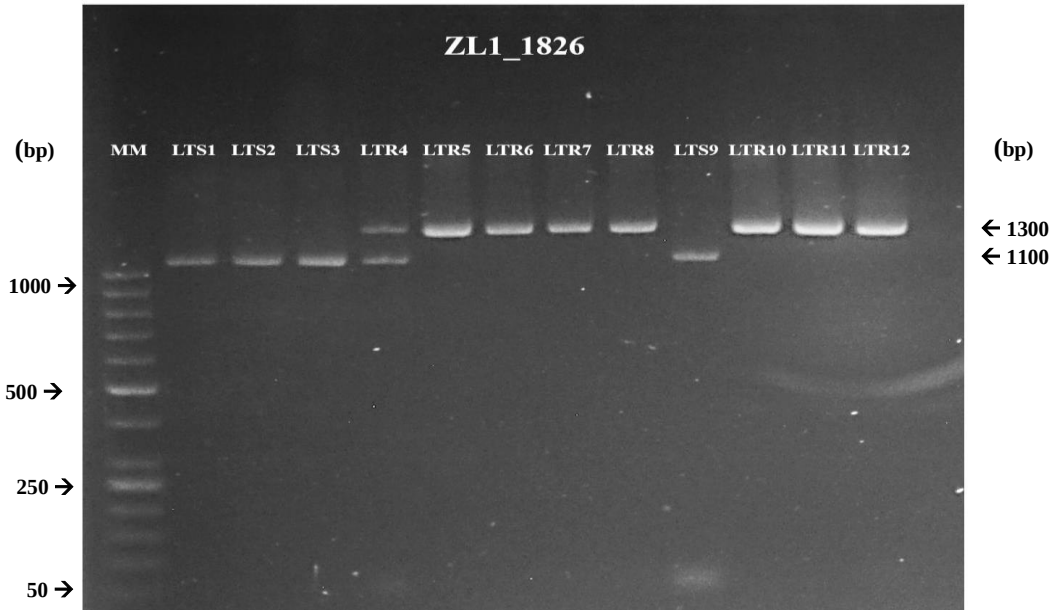


Figure 47. Amplified products of marker ZL1_1826. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder)

Results obtained show positive match between phenotype resistance and expected band with similar fragment size reported in literature (resistant 1,300 bp, susceptible 1,100 bp). Also heterozygous alleles are detected in sample LTR4.

Second molecular marker tested for PCR amplification was CAPS marker CZ2_11628 that included one restriction site recognized by TaqI, but the expected length of the band was not previously known from literature.

Same genotypes set of previous investigation as well as PCR protocol was used.

Resulting digested products are reported in the figure below.

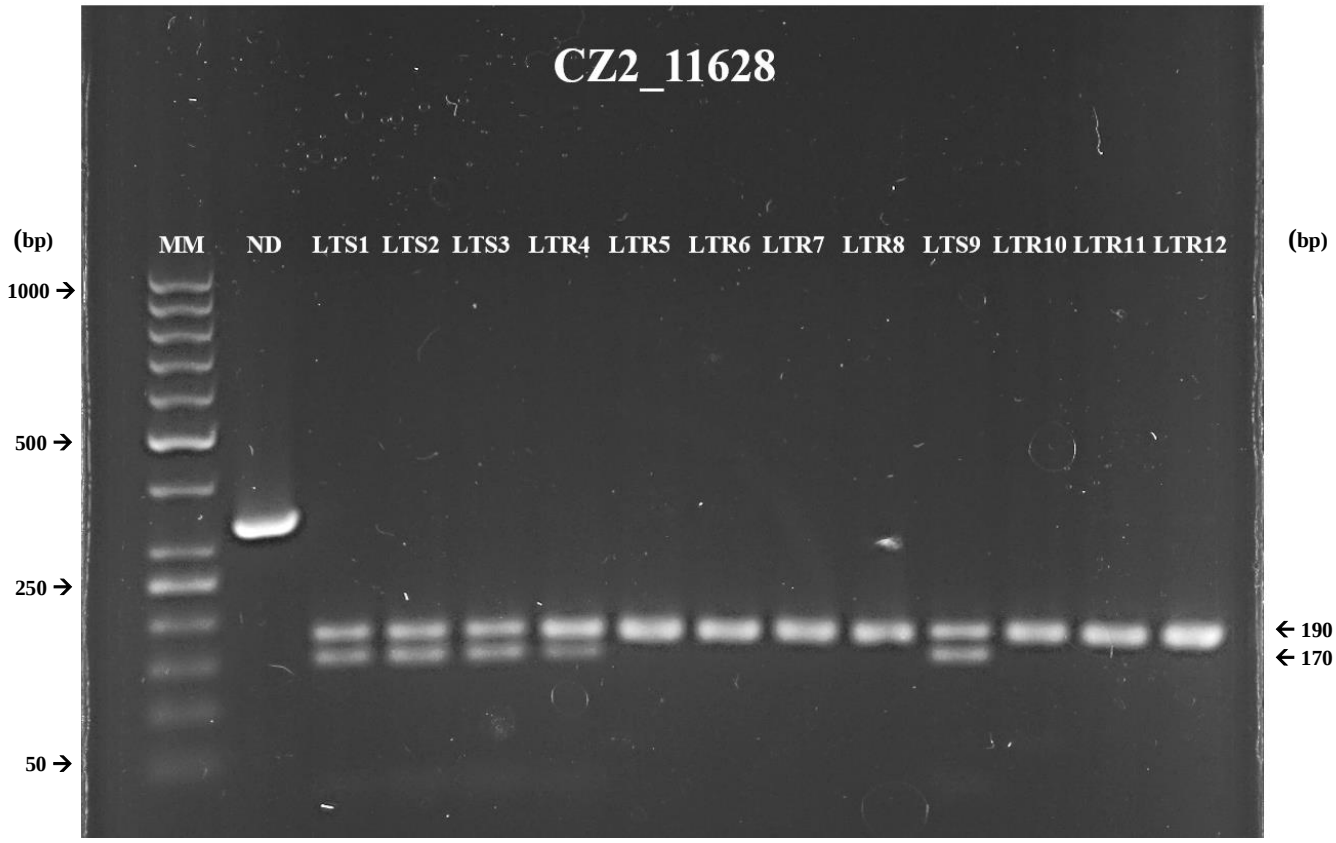


Figure 48. Digested products of marker CZ2_11628 using enzyme restriction TaqI. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder), ND-amplified fragment non digested by restriction enzyme

The results show the cutting site of TaqI for susceptible genotype but heterozygous are not be detected.

Considering also performance of scar ZL1_1826 and less time consuming, the marker CZ2_11628 was excluded.

- ***Tobamovirus (tm:0-3)*** **molecular markers**

L4-SCAR was tested using 10 genotypes of different allelic composition:

- L4-S1 to L4-S5 (susceptible, allele SS);
- L4-R6 to L4-R9 (resistant, allele H);
- L4-R10 and L4-R11 (resistant, allele RR).

Results showed no polymorphisms and therefore marker was rejected.

For marker L-V0-4, a set of following genotypes was tested:

- L4-S1 to L4-S4 and L4-S10 (susceptible, allele SS);
- L4-R5 to L4-R9 (resistant, allele H);
- L4-R11 (resistant, allele RR).

The optimized PCR protocol was as follows:

Reaction mix (µL)	
MIX 1X	
ddH_2O	13,3
Buffer PCR (10X) $MgCl_2$ free	5,0
$MgCl_2$ (1 mM)	1,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24

PCR conditions			
Cycle	**Steps**	**Temperature**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	59 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 23. Reaction mix for L4-SCAR and L-V0-4 markers

Figure 49 report the amplification results.

Figure 49. Amplified products of marker LV-04-SCAR after improved PCR protocol. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder)

Positive match between expected band and allele composition was founded. From literature only deletion of 34 bp is reported (Lee *et al*., 2012).

For LV-04-SCAR, results show following band size:

- Resistant allele: 750 bp;
- Susceptible allele: 800 bp.

That data match from literature (34 bp deletion, Lee *et al*., 2012)

- ***Tobamovirus (tm:0-2)*** **molecular markers**

Amplification for markers L3-SCAR and A339NK was set using 8 genotypes:

- L3-S1 and L3-S2 (susceptible, allele SS);
- L3-R3 and L3-R4 (resistant, allele H);
- L3-R5 to L3-R8 (resistant, allele RR).

However, no correspondence of obtained pattern and phenotypic data and therefore excluded for subsequent analysis.

A panel of 13 genotypes were tested for marker PMFR11.

Listed below the summary of samples loaded and successive amplified sequence resulted:

- L3-S1 to L3-S5 (susceptible, allele SS);
- L3-R6 and L3-R7 (resistant, allele H);
- L3-R8 to L3-R13 (resistant, allele RR).

Considering expected bands (269 bp for resistant genotypes and 283 for susceptibile genotypes) it was decided to separate the amplification products using gel electrophoresis (2% agarose) but no polymorphisms between tested genotypes was detected.

- ***Bacterial spot*** **molecular markers**

Molecular marker tested for the detection of Bacterial spot (*Xanthomonas* spp.) was PR-Bs3 and 14F/14R. Both markers were initially evaluated using 17 genotypes. PrBs3 marker showed no polymorphism among tested genotypes and was thus eliminated. 14F/14R marker requested futher protocol optimization and final PCR mix was as follows:

Reaction mix (µL)	
MIX 1X	
ddH_2O	13,5
Buffer PCR (10X) $MgCl_2$ free	5,0
$MgCl_2$ (1 mM)	1,3
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24

PCR conditions			
Cycle	**Steps**	**Temperature**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	59 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 24. Reaction mix for L4-SCAR and L-V0-4 markers

The genotypes tested were following:

- B-S1 to B-S5 (SET-2), B-S1 to B-S4 (SET-3): susceptible, allele SS;
- B-R6 to B-R8 (SET-2), B-R5 to B-R7 (SET-3): resistant, allele H;
- B-R9 to B-S12 (SET-2), B-R8 to B-S12 (SET-3): resistant, allele RR.

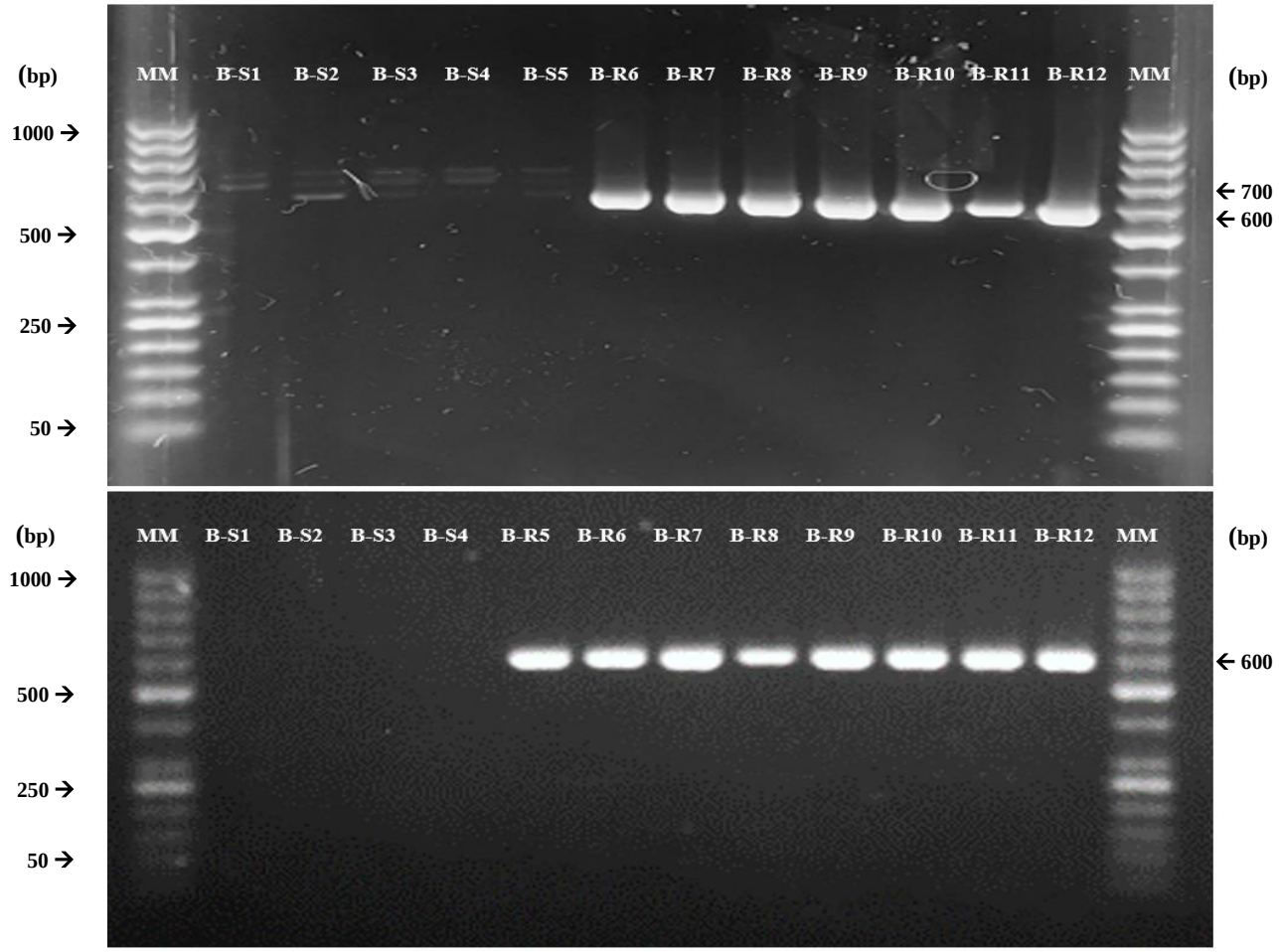


Figure 50. Amplified products of marker 14F/14R. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder)

It can be seen that three light band (600 bp - 650 bp - 700 bp) correspondent to susceptible genotypes while one major amplified product (600 bp) is well defined.

Third protocols (less $MgCl_2$ quantity) show absence of susceptible band confirming resistance allele with size dimension of 600 bp.

Generally, for both improved PCR protocols, positive match was found towards expected allele and fragment size is similar comparing literature. Marker response, as dominant, was:

- resistant allele: 600 bp;
- susceptible allele: aspecific 700 bp.

- ***Cucumber mosaic virus* molecular markers**

Ca-Tm-int1, CAPS A and CAPS B markers for *Cucumber mosaic virus* (CMV) were amplified using same DNA samples:

- C-S1 to C-S4: suscepbile, allele SS;
- C-R5 to C-R12: resistant, allele RR.

The optimized PCR protocol was as follows:

Reaction mix (µL)	
MIX 1X	
ddH_2O	13,3
Buffer PCR (10X) $MgCl_2$ free	5,0
$MgCl_2$ (1 mM)	1,5
dNTP (10 mM)	2,0
Primer forward (10µM)	1,0
Primer reverse (10µM)	1,0
Dream Taq DNA polymerase (5u/µL)	0,2
Total Volume	24,0

PCR conditions			
Cycle	**Steps**	**Temperature**	**Time**
1	Initial denaturation	95 °C	5’
30	Denaturation	95 °C	45’’
	Annealing	59 °C	45’’
	Extension	72 °C	45’’
1	Final extension	72 °C	7’
1	Final Hold	4 C°	overnight

Table 25. Reaction mix for CMV markers

After digestion an expected band of 450 bp for CaTm-int1 was observed but no polymorphic bands between susceptible and resistant parental genotypes, thus the marker was excluded from further analysis.

CAPS A: aspecific common band was observed for all samples with a polymorphism pattern distinctive between susceptible and resistant genotypes.

Digestion with XbaI generated bands of respective size 150 bp and 250 bp in susceptible genotypes, while a band of 350 bp was obtained for resistant ones.

Using XbaI restriction enzyme for CAPS B seem amplified a sequence of 1,500 bp with different cutting site comparing susceptible and resistant: fragments of 250 bp - 600 bp - 750 bp were observed for susceptible while 750 bp - 800 bp for resistant genotypes.

Positive results were thus obtained for both markers with a possible utilization for marker assisted selection program.

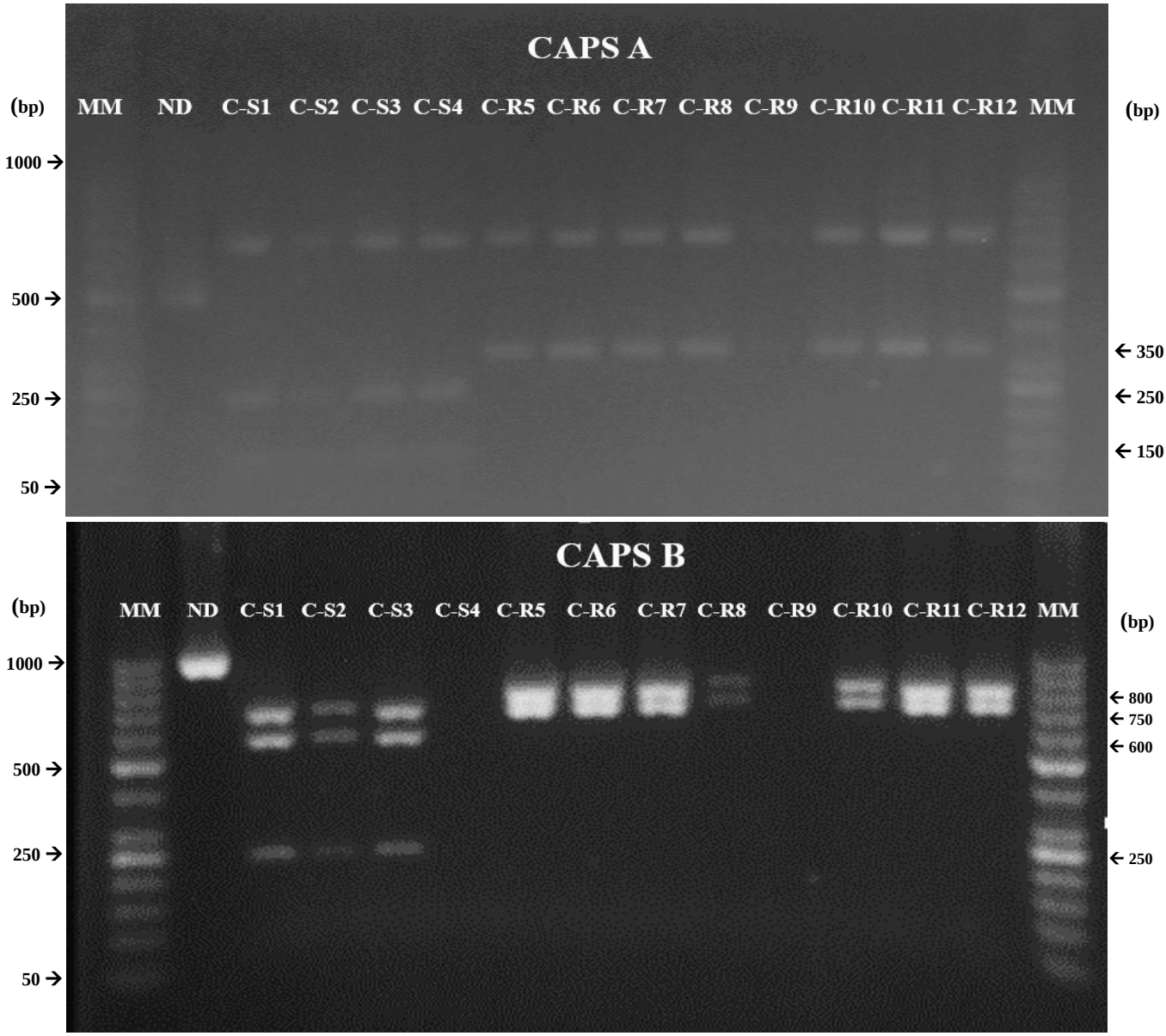


Figure 51. Amplified products of marker of CAPS A and CAPS B. MM-Molecular Weight Marker (GeneRuler 50 bp DNA ladder), ND-amplified fragment non digested by restriction enzyme

- ***Phytophthora capsici* molecular markers**

Markers NBS1-CAPS and P5-SNAP were tested on following genotypes:

- PH-S1 to PH-S5: susceptible, allele SS;
- PH-R6 and PH-R7: resistant, allele H;
- PH-R8 to PH-R11: resistant, allele RR.

No amplification was obtained for NBS1-CAPS, while for P5-SNAP no association with phenotype was observed.

4.4 Conclusions

In order to implement the company marker assisted selection activity, several molecular markers of different types associated to resistance genes for the most important diseases present in the Mediterranean area were identified through an accurate bibliographic search, taking into account also ease of use within the breeding program and therefore giving priority to main characters under monogenic control.

Generally, for each of them, a preliminary set-up of PCR protocols was performed in order to identify both most promising markers and appropriate PCR conditions using a restricted set of resistant and susceptible parental genotypes. Relevant validation on proprietary pepper collection was carried out using a wide set of genotypes with known phenotype expression previously characterize through specific bioassay and selective environments for high pathogenic pressure.

A series of markers were definitely discarded due to lack of polymorphisms or lower performance in term of reliability: S24 and SCARN for nematodes, InDel (Pvr4El) and SCUBC19-1 for *Potyvirus*, CZ2_11628 for *Leveillula taurica*, L4-SCAR/PMFR11/L3 SCAR/A339NK for *Tobamovirus*, PR-Bs3 for Bacterial Spot, CAPS B and CaTm-int1for *Cucumber Mosaic Virus*, P5-SNAP and NBS1-CAPS for *Phytophtora capsici*.

Tab. 26 reports an overview of the main results obtained:

Genes investigated	Gene with positive response	Marker name	Marker type	Restriction enzyme	Marker result	Susceptible sequence lenght (bp)	Resistant sequence lenght (bp)	Notes
Pmr 1	*Pmr 1*	ZL1_1826	scar	-	codominat	1100	1300	
L4	*L4*	L-V0-4	scar	-	codominat	750	800	common aspecific 1000 bp
Bs2	*Bs2*	14F/14R	scar	-	dominant	700	600	susceptible aspecific band
Me1-Me3-Me7-N- *Me 4- Me2 - Mech 1 - Mech 2*	*Me-1*	16880-1-V2	scar	-	codominat	600	500	
Tsw	*Tsw*	SCAC-568	caps	XbaI	dominant	568	320+568	
Pvr4-PVR7	*Pvr-4*	CSO	caps	CaiI (AlwNI)	codominat	480	420	
Cmr1 *- Cmr2*	*Cmr 1*	CAPS A	caps	XbaI	codominat	150+250	350	common aspecific 700 bp
QTL 5.2 - QTL.5.1	*QTL.5.1*	need further PCR set-up / markers validation						
L3	*L3*	need further PCR set-up / markers validation						

Table 26. State of the art about main promising marker identified during screening activity

Is worth noting that, as far as markers associated to resistance genes, for *Nematodes* was emerged the possibility that phenotypic field resistance could be determined both by one single *Me1 gene* (scar marker 16880-1-V2 has been identified as most promising marker for the proprietary collection to be used in future breeding programs) but also by the combination of different resistance loci: combination of *Me1* or *N-gene* with cluster *Me/3Me4/Me7* genes for which associated markers have been identified and tested.

Overall, using the available germplasm collection, the most promising molecular markers for associated resistance to pepper pathogens were identified and validated for effective reliability and related potential utilization in the company breeding program. Positive responses and further integrative activity concerned the improvement of PCR set-up protocol for the most strategic markers. Molecular multi-gene identification useful for marker-assisted breeding has been applied directly on pre-breeding crossed toward gene pyramiding approach allowing simultaneously the beginning application of new innovative breeding scheme. This is of great interest for the Company.

Edward Henry Weston (1886-1958)

Chapter 5

Elite material re-sequencing

5.1 Introduction

In recent years, a rapid increase in genomics has enabled the implementation of novel approaches toward the understanding of the molecular mechanisms underpinning resistance to pathogens. As next-generation sequencing (NGS) costs continue to decline, testing for novel biomarkers will increasingly spread to minor species and small breeding companies. In fact, next-generation sequencing technologies can be applied to provide whole-genome sequencing and to develop high-throughput molecular markers for QTLs mapping and gene discovery. Breeding programs are benefiting from these signs of progress in terms of precision and speediness to achieve results. Indeed, conventional molecular approaches are laborious and time-consuming. NGS-based genotyping produces instead thousands of single nucleotide polymorphism which allow detecting loci involved in the resistance to pathogens narrowing down the regions underlying genes of interest. Reduced representation, sequencing method for genotyping such GBS (genotyping by sequencing) or RADseq (restriction site-associated DNA sequencing) allowing high throughput genome scans at relatively low cost (Tamisier *et al.*, 2020; Jo *et al*., 2017). These NGS technologies, therefore, can be used to generate diagnostic markers able to detect the allelic variation within resistance genes. In addition to genomic-based breeding, NGS can be applied to unravel the diversity of the genome sequences (Kiran *et al.*, 2017). The future of NGS is shifting toward whole-genome sequencing, allowing to resolve key questions related to the genes function as well as the mechanism of resistance. WGS (Whole-genome shotgun) provides a more comprehensive, accurate and efficient tool for plant and animal genomics projects that traditionally have used legacy technology such as genotyping arrays or genotyping-by-sequencing. The WGS approach outperforms these methods by providing an order of magnitude more data, greater statistical power and enhanced variant discovery capabilities, easier analysis all at a lower cost. The variation information such as Single Nucleotide Polymorphism (SNP), Insertion and Deletion (InDel), Copy Number Variation (CNV) and structural variation (SV) obtained through resequencing is used in population genetics research and genome-wide association studies (GWAS) to investigate the causes of diseases, to select plants and animals for agricultural breeding programs and to identify common genetic variations among populations given also a very low cost at which, such analysis can be done nowadays on service, making it affordable to different segments of science and market.

Plant whole genome resequencing projects are carried out as follows:

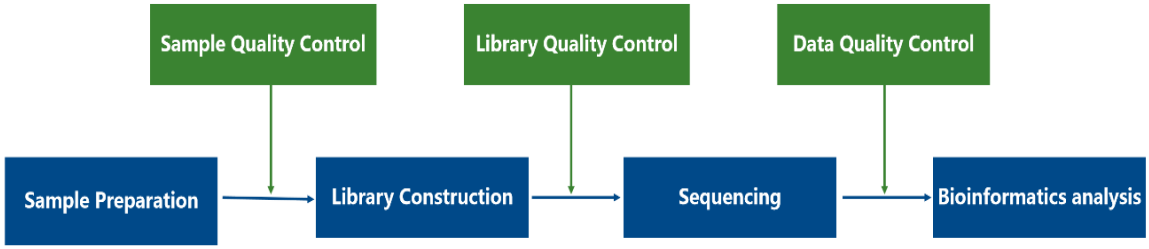


Figure 52. Workflow of resequencing project

In this chapter, we describe the re-sequencing activity of elite material identified through the investigation carried out in the previous work-packages.

Thoroughly characterization of these promising progenitors for future breeding programs determined the study of core collection genetic distances as well the identification of a set of private SNPs for each genotype analyzed, that distinguish them one from each other.

5.2 Materials and methods

5.2.1 Sequencing pipelines

Five genotypes from the company collection (described in Chapter 2), codified as PM, TSW, L3, L4 and CM (due to resistance genes to Powdery Mildew, *Tomato spotted wilt virus*, *L3* and *L4* for *Tobamovirus*, and *Cucumber mosaic virus* they harbor respectively) were chosen. The genomic DNA extracted as described in Chapter 3, was sequenced on service by Novogene (UK) Company Limited on NovaSeq 6000 Illumina Platform (Pair ends 150 bp) according to the following work-flow:

1. Sample Quality Control

(1) DNA degradation and contamination were monitored on agarose gel electrophoresis analysis;

(2) Qubit® DNA concentration was measured using Qubit® DNA Assay Kit in Qubit® 2.0 Flurometer (Life Technologies, CA, USA).

2. Library Construction, Quality Control and Sequencing

The genomic DNA was randomly sheared into short fragments. The obtained fragments were end repaired, A-tailed and further ligated with Illumina adapter.

The fragments with adapters were PCR amplified, size selected and purified following the workflow of library construction (Fig. 53).

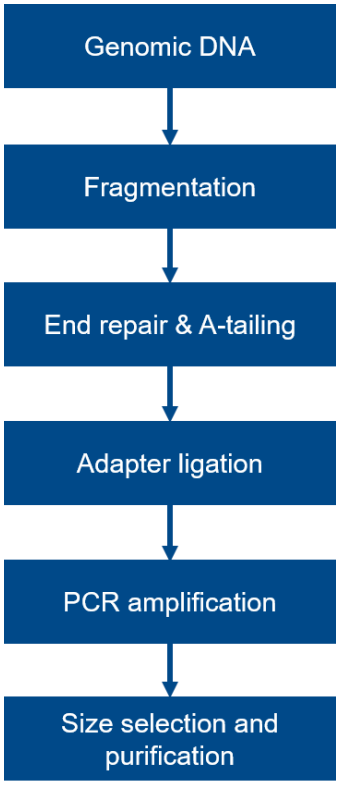


Figure 53. Experimental procedures of library preparation

The library was checked with Qubit and real-time PCR for quantification and bioanalyzer for size distribution detection.

Quantified libraries were pooled and sequenced on NovaSeq 6000 Illumina Platform (Pair ends 150 bp) platform, according to effective library concentration and data amount required.

3. Data quality control

Filtering: the sequencing reads/raw reads often contain low quality reads or reads with adaptors, which will affect the quality of downstream analysis. To avoid this, it is necessary to filter the raw reads and obtain clean reads. Raw reads filtering was as follows:

1) remove the paired reads when either read contains adapter contamination;

2) remove the paired reads when uncertain nucleotides (N) constitute more than 10 percent of either read;

3) remove the paired reads when low quality nucleotides (base quality less than 5, $Q \leq 5$) constitute more than 50 percent of either read.

Quality Distribution: for next-generation sequencing (NGS), the sequencing platform, chemical reactants and sample quality can influence sequencing quality and base error rate. Sequencing quality distribution is examined over the full length of all sequences to detect any sites (base positions) with an unusually low sequencing quality, where incorrect bases may be incorporated at abnormally high levels.

Sequencing Error Rate: sequencing error rate and base quality can be affected by various factors such as sequencing platform, chemical reagent and sample quality. Due to the consumption of chemical reagents, error rate is increasing with read extension, which is a common feature of Illumina high throughput sequencing platforms.

4. Bioinformatics Analysis Pipeline

Bioinformatic analysis with a well-annotated reference genome was performed following the workflow:

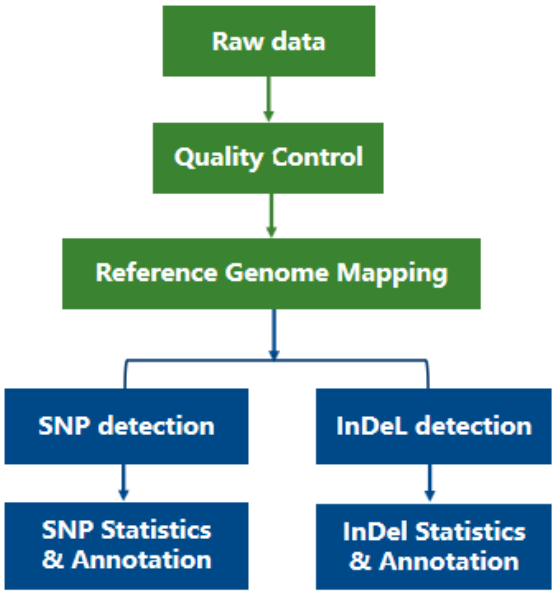


Figure 54. Bioinformatics analysis pipeline

Genetic distances: only polymorphisms restricted to exonic regions and present in all biological replicates were used for further analysis (due to abundance of data). Genetic distances were calculated using PAST software and hierarchical clustering was performed.

Private Polymorphisms evaluation: private (occurring only in one genotype) polymorphism were then inferred, limiting analysis only to polymorpfims inside coding regions that cause stop gain or stop loss. Based on their distribution, SNPs are considered private when are find in a single or a range of samples of a collection and absent in the rest, due to mechanisms such adaptation or as human-based selection.

Even if private SNPs are no reported to be crucial for determining major changes of plant phenome, their investigation could highlight the existence of potential candidate genes for traits of agricultural interest or be applied for a precise genotype identification /traceability/purity. As the pepper genome annotations are not still in a very advances phase (automatic annotation) and new versions still bring to quite profound differences from one version to another.

All polymorphisms identified within putative CDS were controlled manually by search in online databases and confronting pepper genome sequences available at the Sol Genomic Network website in order to eliminate the "false" gene that were not identified as coding sequences by other analysis.

5.3 Results

Raw Data: original image data file from high-throughput sequencing platforms (like Illumina) was transformed to sequenced reads (called Raw Data or Raw Reads) by CASAVA base recognition (Base Calling). Raw data are stored in FASTQ (.fq) format files, which contain sequencing reads and corresponding base quality.

Data Quality Control: the data was filtered, sequencing error rate was evaluated as far as sequencing quality distribution is regarded: if the sequencing error rate is represented by e, and Illumina sequencing quality by Q_{phred}, the quality score of a base (Phred score) is calculated by the following equation: $Q_{phred}=-10\log_{10}(e)$. The correspondence relationship between Illumina sequencing quality and Phred score in base calling by Casava version 1.8 is listed as follows:

Phred score	Error Rate	Correct Rate	Q-score
10	1/10	90%	Q10
20	1/100	99%	Q20
30	1/1000	99.9%	Q30
40	1/10000	99.99%	Q40

Table 27. Sequencing error rate and corresponding base quality value

According to the sequencing feature of Illumina platforms, for paired-end sequencing data we require that Q30 (the percent of bases with phred-scaled quality scores greater than 30) should be above 80%. Statistics of sequencing data are listed in Tab. 28.

Sample name	Raw reads	Raw data (G)	Clean data (G)	Effective (%)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
PM_1	26312405	7.9	7.9	99.71	0.03	97.21	92.32	34.88
PM_2	35440633	10.6	10.6	99.76	0.03	97.01	91.68	34.82
L4_1	26823127	8.0	8.0	99.75	0.03	96.84	91.30	34.97
L4_2	32953176	9.9	9.9	99.75	0.03	97.01	91.64	35.00
L3_1	35454848	10.6	10.6	99.77	0.03	96.23	90.00	35.00
L3_2	29820066	8.9	8.9	99.78	0.03	96.78	91.17	34.91
CM_1	31115011	9.3	9.3	99.78	0.03	96.83	91.24	34.93
CM_2	35796647	10.7	10.7	99.79	0.03	97.00	91.60	34.86
Tsw_1	35046456	10.5	10.5	99.77	0.03	97.06	91.73	34.88
Tsw_2	38118956	11.4	11.4	99.78	0.03	97.01	91.62	34.82

Table 28. Statistics of sequencing Data

The details for the sequencing data statistics are as follows:

(1) Sample name.

(2) Raw reads: the number of sequencing reads pairs; four lines will be considered as one unit according to FASTQ format.

(3) Raw data (G): the original sequence data volume.

(4) Clean data (G): the sequence data volume calculated by clean data.

(5) Effective (%): the ratio of clean data to raw data.

(6) Error (%): overall error rate of base.

(7) Q20 (%): the percentage of bases with higher Phred score than 20.

(8) Q30 (%): the percentage of bases with higher Phred score than 30.

(9) GC: the percentage of G and C in the total bases.

Sequence Alignment: the effective sequencing data was aligned with the reference genome through BWA software (Li H., and Durbin R., 2009) (parameters: mem -t 4 -k 32 -M), and the mapping rate and coverage were counted according to the alignment results.

Reference genome used was Zunla v.1 available at: ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/710/875/GCF_000710875.1_Pepper_Zunla1_Ref_v1.0. Its characteristics are reported in Tab. 29.

Seq number	Total length	GC content (%)	Gap rate (%)	N50 length	N90 length
1,627	2,935,884,163	34.94	3.23	220,335,243	153,299,543

Table 29. Statistics of Reference Genome

The detailed information of reference are as follows:

(1) Seq number: the total number of the assembled genomic sequences.

(2) Total length: the total length of the assembled genomic sequence.

(3) GC content: the GC content of the reference genome.

(4) Gap rate: the proportion of unknown sequence (N) in the reference genome assembly.

(5) N50 length: the length of scaffold N50, of which 50% of the sequence is higher than this level.

(6) N90 length: the length of scaffold N90, of which 90% of sequence is higher than this level.

The mapping rates of samples reflect the similarity between each sample and the reference genome. The depth and coverage are indicators of the evenness and homology with the reference genome.

Sample	Mapped reads	Total reads	Mapping rate (%)	Average depth(X)	Coverage at least 1X (%)	Coverage at least 4X (%)
L3_1	69,565,132	70,747,478	98.33	3.32	89.19	30.47
Tsw_2	74,938,031	76,072,940	98.51	3.53	90.75	34.32
Tsw_1	68,891,040	69,933,638	98.51	3.34	89.55	30.40
CM_1	61,124,442	62,090,386	98.44	3.06	87.39	24.49
PM_1	51,687,818	52,471,638	98.51	2.75	84.12	18.47
PM_2	69,623,698	70,712,424	98.46	3.34	89.58	30.35
CM_2	70,387,974	71,443,758	98.52	3.38	89.91	31.01
L3_2	58,516,258	59,506,162	98.34	2.95	86.18	22.87
L4_1	52,625,642	53,511,830	98.34	2.76	83.32	18.72
L4_2	64,693,894	65,744,832	98.40	3.16	87.82	27.02

Table 30. Statistics of mapping rate, depth and coverage

The details for mapping statistics are as follows:

(1) Sample: sample names.

(2) Mapped reads: the number of clean reads mapped to the reference assembly, including both single-end reads and reads in pairs.

(3) Total reads: total number of effective reads in clean data.

(4) Mapping rate: the ratio of the reference genome assembly mapped reads to the total sequenced clean reads.

(5) Average depth: the average depth of mapped reads at each site, calculated by the total number of bases in the mapped reads dividing by size of the assembled genome.

(6) Coverage at least 1X: the percentage of the assembled genome with more than one read at each site.

(7) Coverage at least 4X: the percentage of the assembled genome with ≥4X coverage at each site.

For the current 2,935,884,163 bp reference genome, the mapping rate of each sample ranged from 98.33% to 98.52%. Refer to the reference genome (without Ns), the average depths were between 2.75X and 3.53X, and 1X coverages range from 83.32% to 90.75%. This result is in the qualified normal range and may serve in the subsequent variation detection and related analyses.

5.3.1 SNP detection & annotation

Single nucleotide polymorphism (SNP) refers to a variation in a single nucleotide which may occur at some specific position in the genome, including transition and transversion of a single nucleotide. We detected the individual SNP variations using SAM tools with the following parameter: 'mpileup -m 2 -F 0.002 -d 1000'.

To reduce the error rate in SNP detection, we filtered the results with the criterion as follows:

(1) the number of supports reads for each SNP should be more than 4;

(2) the mapping quality (MQ) of each SNP should be higher than 20;

ANNOVAR was used for variation annotation with multiple capabilities, including gene-based annotation, region-based annotation, filter-based annotation as well as other functionalities. Novogene use ANNOVAR to do annotation of detected SNPs. The distribution of SNPs in different regions of genome (upstream, exonic, intronic etc) was evaluated and is presented in Tab. 31.

Sample	Upstream	Exonic				Intronic	Splicing	Downstream	Upstream/ Downstream	Intergenic	Total
		Stop gain	Stop loss	Synonymous	Non-synonymous						
PM_1	14236	121	25	2970	5152	37136	60	11882	650	968902	1076700
PM_2	24067	206	52	4971	8471	63894	82	19873	1187	1696134	1879332
L4_1	14569	129	42	3390	5880	43356	65	12275	711	1052762	1171796
L4_2	21559	201	54	4992	8147	63187	81	18575	1109	1582117	1756735
L3_1	24210	181	52	5236	8597	70340	91	20385	1262	1751247	1944495
L3_2	17812	130	41	3949	6386	51813	64	15241	920	1274007	1416936
CM_1	18699	167	38	4019	6678	51908	84	15938	984	1330391	1478926
CM_2	24772	207	53	4929	8220	66356	90	20831	1182	1720002	1908988
Tsw_1	24968	201	50	5091	8384	65953	90	21081	1267	1774991	1962814
Tsw_2	28623	246	54	5657	9676	75292	90	24003	1453	2021975	2235897

Table 31. Statistics of SNP detection and annotation

The details for SNP detection and annotation statistics are as follows:

- Upstream: SNPs located within 1 kb upstream (away from transcription start site) of the gene.
- Exonic: SNPs located in exonic region; Non-synonymous: single nucleotide mutation with changing amino acid sequence; Stop gain/loss: a nonsynonymous SNP that leads to the introduction/removal of stop codon at the variant site; Synonymous: single nucleotide mutation without changing amino acid sequence.
- Intronic: SNPs located in intronic region.
- Splicing: SNPs located in the splicing site (2 bp range of the intron/exon boundary).
- Downstream: SNPs located within 1 kb downstream (away from transcription termination site) of the gene region.
- Upstream/Downstream: SNPs located within the < 2 kb intergenic region, which is in 1 kb downstream or upstream of the genes.
- Intergenic: SNPs located within the > 2 kb intergenic region.
- Others: SNPs located in other region.
- Total: The total number of SNPs.

The whole-genome SNP mutations could be classified into six categories. The frequency of each type observed in our samples is shown in Fig. 55.

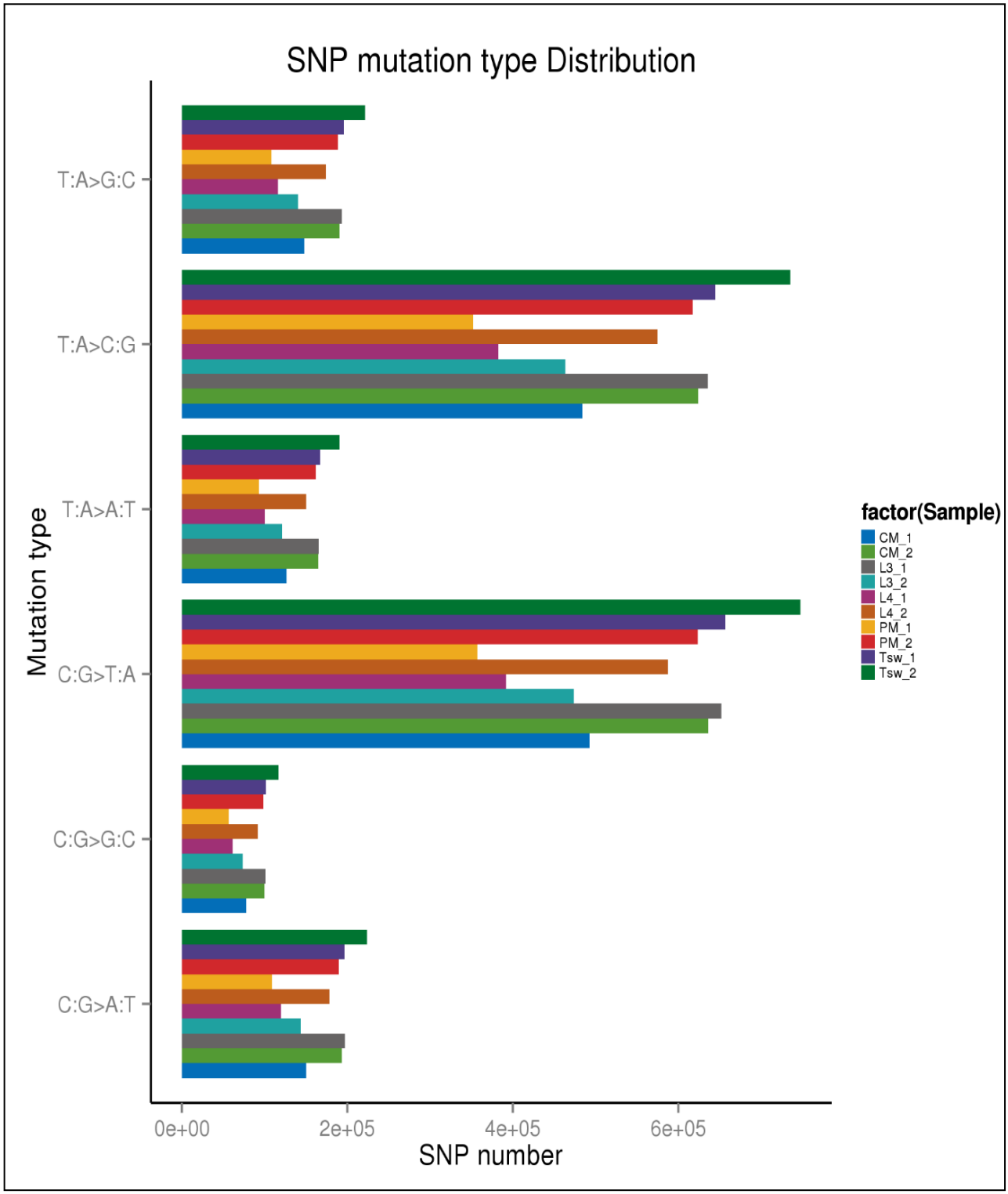


Figure 55. Frequency of SNP mutations. The x-axis represents the number of the SNPs, and y-axis indicates the mutation types

5.3.2 InDel detection & annotation

InDel refers to the insertion or deletion of ≤ 50 bp sequences in the DNA. Novogene uses SAMTOOLS (Li *et al*., 2009) with the following parameter 'mpileup -m 2 -F 0.002 -d 1000' to detect InDels and followed by annotation using ANNOVAR (Wang *et al*., 2010). The filter conditions are same with SNPs.

The distribution of SNPs in different regions of genome (upstream, exonic, intronic etc.) was evaluated and is presented in Tab. 32.

Sample	Up stream	Exonic						Intronic	Splicing	Down stream	Intergenic	Insertion	Deletion	Total
		Stop gain	Stop loss	Frameshift Del	Frameshift In	Non-frameshift Del	Non-frameshift In							
PM_1	1864	6	5	106	97	55	49	4082	10	1508	60366	35705	35909	71629
PM_2	3210	12	3	179	158	96	95	7214	16	2619	106443	62972	63004	126019
L4_1	1841	4	5	109	98	52	66	4724	16	1529	63976	38086	38076	76178
L4_2	2804	7	6	172	120	90	100	7027	21	2288	97093	57490	57791	115293
L3_1	3082	13	2	142	120	90	99	7833	20	2575	107837	63827	64244	128105
L3_2	2223	6	2	117	93	80	82	5703	15	1907	78627	46691	46773	93484
CM_1	2564	9	3	140	98	88	84	5944	23	2002	81354	48705	48310	97037
CM_2	3287	5	1	156	113	99	103	7567	26	2667	107705	63999	63910	127948
Tsw_1	3337	9	3	174	123	93	101	7616	23	2723	110135	65446	65062	130540
Tsw_2	3849	19	3	176	143	116	117	8744	26	3181	126567	75068	74746	149852

Table 32. Statistics of InDel detection and annotation

The details of InDel annotation statistics are as follows:

(1) Sample: sample names.

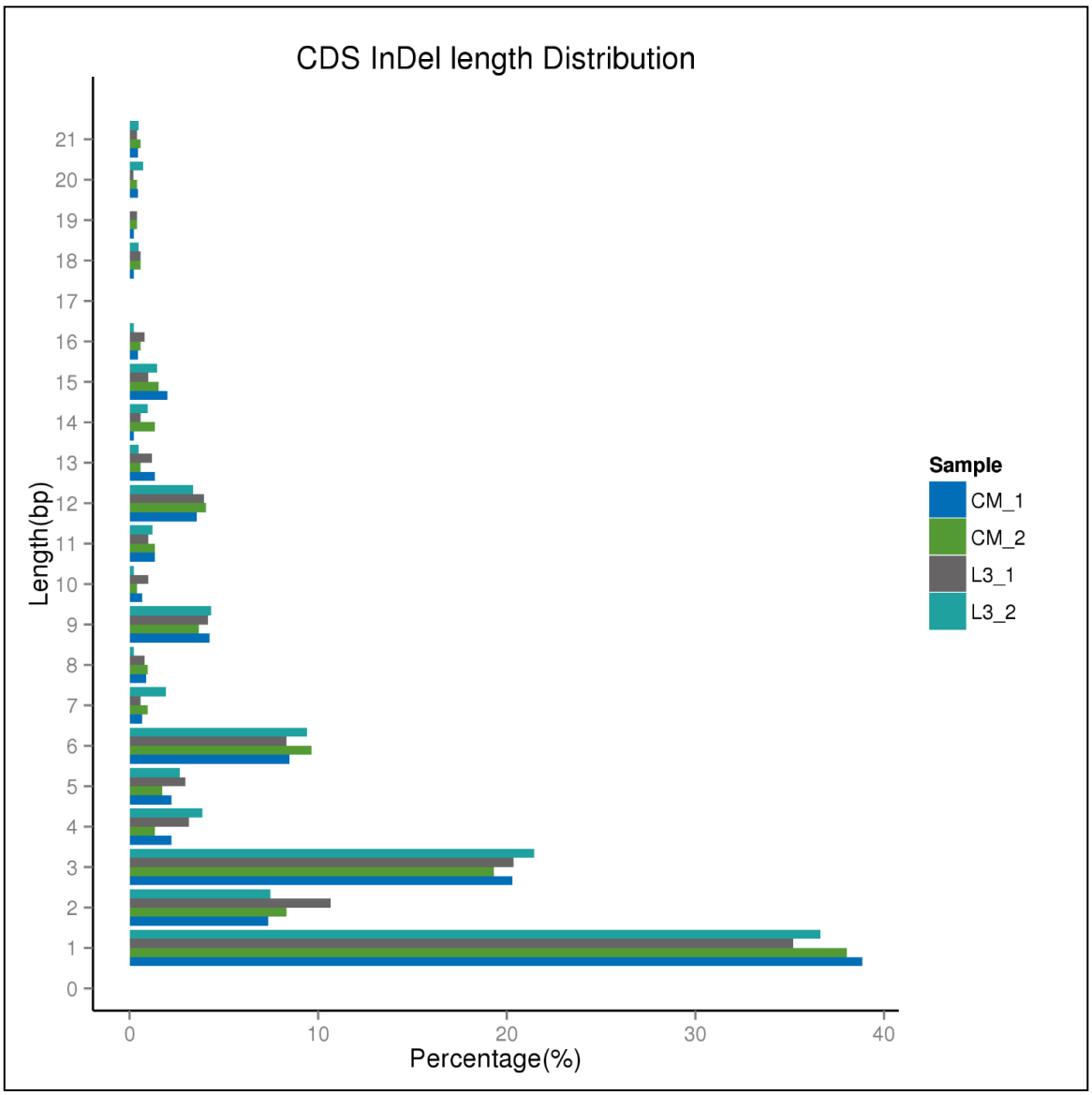
(2) Upstream: InDels located within 1 kb upstream (away from transcription start site) of the gene.

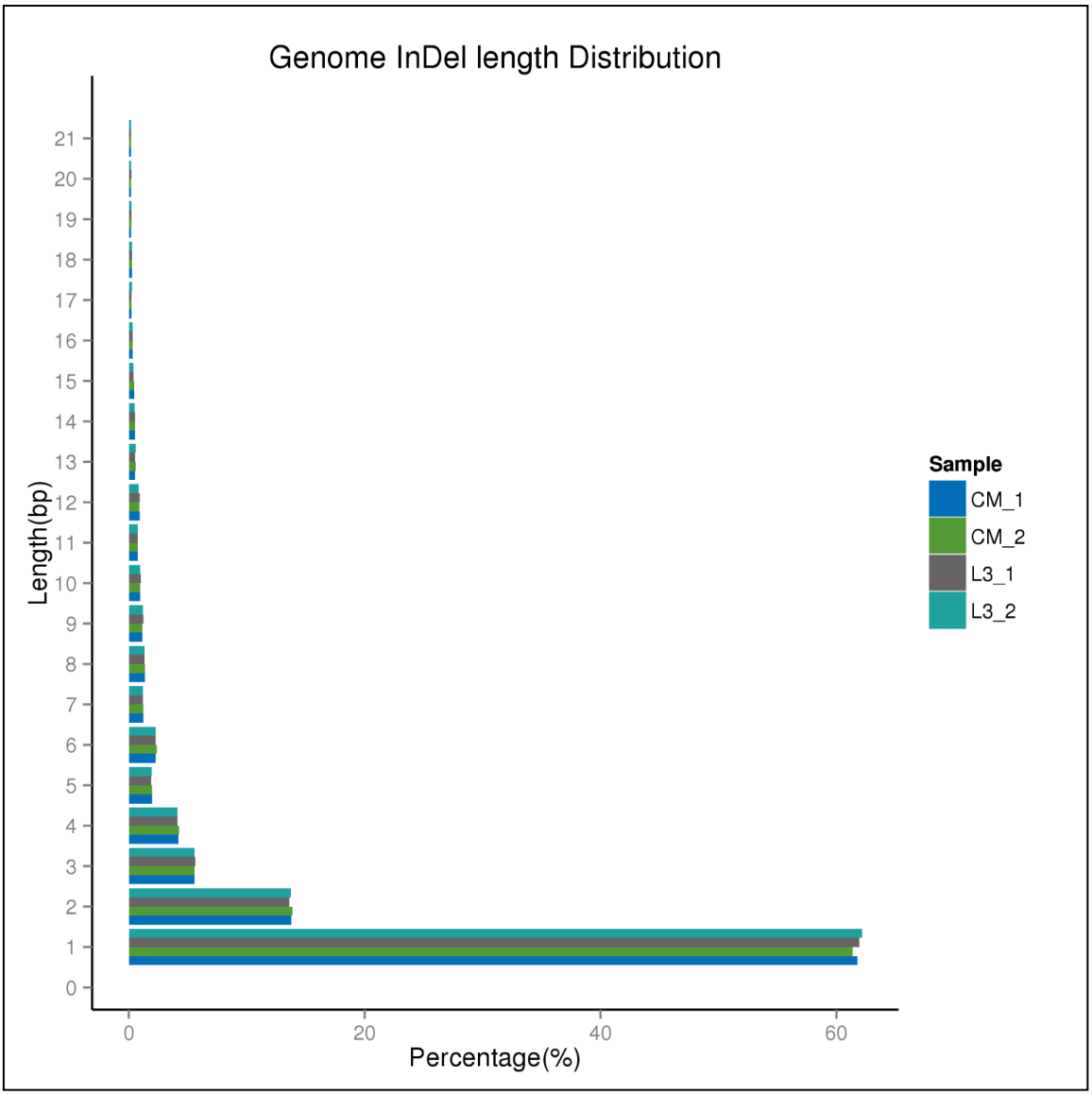
(3) Exonic: InDels located in exonic region; Stop gain/loss: InDel that leads to the introduction/removal of stop codon at the variant site; Frameshift deletion/insertion: InDel mutation changing the open reading frame with deletion or insertion; Non-Frameshift deletion/insertion: InDel mutation without changing the open reading frame with deletion or insertion sequences of 3 or multiple of 3 bases.

(4) Intronic: InDel located in intronic region;

(5) Splicing: InDel located in the splicing site (2 bp range of the intron/exon boundary).
(6) Downstream: InDel located within 1 kb downstream (away from transcription termination site) of the gene region).
(7) Upstream/Downstream: InDel SNPs located within the < 2 kb intergenic region, which is in 1 kb downstream or upstream of the genes.
(8) Intergenic: InDel located within the > 2 kb intergenic region.
(9) Total: The total number of InDels.

The Fig. 56-57 show the example length distribution of InDels for the random four samples (full data not shown for the readability of the figures) list for coding and genomic regions.





Figures 56-57. Length distribution of InDels. The x-axis represents the proportion of the InDels with a certain length, and y-axis indicates the length of the InDels.

5.3.3 SNP/InDel density

The SNP/InDel density for each chromosome was evaluated showing a relatively very low level of polymorphism respect reference genome for all genotypes (very few red lines) and the absence of "red blocks" typical for introgression events in pedigree. Figures 58 and 59 presents example results obtained for sample PM1

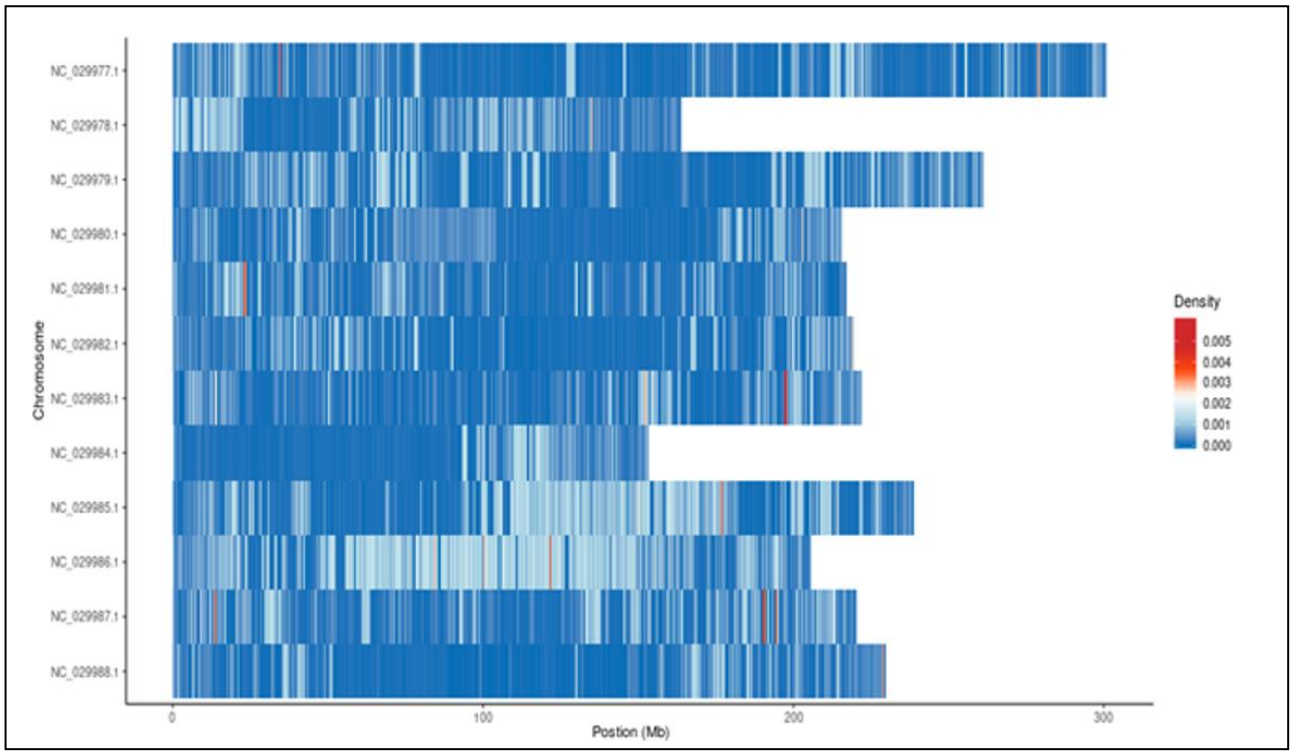


Figure 58. The PM_1 sample SNP density. The SNP density on each window was counted with 100kb as the window

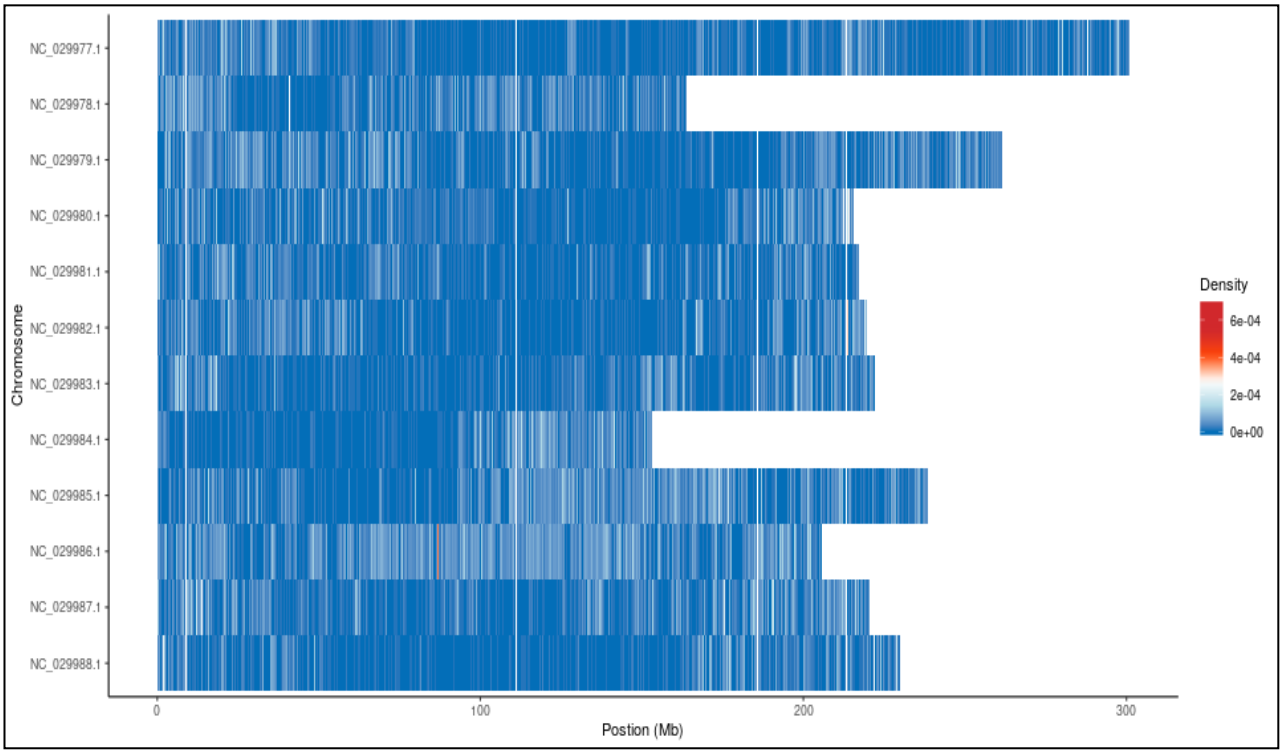


Figure 59. The InDel density insample PM_1. The InDel density on each window was counted with 100kb as the window

The list of all software used for the bioinfomatic analysis is presented in Tab. 33.

Analysis	Software	Usage	Version
Quality control	Fastp	Quality control	0.20.0
Mapping	BWA	Mapping clean reads to the reference genome and generation of bam result files.	0.7.8-r455
	SAMtools	Sorting the bam files and removing duplication reads.	1.3.1
	Picard	Merging the bam files from the same sample.	1.111
SNP/InDel Detection	SAMtools	Detection and filtration of SNPs and InDels.	1.3.1
Variation Annotation	ANNOVAR	Annotation of the detected variations.	2015Dec14

Table 33. List of softwares in WGS analyses

5.3.4 Genetic distances

Only polymorphisms restricted to exonic regions and present in all biological replicates were used for further analysis (due to abundance of data).

The analysis of genetic distances based on polymorphisms identified, done with PAST (Paleontological and statistics) software hierarchical clustering option (Fig. 60), highlighted a very narrow genetic base of the cultivated varieties used in the breeding programs in confront with reference Zunla (chinese cv.) genome.

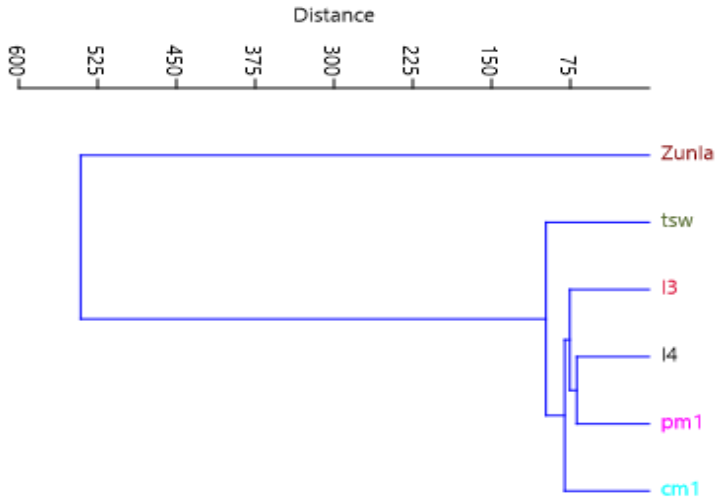


Figure 60. Analysis of the genetic distances using Past software

5.3.5 Private SNP evaluation

A broader set of private polymorphisms (only those located within exonic regions with the priority given to non-synonymous mutations) were obtained for the further studies of candidate genes (data not shown).

A restricted panel of private SNPs distributed on most of the chromosomes and giving origin to codon stop gain or loss (that leads to shorter or highly modified protein that most probably loose or change their function) were identified (Tab. 34).

	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8	Ch9	Ch10	Ch11	Ch12
PM		1 gain	1gain	1 gan /1 loss	1 loss	1 gain /1 loss	1 gain		1 gain	1 gain		
TSW		1 loss	1 gain			4 gain			1 gain		1 gain	
CM	1 gain/ 1 loss		1 gain/ 2 loss	2 gain			1 gain		1 gain		1 gain	
L3	1 gain	2 gain	1 loss		2 gain	1 gain	1 gain	1 loss				1 loss
L4	1gain	3 loss	1 loss						1 loss	1 gain/ 1 loss		

Table 34. Private SNP identified

5.4 Conclusions

Relatively few private polymorphisms were identified during this work. As the genetic basis of pepper species is quite narrow, most mutations (respect to reference genome) are shared with at least one genotype, and this was highlighted by genetic distances calculation.

Private (occurring only in one genotype) polymorphism were then inferred. Based on their distribution, SNPs are considered private when they are found in a single or a range of samples of a collection and absent in the rest, due to mechanisms such adaptation or as human-based selection.

Even if private SNPs are not reported to be crucial for determining major changes of plant phenome, their investigation could highlight the existence of potential candidate genes for traits of agricultural interest or be applied for a precise genotype identification /traceability/purity.

A broader sets of private polymorphisms (only those located within exonic regions with the priority given to non-synonymous mutations) were obtained for the further future studies of candidate genes. Most of them are annotated as "unknown" however the annotation should be check regularly as the genome versions are being updated.

A restricted panel of private SNPs distributed on most of the chromosomes and giving origin to codon stop gain or loss (that leads to shorter or highly modified protein that most probably loose or change their function) were identified to be applied in the company for the traceability and variety identification (possible patent application). Noteworthy a stable number of 8-10 private SNP_S leading to stop/loss were identified in all genotypes.

With a nowadays costs of resequencing, a seed company may characterize easily complete genomic profiles of their varieties, and mine those data for genes/markers useful for breeding programs but also to protect their intellectual property.

Edward Henry Weston (1886-1958)

Chapter 6

Inter-placing selection breeding scheme in pepper

6.1 Introduction

Most of the pepper breeding program have always addressed the development of cultivars or hybrids able to meet specific parameters regarding both changing and improving agricultural practices and qualitative parameters required by final consumer. During the last century, introgression of resistances to a wide range of pathogens and pests have been deep investigated while strategy related the creation of superior genotypes for adaptability performance lagged behind (Roy, D. and Kharkwal M. 2004).

The success of a new variety released depends thus from many factors but key role is often related to capacity to adapt it in different growing condition in terms of abiotic stresses. As for other hybrids crops, efficient breeder program oriented towards parental lines with enhanced genetic background for multi-environment adaptation could represent a winning strategy.

The long-term objective of the present study is to conduct an efficient breeding program following a new proposal Inter-placing selection breeding scheme.

Set to enhance the genetic variability exploration and subsequent selection in *Capsicum* species, the Inter-placing scheme allow the multiple accumulation of favorable genes in a single genotype with an improved accuracy and efficiency due to pyramiding of genes for adaptability in the background to improve a genotype by multi-trait and multi-environment selection.

A scheme that lead the selection taking advantages of the different cultivation cycles which is subject pepper crop such as protected (winter, spring, fall greenhouse) and open field is being proposed in the present work (Fig. 61). The seasonal and environmental conditions determine specific abiotic pressure (e.g. temperature range, humidity, soil composition, brightness etc.) at different plant development stages. This affects a breeding selection oriented towards specific plant characteristic in that environment such as plant habit, fruit setting capacity, leaf coverage, length of internodes, qualitative parameters considering strong interaction GxE in pepper specie. Correct subsequent multi-environment cycles of selection allow to introgress adaptability characteristics in each generation, generating environment-adapted and fixed germplasm.

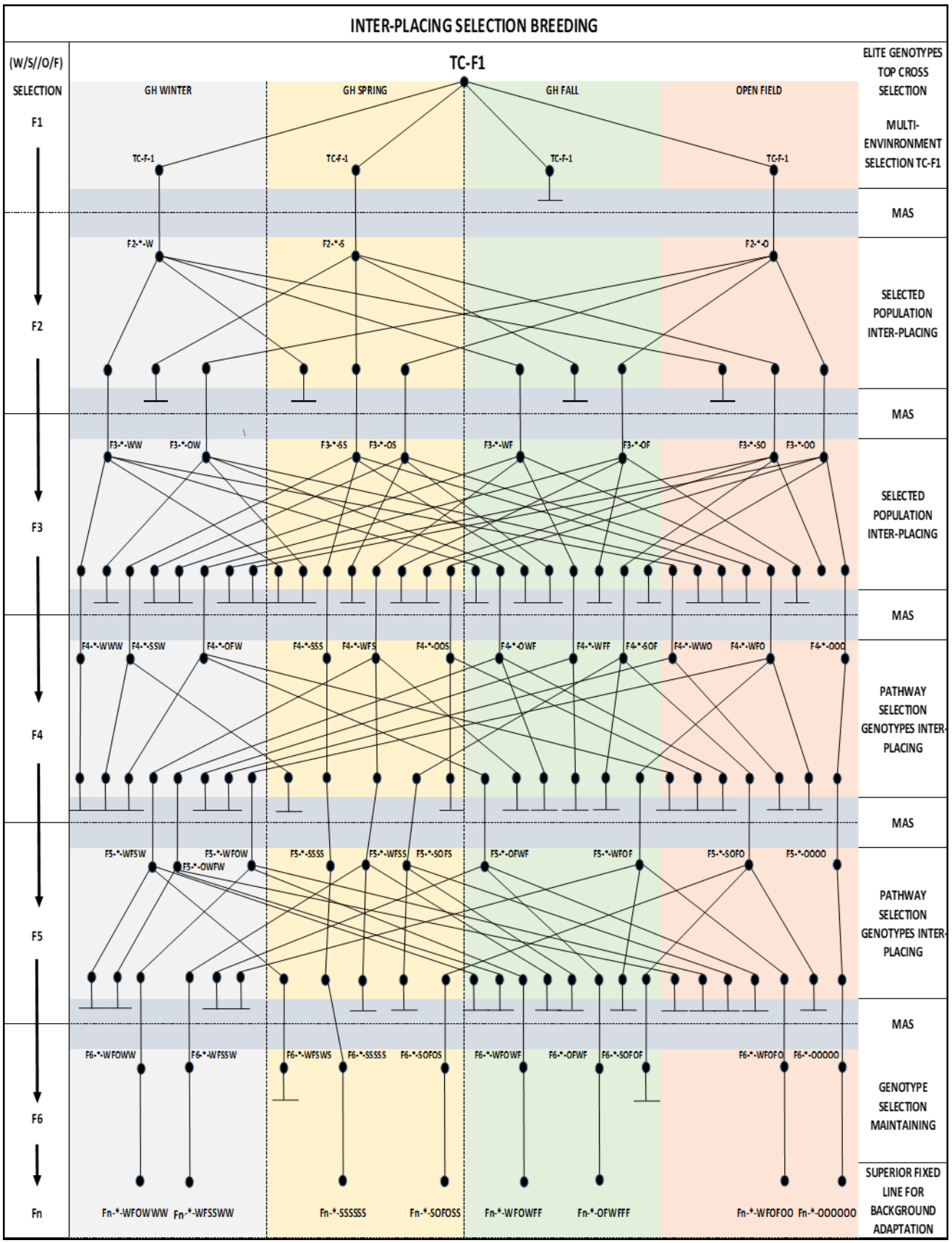


Figure 61. Schematic representation of Inter-placing selection (TC=top cross, GH= greenhouse, W=winter, S=spring, F=fall, O=open field)

6.2 Materials and Methods

6.2.1 Elite genotypes crosses and TC-F1 multi-environment selection

In order to enhance variability within breeding program, starting cross for the creation of pre-breeding/breeding population foresees the choice of elite genotypes for respective top performance in four different growing conditions: greenhouse W:winter, S:spring, F:fall and O:open field. Top cross (TC-F1) with W/S//O/F pedigree is split in each experimental field and subjected to selection. Considering variability related to parental line used, a consistent populations have to be analyzed and characterized (at least transplant of 2-300 individual). Marker assisted selection for traits of interest of each plants is then conducted in order to add value in terms of gene pyramiding approach. Keeping a mild phenotype selection, bulk methods of best plant in relation to trait of interest is suggested to advance the generation.
A summary of phenotypic traits useful for plant selection are report below in table 35.

General Information	Value	Value = 0	Value = 1	Value = 2	Value = 3	Value = 4	Value = 5
	Type of Pepper	Lamuyo	Dulce Italiano	Charliston	Sivri	Dolma	Blocky
	Color	Not Relevant	Light green	Dark green	Red	Yellow	Orange
Plant Characteristics	Plant vigor	Not relevant	Very weak	Weak	Medium	Strong	Very strong
	Internode length.	Not relevant	Very short	Short	Medium	Long	Very long
	Leaf coverage of fruit	Not relevant	Open plant	Some coverage	Medium coverage	Good coverage	Full coverage
	Side branches	Not relevant	Very few	Few	Average	Many	Too many
	Plant continuance	Not relevant	Very poor	Poor	Medium	High	Very High
	Fruit setting under low temperatures	Not relevant	Very poor	Poor	Medium	Good	Very good
	Fruit setting under high temperatures	Not relevant	very poor	poor	Medium	Good	Very good
	Harvesting (Maturity): Early-Late vs Control	Not relevant	Very early	Early	Medium	Late	Very late
	Ability to hold red fruit (vine storage)	Not relevant	Very poor	Poor	Medium	Good	Very good
	Yield visual	Not relevant	Very low	Low	Medium	High	Very high
Fruit Characteristics	Fruit weight (grams)	Not relevant					
	Fruit weight uniformity	Not relevant	Very poor	Poor	Medium	Good	Very good
	Fruit diameter (mm)	Not relevant					
	Fruit length (cm)	Not relevant					
	Fruit shape uniformity	Not relevant	Very poor	Poor	Medium	Good	Very good
	Fruit wall thickness	Not relevant	Very thin	Thin	Medium	Thick	Very thick
	Depth of shoulders	Not relevant	Very deep	Deep	Normal	Flat	Sloped
	Number of lobes on fruit	Not relevant					
	Blossom end shape	Not relevant	Very pointy	Not very pointy	Normal	Concave shape	Very deep concave
	Blossom end cracking	Not relevant	Very big scars	Big scars	Medium scars	Small scars	No scars
	Cracking sides-shoulders	Not relevant	Very large	Large	Medium	Low amount	No cracking

	Micro-cracking on sides & shoulders	Not relevant	Very large amount	Large amount	Medium	low amount	No cracking
	Calyx: touching or not touching fruit shoulders	Not relevant	Not touching	Partly touching	Normal	Almost fully touching	Fully touching
	Silvering	Not relevant	very large amount	large amount	some	almost none	none
	Papule - small bumps on skin	Not relevant	very large amount	large amount	medium	low amount	none
	Green color before maturity	Not relevant	Very light green	Light	Medium	Dark	Very dark green
	Final fruit color.	Not relevant	Very poor	Poor	Medium	Good	Very good
	Firmness when fruit is ready to pick	Not relevant	Very soft	Soft	Medium	Firm	Very firm
Post Harvest	Harvesting stage of fruit being checked	Not relevant	Green	25%	50%	75%	90%+
	Shelf life	Not relevant	Very short	Short	Medium	Long	Very long
	Firmness	Not relevant	Very soft	Soft	Medium	Good	Very good
	Final fruit color	Not Relevant	Very poor	Poor	Average	Good	Very good
	Green stem / peduncle freshness	Not relevant	Not fresh at all	Not very fresh	Medium	Fresh	Very fresh
	General impression of post harvest fruit	Not relevant	Very poor	Poor	Medium	Good	Very good
Conclusions	Overall general impression	Not relevant	Very poor	Poor	Medium	Good	Very good
	Conclusion						

Table 35. List of phenotypic traits usable useful for plants characterization and selection

6.2.2 Inter-placing selection F_2 to F_6

F_2 bulk seed generated from every field is subject again to an Inter-placing positioning. Each subsequent selection will take a name/code corresponding to progressive environment selection. In order to employ a high segregation level, at-least 150-200 individual must be evaluated for each selection. After MAS screening, best plants are selected in bulk and harvest. F_3 seed collected is again split via Inter-placing criteria and selected for best performance adaptability.

The selection level is raised in order to keep size population controlled, especially regarding the advancement for F_3 to F_4 where single plant choice is preferable. At this stage are still present genotypes coming from multi-environment selection. Step by step, advanced material will reach a certain stability and for this reason stage F_4 to F_6 envolves positioning of genotype based on previous breeding field pedigree.

For instance a genotype F_4-*-WFF (selected respectively in winter and fall) will be divided and assigned for subsequent selection separately in those environments, following a pathway selection. For enhancing the fixity of plant materials it is very important to increase the selection rate with an advancement for single plant.

6.2.3 Selection maintaining F_6 to Fn

Reached F_6 generation, maintaining selection is applied and genotype is advanced keeping the last one enviroment of selection (form previous step) to stabilize the background adaptation. In this way stability is reached, after n generation (Fn) of selection through a gene piramyding for traits related to adaptability (to different conditions in the firts steps) and marker assisted selection for targeted alleles (e.g. genes for resistance to biotic stresses). Inter-placing selection generate superior elite genotypes with different genetic background adaptation (Tab. 36) constituting new advanced elite materials that will confer to hybrids succesful factors in term of adaptation "plasticity" in different growing conditions. Starting from a single pre-breeding cross, this pipeline, allow to maximize the exploration of variability creating the condition for a selection aimed to enhancing the genetic distance between new high value genotypes.

SUPERIOR ELITE GENOTYPES	GENETIC BACKGROUND ADAPTATION
Fn-*-WFOWWW	WINTER (fall, open field)
Fn-*-WFSSWW	WINTER (fall, spring)
Fn-*-SOFOSS	SPRING (open field, fall)
Fn-*-WFOWFF	FALL (winter, open field)
Fn-*-OFWFFF	FALL (winter, open field)
Fn-*-WFOFOO	OPEN FIELD (fall, winter)

Table 36. Superior elite genotypes with genetic background adaptation

6.3 Results

A total of 3 top cross (mating 4 way combination) aimed for the creation of pre-breeding and breeding populations was performed in the company facilities using elite genotypes with top performance for different growing condition as well harboring resistance genes of interest. For TC1 to TC3 were respectively placed (200 plants each) within different breeding field characterized by the four different environmental conditions previously described. At seedling stage, each plants were analyzed using molecular markers for main resistance genes (*L4*, *tsw*, *pmr1*), subsequent selection was carry out through separate bulk for provenience origin, choosing best plants for habit behaviour as well quality of fruit and taking care of MAS information (grouping for homozygous or heterozygous gene).

F_2 collected seed (about 2ks) was then re-deployed in each selective field (150 genotypes per each selected group) and analyzed with "integrative" marker assisted selection (i.e. using marker per heterozygous subgroup). Advanced selected plants was collected in bulk and F_3 seed was extracted. Tab. 37 show realized selection process.

GENERATION/ CYCLE	WINTER		SPRING		FALL		OPEN FIELD	
	id	selection name	id	selection name	id	selection name	id	selection name
F1 to F2	tc-F1-1	**1-F2-W**	tc-F1-1	**1-F2-S**	tc-F1-1	-	tc-F1-1	**1-F2-O**
	tc-F1-2	**2-F2-W**	tc-F1-2	-	tc-F1-2	**2-F2-F**	tc-F1-2	-
	tc-F1-3	**3-F2-W**	tc-F1-3	**3-F2-S**	tc-F1-3	-	tc-F1-3	-
F2 to F3	1-F2-W	**1-F3-*-WW**	1-F2-W	-	1-F2-W	-	1-F2-W	-
	2-F2-W	**2-F3-*-WW**	2-F2-W	**2-F3-*-WS**	2-F2-W	-	2-F2-W	-
	3-F2-W	-	3-F2-W	**3-F3-*-WS**	3-F2-W	-	3-F2-W	**3-F3-*-WO**
	1-F2-S	-	1-F2-S	**1-F3-*-SS**	1-F2-S	-	1-F2-S	**1-F3-*-SO**
	3-F2-S	**3-F3-*-SW**	3-F2-S	-	3-F2-S	**3-F3-*-SF**	3-F2-S	-
	2-F2-F	**2-F3-*-FW**	2-F2-F	**2-F3-*-FS**	2-F2-F	**2-F3-*-FF**	2-F2-F	**2-F3-*-FO**
	1-F2-O	**1-F3-*-OW**	1-F2-O	**1-F3-*-OS**	1-F2-O	**1-F3-*-OF**	1-F2-O	**1-F3-*-OO**

Table 37. Application of first selection stages using Inter-placing scheme which details naming assignment for each chosen genotype in relation to generation and selection environment

During third year and raising the selective pressure, F_3 populations were managed again via interplacing scheme through same procedure. Generation was advanced via single plant collection producing F_4 (21 population was collected).

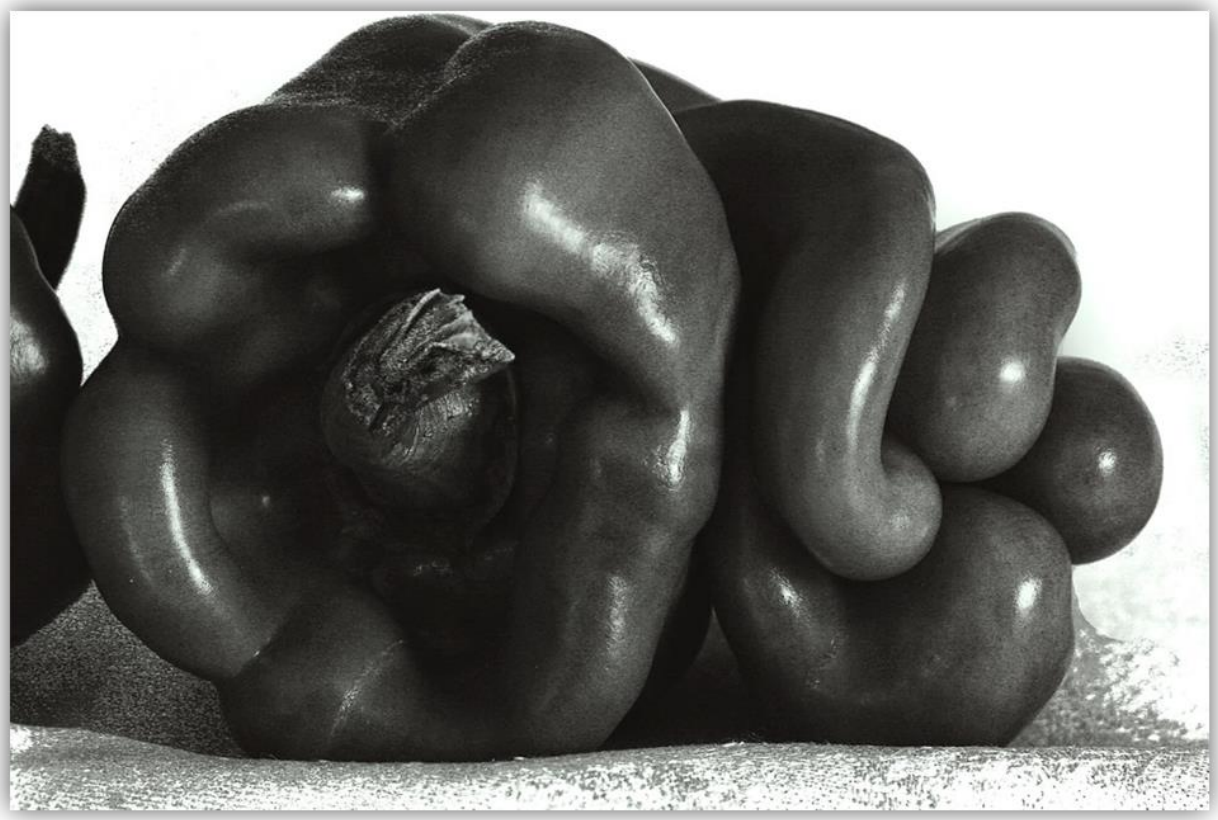
F_4 selection is currently ongoing following the pathway pedigree: each genotype has now been placed only in the fields where it had been advanced in the previous selections.

6.4 Conclusions

Multidisciplinary field of modern plant breeding combines molecular methods and tools withmore conventional approaches in order to obtain improvement of crop plants. Genomics, phenomics, enviromics are often interconnected and their use can be optimized based on needs resources of a breeding program. There is a strong need for using new technologies to accelerate the progress of plant breeding in order to meet the well know challenges of food and nutrition quality security. Genetic improvement process is speed up by recent advances in "omics" field, enabling a great step forward in a new plant breeding era (Kumar *et al*., 2021). Furthermore, the success or failure of the whole, years-lasting program may depend on the correct choice of breeding strategy. Basical approach is thus based on breeder expertise in careful observation matched together with behaviours understanding, multiple genotypes evaluation (parents and progenies), selection at different generation levels and wide enviromental test for adaptability.

Efficient seed delivery is the principal aim of the industrial seed company and to reach this objective, plant breeder must adopt a continuous updating philosophy in order to improve breeding schemes design deploying the novel breeding methodologies. Unfortunately, breeding schemes are rarely defined, not stored in a database and not expressed in a quantifiable format. Furthermore, the continuous implementation and updating of breeding approach is not routinely conduced in many breeding programs (Matthews *et al*., 2019, Covarrubias-Pazaran *et al.*, 2022).

The proposed Inter-placing method allow the generation of high value germplasm through the creation of enriched parental line with characteristics deriving from accurate management of subsequent cycles of selection. Optimizing the breeding pipeline for efficiency and efficacy is the key step for maximize breeding program resources and increase genetic gains useful for new varieties creation. These methods can be incorporated into conventional breeding programs or help to re-design established breeding process with also further applicability and possibilities concerning genomic support through the development of training population for genomic selection or the increasing of genetic distance of a germplasm collection via pre-breeding perspective.



Edward Henry Weston (1886-1958)

Chapter 7

General discussion and future perspectives

Pepper (*Capsicum* spp.), that belongs to the Solanaceae family, is among the most economically important vegetable crops mainly because of the high demand and its spicy taste, popular in many parts of the world (Pinto *et al*., 2016). The genus includes over 30 species, five of which (*C. annuum, C. frutescens, C. chinense, C. baccatum*, and *C. pubescens*) are domesticated and grown for food and for non-food purposes (e.g. cosmetics) (Parisi *et al*., 2020). Given its very large variability and geographical distribution, pepper (*Capsicum annuum* L.) includes a large number of accessions with a considerable variation for several traits including growth habit, seeds, flower and fruit color, shape and size (Di Dato *et al*., 2015).

While China and India are the biggest producers of pepper (dry and fresh) with more than 50% of the world production, the Mediterranean area itself represents about 20% of the world harvest (Faostat 2020).

The sustainability of agricultural production, food security and high yield due to the increasing of consumers demand remain the main challenges for pepper crop improvement. We have to consider, however that a great range of pathogens affect pepper production, including fungal (anthracnose and powdery mildew), oomycete (*Phytophthora* root rot), viral (*Cucumber mosaic virus, Tobamoviruses, Potyviruses, Tomato spotted wilt virus* etc.), bacterial (bacterial spot and bacterial wilt) and nematode (root-knot nematodes) (Barchenger *et al*., 2019; Barka, G. and Lee, J., 2020; Parisi *et al*., 2020) that cause economic damage and crop loss.

Considering that chemical control is neither environmentally nor economically sustainable, breeding is the best strategy to manage the diseases by developing resistant varieties/hybrids, however the exploitation of *Capsicum* germplasm for resistances still represent challenging tasks (Babu *et al*., 2011). Indeed, despite continuous new scientific findings, there is still a need of developing new cultivars resistant to various pests and adapted to varied pedoclimatic conditions (Di Dato *et al*., 2015).

In order to obtain the genetic resistances without hampering the quality, marker assisted selection is the best method to speed up the genetic improvement process, moreover pepper breeding for multiple resistances to multiple diseases is highly required (Wiesner-Hanks, T. and Nelson, R., 2016).

The availability of molecular markers linked to identified/cloned resistance genes, enabled the accumulation of disease resistance genes through the use of marker-assisted selection (MAS) facilitating screening of breeding populations for resistance alleles respect methods that require pathogen inoculation (Cobb *et al*., 2019; Barka, G. and Lee, J., 2020).

The aim of the present work was to introduce latest advancements in marker development and deployment for innovation of molecular characterization of company's germplasm collection in order to release new, superior, resistant varieties through the improvement of breeding strategy as well as the management of genetic resources.

The success of any breeding program depends mainly on two factors: high-value breeding material and a correct breeding strategy with an accurate breeding scheme. Our activity started thus with the construction and characterization of the company's proprietary germplasm collection as the deep knowledge is a base of a proper choice of parental genotypes (Lozada *et al*., 2022).

The selection of parental lines with specific characteristics is critical for effective crop breeding schemes, which are highly dependent on phenotypic selection after the development of breeding populations. Once the parental lines are determined based on a target phenotype, genotypic variation underlying it become informative for developing polymorphic molecular markers that can be used to trace genes responsible for observed genetic variation without the need of constant phenotyping (Kang *et al*., 2016).

The long-term objective of the present study is to conduct an efficient breeding program following a novel proposed Inter-placing selection breeding scheme (Chapter 6). Set to enhance the genetic variability exploration and subsequent selection in *Capsicum* species, the Inter-placing scheme allow the multiple accumulation of favorable alleles in a single genotype with an improved accuracy and efficiency due to pyramiding of genes for adaptability by multi-trait and multi-environment selection.

Inter-placing selection generate superior elite genotypes with different genetic background adaptation constituting new advanced elite materials that will confer to hybrids a kind of adaptation "plasticity" to different growing conditions. The Inter-placing method allow the generation of high value germplasm through the creation of enriched parental line with this characteristic deriving from accurate management of subsequent cycles of selection.

Optimizing the breeding pipeline for efficiency and efficacy is the key step for maximize breeding programs resources and increase genetic gains useful for new varieties creation. These methods can be incorporated into conventional breeding programs or help to re-design established breeding process with also further applicability and possibilities concerning genomic support through the development of training population for genomic selection or the increasing of genetic distance of a germplasm collection via pre-breeding perspective.

Molecular multi-gene identification, useful for marker-assisted breeding, is being/will be applied directly on pre-breeding crosses toward gene pyramiding approach allowing simultaneously the beginning application of new innovative breeding scheme of the company.

A panel consisting on 185 *Capsicum annum* genotypes representative of all main segments cultivated in the Mediterranean area such as blocky, lamuyo, white type, bull horn, kapia, charleston, sivri, topepo and hot cayenna was set up recovering seeds from company's breeding fields placed in different country (Spain, Italy, Hungary), creating subgroups of resistant/susceptible genotypes for each of the diseases of interest.

Application of 48 KASP molecular markers enabled the evaluation of homozygosity/heterozygosity level of the collection, fundamental for the correct breeding scheme application.

The evaluation of fixity stage of parental lines introduced to breeding programs is especially important for the introgression of recessive characters for obtaining heterosis effect of F_1 hybrids (Acquadro *et al*., 2020).

This activity was propaedeutic for the PhD project enabling the identification of elite materials useful for the successive step of research activity, while representing an immediate and important tool for germplasm collection management within company breeding program since clustering analysis allowed to define a molecular pattern that can serve as reading key for the pepper typology, fruit structure and provenience area.

Subsequently, in order to give additional information on the genotypes under study as well as to detect and transfer useful alleles from different backgrounds, development of internal protocols based on known markers and screening of the germplasm collection with markers linked to major disease resistance genes was performed.

In order to implement the company marker assisted selection activity, several molecular markers of different types associated to resistance genes for the most important diseases present in the Mediterranean area were identified.

The ease of use within the breeding program was taken into account giving priority to main characters under monogenic control and user friendly protocols (agarose gel) that could be easily applied also in a small laboratory that every, even small seed company, can equipped on their premises.

It will enable the company to perform the screening activity of breeding population, in a "on the fly" manner, *in situ,* without relaying on external services and thus improving the speed and efficiency of selection.

The bibliographic search performed highlighted that the publicly available markers are not numerous (many of them being protected by a patent) and since the majority of them were developed using particular local landraces/varieties (mainly Asian) before the availability of reference genome sequence, often the primers could not be aligned on genome sequence due to the apparent differences in our breeding material.

The advances of next-generation sequencing (NGS) technologies by whole-genome resequencing (WGRS), (Baird *et al.,* 2008), genotyping-by-sequencing (GBS) (Poland, J., and Rife, T., 2012) and sequence-based genotyping (SBG) (Truong *et al.* 2012) have provided fabulous means that enable identification of large number of genetic markers throughout the genome.

Recently, the rapid detection of single nucleotide polymorphism (SNP) markers associated with disease resistance genes by the high-throughput genotyping methods combined with the next-generation sequencing (NGS) technologies has substantially shortened the time required for genetic map construction, quantitative trait loci (QTL) analysis, candidate gene identification and marker assisted breeding.

They find also application in identifying variation, investigating genetic diversity, genotypic fingerprinting and phylogenetic analysis (Rafalski, A. 2002; Varshney *et al*., 2009; Kumar *et al*. 2012; Thomson, M.J. 2014; Xu *et al*., 2016, Barka, G. and Lee, J., 2020).

The worldwide accessibility to sequencing platforms as well the pipelines for analysis and management of data make the application of those technologies accessible to scientists of different branches as pathologists, geneticists, and breeders for better control of diseases (Parisi *et al*., 2020).

The availability of NGS technologies and *C.annuum* reference genome provides us the opportunity to employ the resequencing strategy facing some specific and practical issues in genome-assisted breeding schemes (Kang *et al*., 2016).

In plant breeding, the availability of a contiguous genome represent a tool that underlie genetics traits and how they interact with their environment in different backgrounds. At the simplest level, it allows the association of genetic markers for selection and introgression of traits in order to enable crop improvement.

A high-quality genome serves as a tool for studies with higher statistical power towards the identification of genomic segments and genes responsible for economically important traits (Hulse *et al*., 2018). The first whole-genome sequences of *C. annuum* Zunla-1 and Chiltepin (*C. annuum* var. *glabriusculum*) were released by Qin *et al*. (2014).

The development of resequencing technology, give us thus the possibility to identify all possible polymorphisms between target parental lines and select highly informative variations basing on available knowledge, such as known gene function and/or collinearity (Kang *et al*., 2016).

Keeping also in mind that this technique can be applied easily to the large number of accessions of domesticated and wild species stored in the world seed banks that represent a valuable resource for breeding in term of resistance traits to various biotic stresses (Parisi *et al*., 2020).

In this context we resequenced, as a pilot analysis, five genotypes from the company collection (described in Chapter 5), codified as PM, TSW, L3, L4 and CM (due to resistance genes to Powdery mildew, *Tomato spotted wilt virus*, *L3* and *L4* for *Tobamovirus* and *Cucumber mosaic virus*, they harbor respectively).

This analysis is going to be in the future applied to a broader selection of most promising genotypes in the company collection. A global picture of variation between selected genotypes was obtained including SNP and InDel polymorphisms, both in coding and intergenic/intronic, putatively regulative regions.

Molecular markers can be rapidly developed for SNPs, which are known to be the most abundant polymorphisms with unlimited nucleotide variations between individual organisms, even within the same species (Rafalski, A. 2002). In fact, in our study more than one milion of SNP were detected for each genome. After single nucleotide polymorphisms (SNPs), InDels are known to be the second most common type of genetic variation and are also ubiquitously distributed, with massive variation in size (Păcurar *et al*., 2012).

The analysis of genetic distances based on polymorphisms identified, performed with PAST (Paleontological and Statistics) software, highlighted a very narrow genetic base of the cultivated varieties used in the breeding program. Private (occuring only in one genotype) polymorphism were then inferred. Based on their distribution, SNPs are considered private when are find in a single or a range of samples of a collection and absent in the rest, due to mechanisms such adaptation or as human-based selection (Acquadro *et al*., 2020). Even if private SNPs are not reported to be crucial for determining major changes of plant phenome,

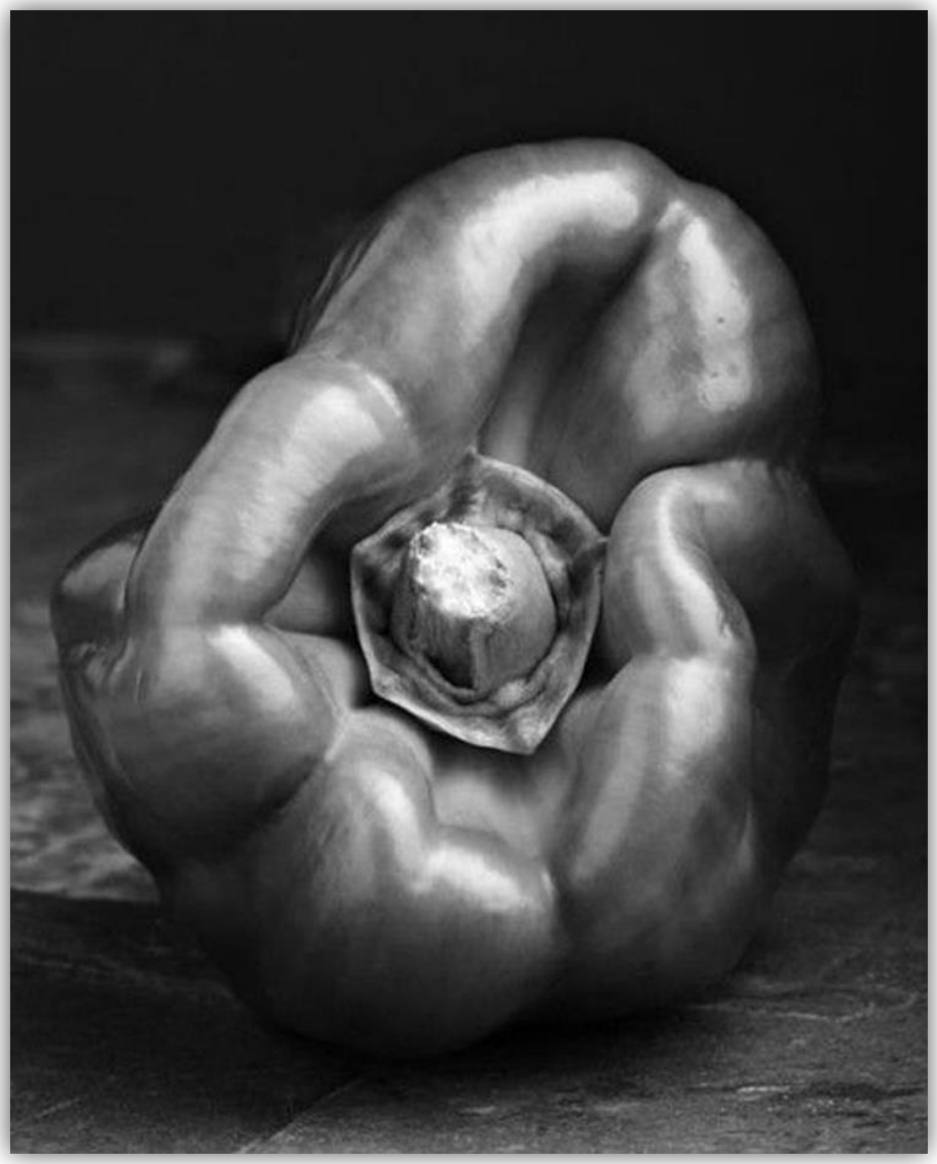
their investigation could highlight the existence of potential candidate genes for traits of agricultural interest or be applied for a precise genotype identification/traceability/purity.

Relatively few private polymorphisms were identified in the present study as most of mutations (respect to reference genome) are shared with at least one other genotypes suggesting the need to broaden the company's germplasm genetic base by introducing new material of different origin.

A broader sets of private polymorphisms limited however only to those located within exonic regions with the priority given to non-synonymous mutations were obtained for the further future studies of candidate genes. Most of them are annotated as "unknown"; however, as different pepper genome sequences are available, the private SNPs should be check regurarly along with the information related if they are expressed or not.

A restricted panel of 8-10 private SNPs giving origin to codon stop gain or loss (that lead to shorter or highly modified protein that most probably loose or change their function) were identified in each genotype distributed on most of the chromosomes. Those SNPs after being transformed into molecular markers can be applied in the company for the traceability and variety identification. With a nowadays costs of resequencing a seed company may characterize easily complete genomic profiles of their varieties, and mine those data for genes/markers useful for breeding programs but also to create a kind of "varietal identity card" to protect their intellectual property. In the present work, we considered private SNPs those identified only in one of the genotypes analyzed as a kind kick-off set for proprietary collection characterization. We are aware that molecular characterization of further genotypes will restrict the number of polymorphisms limited to just one accessions, however a set of 46 identified polymorphisms will enable to obtain absolutely unrepeatable profiles for all accessions in the collection.

The activity done and the outcomes so far gathered can be considered of high importance because they represent one of the few examples of practical molecular work carried out by a private Italian seed company on its genetic materials with aid of academic research. The hope is that it will be possible to increase the genetic gain with effects at the commercial level.



Edward Henry Weston (1886-1958)

.

Chapter 8

References

Acquadro, A., Barchi, L., Portis, E., Nourdine, M., Carli, C., Monge, S., Valentino, D., Lanteri, S. (2020). Whole genome resequencing of four Italian sweet pepper landraces provides insights on sequence variation in genes of agronomic value. Sci Rep. 2020; 10: 9189.

Agarwal, M., Shrivastava, N. (2008). Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. v. 27, 617–631.

Ahn, Y., Manivannan, A., Karna, S., Jun, T., Yang, E., Choi, S., Lee, E. (2018). Whole genome resequencing of capsicum baccatum and capsicum annuum to discover single nucleotide polymorphism related to powdery mildew resistance. Scientific Reports, 8(1) doi:10.1038/s41598-018-23279-5.

Anand, N., Deshpande, A.A., Sridhar, T.S. (1987). Resistance to powdery mildew in an accession of Capsicum frutescens and its inheritance pattern. Capsicum Eggplant Newsl 6, 77–78.

Andrés, A., Luis Arteaga, M., Gil Ortega, R. (2004). New genes related to PVY resistance in C. annuum L. 'Serrano Criollo de Morelos-334'. In Proceedings of the XIIth EUCARPIA Meeting on genetics and breeding of Capsicum and eggplant, Noordwijkerhout, The Netherlands, R.E., Ed., Pubblisher Plant Research International: Wageningen UR, The Netherlands, pp. 134–138.

Antonio, A. S., Wiedemann, L. S. M., Veiga Junior, V. F. (2018). The genus: Capsicum: a phytochemical review of bioactive secondary metabolites. RSC Adv. 8, 25767–25784.

Arnedo-Andrés, M. S., Gil-Ortega, R., Luis-Arteaga, M., Hormaza, J. I. (2002). Development of RAPD and SCAR markers linked to the Pvr4 locus for resistance to PVY in pepper (Capsicum annuum L.). Theor. Appl. Genet. 105, 1067–1074.

Atkinson, J. A., Jackson, R. J., Bentley, A. R., Ober, E., and Wells, D. M. (2018). Field phenotyping for the future. Annu. Plant Rev. Online. doi: 10.1002/9781119312994.apr0651.

Avilla, C., Collar, J.L., Duque, M., Pérez, P., Fereres, A. (1997). Impact of floating row covers on bell pepper yield and virus incidence. Hortsci. v. 32, 882–883.

Babu, B.S., Pandravada, S.R., Pasada Rao, R.D.V.J., Anitha, K., Chakrabarty, S.K., Varaprasad, K.S. (2011). Global sources of pepper genetic resources against arthropods, nematodes and pathogens. Crop Prot 30:389–400.

Baenziger, P. S. (2006). Plant breeding training in the US. Hortscience 41, 40–44. doi: 10.21273/HORTSCI.41.1.40.

Bai, Y., Pavan, S., Zheng, Z., Zappel, N.F., Reinstadler, A., Lotti, C., De Giovanni, C., Ricciardi, L., Lindhout, P., Visser, R., Theres, K., Panstruga, R. (2008). Naturally

occurring broad-spectrum powdery mildew resistance in a Central American tomato accession is caused by loss of mlo function. Mol Plant Microbe Interact 21:30–39.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3: e3376. doi: 10.1371/journal.pone.0003376.

Barbary, A., Djian-Caporalino, C., Marteu, N., Fazari, A., Caromel, B., Castagnone-Sereno, P., Palloix, A. (2016). Plant genetic background increasing the efficiency and durability of major resistance genes to root-knot nematodes can be resolved into a few resistance QTLs. Front. Plant Sci. 7, 632.

Barchenger, D.W., Naresh, P., Kumar, S. (2019). Genetic resources of Capsicum. The Capsicum Genome, Compendium of Plant Genomes. p. 9-23. Springer Nature Switzerland AG, Switzerland.

Barka, G.D., Lee, J. (2020). Molecular Marker Development and Gene Cloning for Diverse Disease Resistance in Pepper (Capsicum annuum L.): Current Status and Prospects. Plant Breed. Biotech. 2020 (June) 8(2):89~113. https://doi.org/10.9787/PBB.2020.8.2.89.

Ben Chaim, A., Grube, R.C., Lapidot, M., Jahn, M., Paran, I. (2001). Identification of quantitative trait loci associated with resistance to Cucumber mosaic virus in Capsicum Annum. Theor. Appl. Genet, 102, 1213–1220.

Ben-Chaim, A., Brodsky, Y., Falise, M., Mazourek, M. (2006). QTL analysis for capsaicinoid content in Capsicum. Theo. Appl. Genet. 113: 1481-1490.

Berger, A., Henderson, M., Nadoolman, W., Duffy, V. (1996). Oral capsaicin provides temporary relief for oral mucositis pain secondary to chemotherapy/radiation therapy. J Pain Symptom Manage. 10: 243-248.

Bernardo, R. (2002). Breeding for Quantitative Traits in Plants. Stemma Press, Woodbury, Minnesota, 369 pp.

Bhat, R.A., Miklis, M., Schmelzer, E., Schulze-Lefert, P., Panstruga, R. (2005). Recruitment and interaction dynamics of plant penetration resistance components in a plasma membrane microdomain. Proc Natl Acad Sci USA 102:3135–3140.

Boiteux L.S. (1995). Allelic relationships between genes for resistance to tomato spotted wilt tospovirus in Capsicum chinense. Theor Appl Genet 1995 Jan;90(1):146-9. doi: 10.1007/BF00221009.

Bonnet, D.G., Rebetzke, G.J. and Spielmeyer, W. (2005). Strategies for efficient implementation of molecular markers in wheat breeding. Molecular Breeding 15, 75-85.

Bosland, P.W. (1992). "Chiles: a diverse crop." HortTechnology, 2 (1): 6-10.

Boukema, I.W. (1980). Allelism of genes controlling resistance to TMV in Capsicum, L. Euphytica 29, 433–439.

Bosland, P.W., Votava E.J. (2000). "Peppers: vegetable and spice Capsicums", Departement of Agronomy and Horticulture, New Mexico State University, Las Cruces, USA. CABI publishing.

Caranta, C., Palloix, A., Lefebvre, V., Daubéze, A.M. (1997). QTLs for a component of partial resistance to Cucumber mosaic virus in pepper: Restriction of virus installation in host-cells. Theor. Appl. Genet., 94, 431–438.

Caranta, C., Thabuis, A., Palloix, A. (1999). Development of a CAPS marker for the Pvr4 locus: A tool for pyramiding potyvirus resistance genes in pepper. Genome 42, 1111–1116.

Caranta, C., Pflieger, S., Lefebvre, V., Daubèze, A.M., Thabuis, A., Palloix, A. (2002). QTLs involved in the restriction of Cucumber mosaic virus (CMV) long-distance movement in pepper. Theor. Appl. Genet. 104, 586–591.

Cebolla-Cornejo, J., Soler, S., Gomar, B., Soria, M.D., Nuez, F. (2003). Screening Capsicum germplasm for resistance to Tomato spotted wilt virus (TSWV). Ann. Appl. Biol. 143, 143–152.

Celik, I., Sogut, M. A., Ozkaynak, E., Doganlar, S., Frary, A. (2016). Physical mapping of NBS coding resistance genes to the Me-gene cluster on chromosome P9 reveals markers tightly linked to the N gene for root-knot nematode resistance in pepper. Mol. Breed. 36, 1–7.

Changkwian, A., Venkatesh, J., Lee, J.H., Han, J.W., Kwon, J.K., Siddique, M.I., Solomon, A.M., Choi, G.J., Kim, E., Seo, Y., Kim, Y.H., Kang, B.C. (2019). Physical localization of the root-knot nematode (Meloidogyne incognita) resistance locus Me7 in pepper (Capsicum annuum). Frontiers in Plant Science, 10, art. no. 886.

Charmet, G., Robert, N., Perretant, M.R., Gay, G., Sourdille, P., Groos, C., Bernard, S. and Bernard, M. (1999). Marker-assisted recurrent selection for cumulating additive and interactive QTLs in recombinant inbred lines. Theoretical and Applied Genetics 99, 1143–1148.

Cheng, S.S. (1989). "The use of Capsicum chinense as sweet pepper cultivars and source for gene transfer." In: Griggs, T.D. e McLean B.T. (eds.) Tomato and Pepper Production in the Tropics. Asian Vegetable Research & Developement Center, Taiwan, pp. 55-62.

Choi, H.W., Kim, Y.J., Lee, S.C., Hong, J.K., Hwang, B.K. (2007). Hydrogen peroxide generation by the pepper extracellular peroxidase CaPO2 activates local and systemic cell death and defense response to bacterial pathogens. Plant Physiol. 145, 890–904.

Choi, S., Lee, J.H., Kang, W.H., Kim, J., Huy, H.N., Park, S.W., Son, E.H., Kwon, J.K., Kang, B.-C. (2018). Identification of cucumber mosaic resistance 2 (cmr2) that confers resistance to a new cucumber mosaic virus isolate p1 (cmv-p1) in pepper (capsicum spp.). Frontiers in Plant Science, 9, art. no. 1106.

Cobb, J.N., Biswas, P.S., Platten, J.D. (2019). Back to the future: revisiting MAS as a tool for modern plant breeding.Theor. Appl. Genet. 132: 647-667.

Cobb, J.N., Juma, R.U., Biswas, P.S., Arbelaez, J.D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., Ng, E.H. (2019). Enhancing the Rate of Genetic Gain in Public-Sector Plant Breeding Programs: Lessons from the Breeder's Equation. Theor. Appl. Genet 2019, 132, 627–645.

Collard, B.C., Jahufer, M.Z., Brouwer, J.B. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. 79 Euphytica v. 142, 169–196.

Collard, B.C.Y., Mackill, D.J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of The Royal Society B 363, 557–572.

Cooper, M., Technow, F., Messina, C., Gho, C., and Radu Totir, L. (2016). Use of crop growth models with whole-genome prediction: application to a maize multienvironment trial. Crop Sci. 56, 2141–2156. doi: 10.2135/cropsci2015.08.0512.

Costa-Neto, G., Fritsche-Neto, R., Crossa, J. (2020). Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. Heredity (Edinb). 126, 92–106. doi: 10.1038/s41437-020-00353-1.

Covarrubias-Pazaran, G., Zelalem G., Dorcus G., Christian W., Marlee L., Solomon S., Peter C., Ismail R., Siraj I., Elizabeth P., Edward K., Edwige G., Afolabi A., Peter K., Michael Q., Jan D. (2022). Breeding schemes: what are they, how to formalize them, and how to improve them? Frontiers in Plant Science, 21 January 2022, doi:10.3389/fpls.2021.791859.

CPVO, Community Plant Variety Office. Available on line: https://cpvo.europa.eu/en/applications-and-examinations/technical-examinations/technical-protocols.

Csilléry, G. (2006). Pepper taxonomy and the botanical description of the species. Acta Agron. Hungarica 54, 151–166.

Cvetkovi, T., Ranilovi, J., Joki, S. (2022). Quality of Pepper Seed By-Products: A Review. Foods 2022, 11, 748.

Daubèze, A.M., Hennart, J.W., Palloix, A. (1995). Resistance to Leveillula taurica in pepper (Capsicum annuum) is oligogenically controlled and stable in Mediterranean regions. Plant Breed. 114, 327–332.

De Almeida-Engler, J., Favery, B., Engler, G., Abad, P. (2005). Loss of susceptibility as an alternative for nematode resistance. Curr Opin Biotechnol 16:112–117.

De Los Campos, G., Pérez-Rodríguez, P., Bogard, M., Gouache, D., and Crossa, J. (2020). A data-driven simulation platform to predict cultivars' performances under uncertain weather conditions. Nat. Commun. 11:4876. doi: 10.1038/s41467-020-18480-y.

Devran, Z., Kahveci, E., Özkaynak, E., Studholme, D. J., Tör, M. (2015). Development of molecular markers tightly linked to Pvr4 gene in pepper using next-generation sequencing. Mol. Breed. 35 .

Di Dato, F., Parisi, M., Cardi, T., Tripodi, P. (2015). Genetic diversity and assessment of marker linked to resistance and pungency genes in Capsicum germplasm. Euphytica (2015) 204:103-119. DOI 10.1007/s 10681-014-1345-4.

Dilday, R.H. (1990). Contribution of ancestral lines in the development of new cultivars of rice. Crop Science 30, 905–911.

Dirks, R.H.G., Dun, C.M.P.V., Reinink, K., Wit, J.P.C.D. (2006). Reverse Breeding. US20060179498A1. United States, United States Patent and Trademark Office.

Djian-Caporalino, C., Pijarowski, L., Fazari, A., Samson, M., Gaveau, L., O'Byrne, C., Lefebvre, V., Caranta, C., Palloix, A., Abad, P. (2001). High-resolution genetic mapping of the pepper (Capsicum annuum L.) resistance loci Me3 and Me4 conferring heat-stable resistance to root-knot nematodes (Meloidogyne spp.). Theor. Appl. Genet.103, 592–600.

Djian-Caporalino, C., Fazari, A., Arguel, M.J., Vernie, T., VandeCasteele, C., Faure, I., Brunoud, G., Pijarowski, L., Palloix, A., Lefebvre, V., Abad, P. (2007). Root-knot nematode (Meloidogyne spp.) Me resistance genes in pepper (Capsicum annuum L.) are clustered on the P9 chromosome. Theoretical and Applied Genetics, 114 (3), pp. 473-486.

Djian-Caporalino, C., Palloix, A., Fazari, A., Marteu, N., Barbary, A., Abad, P., Sage-Palloix, A.M., Mateille, T., Risso, S., Lanza, R., *et al*. (2014). Pyramiding, alternating or mixing: Comparative performances of deployment strategies of nematode resistance genes to promote plant resistance efficiency and durability. BMC Plant Biol. 2014, 14, 53.

Dombrovsky, A., Smith, E. (2017). Seed Transmission of Tobamoviruses: Aspects of Global Disease Distribution. Adv. Seed Biol. (2017). doi:10.5772/intechopen.70244.

Dudley, D.N., Saghai Maroof, M.A. and Rufener, G.K. (1991). Molecular markers and grouping of parents in a maize breeding program. Crop Science 31, 718–723.

Edwards, M., Johnson, L. (1994). RFLPs for rapid recurrent selection. In: Proceedings of Symposium on Analysis of Molecular Marker Data. American Society of Horticultural Science and Crop Science Society of America, Corvallis, Oregon, pp. 33–40.

Elad, Y., Messika, Y., Brand, M., David, D. R., Sztejnberg, A. (2007). Effect of microclimate on Leveillula taurica powdery mildew of sweet pepper. Phytopathology 97, 813–824.

Endelman, J.B., Atlin, G.N., Beyene, Y., Semagn, K., Zhang, X., Sorrells, M.E., Jannink, J. (2014). Optimal Design of Preliminary Yield Trials with Genome-Wide Markers. Crop Sci., 54, 48–59.

Esquinas-Alcazar, J.T. (1993). Plant genetic resources. In: Hayward, M.D., Bosemark, N.O. and Romagosa, I. (eds) Plant Breeding: Principles and Prospects. Chapman & Hall, London, pp. 33–51.

Eun, M. H., Han, J. H., Yoon, J. B., Lee, J. (2016). QTL mapping of resistance to the Cucumber mosaic virus P1 strain in pepper using a genotyping-by-sequencing analysis. Hortic. Environ. Biotechnol. 57, 589–597.

Falconer, D.S.; Mackay, T.F.C. (1996). Introduction to Quantitative Genetics. In Harlow Essex UK Longmans Green; Pearson: Essex, UK; Volume 3.

Faostat 2019-2020. Available online: http://www.fao.org/ (accessed on March 2022).

Fazari, A., Palloix, A., Wang, L., Yan Hua, M., Sage-Palloix, A.-M., Zhang, B.X., Djian-Caporalino, C. (2012). The root-knot nematode resistance N-gene co-localizes in the Me-genes cluster on the pepper (Capsicum annuum L.) P9 chromosome. Plant Breeding, 131 (5), pp. 665-673.

Frisch, M., Bohn, M., Melchinger, A. E. (1999). Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci. 39, 1295–1301.

Geleta, L.F., Labuschagne, M.T., Viljoen, C.D. (2004). Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. Plant Breeding 123, 467–473.

Gaynor, R. C., Gorjanc, G., and Hickey, J. M. (2021). AlphaSimR: an R package for breeding program simulations. G3 11:jkaa017. doi: 10.1093/g3journal/jkaa017.

Gene Script software, https://www.genscript.com/tools/restrictionenzyme-map-analysis.

Gepts, P., Hancock, J. (2006). The future of plant breeding. Crop Sci. 46, 1630–1634. doi: 10.2135/cropsci2005-12-0497op.

Gianessi, L. P., Reigner, N. (2005). The Value of Fungicides in U.S. Crop Production. Crop Prot. Res. Inst. 1–243.

Gizlice, Z., Carter, T.E., Jr and Burton, J.W. (1993). Genetic diversity in North American soybean: II. Prediction of heterosis in F2 populations of southern founding stock using genetic similarity measures. Crop Science 33, 620–626.

Godshalk, E.B., Lee, M. Lamkey, K.R. (1990). Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. Theoretical and Applied Genetics 80, 273–280.

Grube, R.C., Zhang, Y., Murphy, J.F., Loaiza-Figueroa, F., Lackney, V.K., Provvidenti, R., Jahn, M.K. (2000). New source of resistance to Cucumber mosaic virus in Capsicum frutescens. Plant Dis. 84, 885–891.

Guan, Y., Wang, B., Feng, Y., LI, P. (2015). Development and application of marker-assisted reverse breeding using hybrid maize germplasm. Journal of Integrative Agriculture 2015, 14(12): 2538–2546).

Guo, G.J.,Wang, S.B., Liu, J.B., Pan, B.G., Diao,W.P., Ge,W., Gao, C.Z., Snyder, J.C. (2017). Rapid identification of QTLs underlying resistance to Cucumber mosaic virus in pepper (Capsicum frutescens). Theor. Appl. Genet. 130, 41–52.

Hardham, A.R., Jones, D.A., Takemoto, D. (2007). Cytoskeleton and cell wall function in penetration resistance. Curr Opin Plant Biol 10:342–348.

Hayden, M. J., Nguyen, T. M., Waterman, A., Chalmers, K. J. (2008). Multiplex-Ready PCR: A new method for multiplexed SSR and SNP genotyping. BMC Genomics 9.

Heiser, C.B. (1976). "Pepper Capsicum (Solanaceae)." In: Simmonds, N.W. (ed.) Evolution of Crop Plants. Longman Press. London, pp.265-268.

Henryon, M., Berg, P., and Sørensen, A. C. (2014). Animal-breeding schemes using genomic information need breeding plans designed to maximise long-term genetic gains. Livest. Sci. 166, 38–47. doi: 10.1016/j.livsci.2014.06.016.

Heslot, N., Akdemir, D., Sorrels, M. E., and Jannink, J.-L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Gen. 127, 463-480. doi: 10.1007/s00122-013-2231-5.

Hiroshi, M., Takeo, S., Masashi, H., Tsukasa, N., Tatemi, Y. (2003). Dna markers linked to pepper mild mottle virus (PMMoV) resistant locus (L4) in Capsicu. J.Japan Soc. Hort. Sci, 72(3): 218 220.

Hoang, N., Yang, H.B., Kang, B.C. (2013). Identification and inheritance of a new source of resistance against Tomato spotted wilt virus (TSWV) in Capsicum. Sci. Hortic. 161, 8–14.

Hoefler, R., González-Barrios, P., Bhatta, M., Nunes, J.A.R., Berro, I., Nalin, R.S., Borges, A., Covarrubias, E., Diaz-Garcia, L., Quincke, M. (2020). Do Spatial Designs Outperform Classic Experimental Designs? J. Agric. Biol. Environ. Stat., 25, 523–552.

Holdsworth, W. L., Mazourek, M. (2015). Development of user-friendly markers for the pvr1 and Bs3 disease resistance genes in pepper. Molecular Breeding, 35(1). doi:10.1007/s11032-015-0260-2.

Holland, J. B. (2004). Implementation of molecular markers for quantitative traits in breeding programs-challenges and opportunities. In Proc. 4th Int. Crop Sci. Congress., Brisbane, Australia, 26 September—1 October.

Hospital, F., Charcosset, A. (1997). Marker-assisted introgression of quantitative trait loci. Genetics 147, 1469–1485.

Hospital, F., Goldringer, I. and Openshaw, S. (2000). Efficient marker-based recurrent selection for multiple quantitative trait loci. Genetical Research 75, 1181–1189.

Hulse-Kemp, A. M., Maheshwari, S., Stoffel, K., Hill, T. A., Jaffe, D., Williams, S. R., Van Deynze, A. (2018). Reference quality assembly of the 3.5-Gb genome of Capsicum annum from a single linked-read library. Horticulture Research, 5. doi:10.1038/s41438-017-0011 0.

Humphry, M., Consonni, C., Panstruga, R. (2006). mlo-based powdery mildew immunity: silver bullet or simply nonhost resistance? Mol Plant Pathol 7:605–610.

IBPGR (1983). "Genetic Resources of Capsicum: A Global Plan of Action." International Board for Plant Genetic Resources. AGPG/IBPGR/82/12, Rome, Italy, pp. 49.

Invitrogen, https://www.thermofisher.com.

ISTAT, **2020**. www.istat.it/it/agricoltura.

Jacquin, N.J. (1776). "Hortus Botanicus Vindobonensis." Vienna.

Jamann, T. M., Poland, J. A., Kolkman, J. M., Smith, L. G., Nelson, R. J. (2014). Unraveling genomic complexity at a quantitative disease resistance locus in maize. Genetics 198, 333– 344.

Janzac, B., Fabre, M.F., Palloix, A., Moury, B. (2008). Characterization of a new potyvirus infecting pepper crops in Ecuador. Arch. Virol. 2008, 153, 1543–1548.

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucour, J., Lorgeou, J. (2014). A reaction norm model for genomic selection using highdimensional genomic and environmental data. Theor. Appl. Genet. 127, 595–607. doi: 10.1007/s00122-013-2243-1.

Jarvis, P., Lister, C., Szabo, V. (1994). Integration of CAPS markers into the RFLP map generated using recombi- nant inbred lines of Arabidopsis thaliana. Plant Mol Biol. v.24, 685–687.

Jo, J., Purushotham, P.M., Han, K., Lee, H.R., Nah, G., Kang, B.C. (2017). Development of a genetic map for onion (Allium cepa L.) using reference-free genotyping-by-sequencing and SNP assays. Front. Plant Sci. 14, 1606.

Jo, J., Venkatesh, J., Han, K., Lee, H.Y., Choi, G.J., Lee, H.J., Choi, D., Kang, B.C. (2017). Molecular mapping of PMR1, a novel locus conferring resistance to powdery mildew in pepper (Capsicum annuum). Frontiers in Plant Science, 8, art. no. 2090.

John, G.K., Williams, A., Kubelik, R., Kenneth, J., Antoni J. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. v. 18, 6531–6535.

Johnson, G.R. (2004). Marker assisted selection. Plant Breeding Reviews 24, 293–310.

Jones, J.B., Lacy, G.H., Bouzar, H., Stall, R.E., Schaad, N.W. (2004). Reclassification of the xanthomonads associated with bacterial spot disease of tomato and pepper. Syst. Appl. Microbiol. 2004, 27, 755–762.

Kalia, R.K., Rai, M.K., Kalia, S. (2011). Microsatellite markers: an overview of the recent progress in plants. Euphytica. Euphytica. v. 177, 309–334 (2011).

Kang, B.C., Nahm, S.H., Huh, J.H., Yoo, H.S. (2001). An interspecific (Capsicum annuum x C. Chinense) F2 linkage map in pepper using RFLP and AFLP markers. Theo. Appl. Genet. 102: 531-539.

Kang, B. C., Yeam, I., Frantz, J. D., Murphy, J. F., Jahn, M. M. (2005). The pvr1 locus in Capsicum encodes a translation initiation factor elF4E that interacts with Tobacco etch virus VPg. Plant J. 42, 392–405.

Kang, W.H., Hoang, N.H., Yang, H.B., Kwon, J.K., Jo, S.H., Seo, J.K., Kim, K.H., Choi, D., Kang, B.C. (2010). Molecular mapping and characterization of a single dominant gene controlling CMV resistance in peppers (Capsicum annuum L.). Theoretical and Applied Genetics, 120 (8), pp. 1587-1596.

Kang, Y.J, Ahn, Y.K., Kim, K.T., Jun, T.H. (2016). Resequencing of Capsicum annuum parental lines (YCM334 and Taean) for the genetic analysis of bacterial wilt resistance. BMC Plant Biology volume 16, article number: 235.

Kay, S., Hahn, S., Marois, E., Hause, G., Bonas, U. (2007). A bacterial effector acts as a plant transcription factor and induce a cell size regulator. Science 318:648–651.

Kenyon, L., Kumar, S., Tsai, W. S., Hughes, J. D. A. (2014). Virus Diseases of Peppers (Capsicum spp.) and Their Control. Advances in Virus Research 90 (Elsevier Inc.).

Kesawat, M. S. and Das Kumar, B. (2009). Molecular markers: It's application in crop improvement. J. Crop Sci. Biotechnol. 12, 169–181.

Kim, D.S., Hwang, B.K. (2012). The pepper MLO gene, CaMLO2, is involved in the susceptibility cell death response and bacterial and oomycete proliferation. PlantJ. 72, 843–855.

Kim, S.B., Kang,W.H., Huy, H.N., Yeom, S.I., An, J.T., Kim, S., Kang, M.Y., Kim, H.J., Jo, Y.D., Ha, Y. (2017). Divergent evolution of multiple virus-resistance genes from a progenitor in Capsicum spp. New Phytol. 213, 886–899.

Kim, H., Yoon, J. B., and Lee, J. (2017). Development of fluidigm SNP type genotyping assays for marker-assisted breeding of chili pepper (Capsicum annuum L.). Korean J. Horticult. Sci. Technol. 35, 465–479. doi: 10.12972/kjhst.20170050.

Kim, H.J., Nahm, S.H., Lee, H.R., Yoon, G.B., Kim, K.T., Kang, B.C., Choi, D., Kweon, O.Y., Cho, M.C., Kwon, J.K. (2008). BAC-derived markers converted from RFLP linked to Phytophthora capsici resistance in pepper (Capsicum annuum L.). Theor. Appl. Genet. 118, 15–27.

Kim, H.J., Han, J.H., Yoo, J.H., Cho, H.J., Kim, B.D. (2008). Development of a sequence characteristic amplified region marker linked to the L4 locus conferring broad spectrum resistance to tobamoviruses in pepper plants. Molecules and Cells, 25 (2), pp. 205-210.

Kim, N., Kang, W.H., Lee, J. and Yeom, S. I. (2019). Development of Clustered Resistance Gene Analogs-Based Markers of Resistance to Phytophthora capsici in Chili Pepper. Biomed Res. Int. 2019.

Kim, S.B., Lee, H.Y., Seo, S., Lee, J.H., Choi, D. (2015). RNA Dependent RNA Polymerase (NIb) of the Potyvirusesis an avirulence factor for the broadspectrum resistance gene Pvr4 in Capsicum annuum cv. CM334. PLoS One v.10.

Kiran, K., Rawal, H.C., Dubey, H., Jaswal, R., Bhardwaj, S.C., Prasad, P., Pal, D., Devanna, B.N., Sharma, T.T. (2017). Dissection of genomic features and variations of three pathotypes of Puccinia striiformis through whole genome sequencing. Sci. Rep. 7, 42419.

Kloppers, F. J., Pretorius, Z. A. (1997). Effects of combinations amongst genes Lr13, Lr34 and Lr37 on components of resistance in wheat to leaf rust. Plant Pathol. 46, 737–750. (doi:10.1046/j.1365-3059.1997.d01-58.x).

Kumar, S., Banks, T.W., Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. Int. J. Plant Genomics 2012: 831460.

Kumar, R., Sharma, V., Suresh, S., Ramrao, D.P., Veershetty, A., Kumar, S., Priscilla, K., Hangargi, B., Narasanna, R., Pandey, M.K., Naik, G.R., Thomas, S., Kumar, A. (2021). Understanding Omics Driven Plant Improvement and de novo Crop Domestication: Some Examples. Front. Genet. Sec. Systems Biology Archive Volume 12-2021. https://doi.org/10.3389/fgene.2021.637141.

Lamour, K. H., Daughtrey, M. L., Benson, D. M., Hwang, J., Hausbeck, M. K. (2003). Etiology of Phytophthora drechsleri and P. nicotianae (P. parasitica) diseases affecting floriculture crops. Plant Dis. 87, 854–858.

Lantieri, S., Pickersgill B. (1993). Chromosomal structural changes in Capsicum annum L. and Capsicum annum Jacq. Euphytica, 67:155-160.

Lee, O.H., Hwang, H.S., Kim, J.Y., Han, J.H., Yoo, Y.S., Kim, B.S. (2001). A search for sources of resistance to powdery mildew (Leveillula taurica (Lev.) Arn.) in pepper (Capsicum spp.). Korean J. Hort. Sci. Tech. 19.

Lee, M.Y., Lee, J.H., Ahn, H.I., Yoon, J.Y., Her, N.H., Choi, J.K., Choi, G.S., Kim, D.-S., Ryu, G.H. (2006). Identification and sequence analysis of RNA3 of a resistance-breaking Cucumber mosaic virus isolate on Capsicum annuum. Plant Pathol. J. 22, 265–270.

Lee, J., Han, J.H., Yoon, J.A. (2012). Set of Allele-specific Markers Linked to L Locus Resistant to Tobamovirus in Capsicum spp.(2012) Kor. J. Hort. Sci. Technol. 30(3):286-293 DOI http://dx.doi.org/10.7235/hort.2012.12018.

Lee, M., Godshalk, E.B., Lamkey, K.R. Woodman, W.L. (1989). Association of restriction length polymorphism among maize inbreds with agronomic performance of their crosses. Crop Science 29, 1067–1071.

Lee, M. (1995). DNA markers and plant breeding programs. Advances in Agronomy 55, 265–344.

Lee, Y.H., Kim, H.S., Kim, J.Y., Jung, M. (2004). A new selection method for pepper transformation: callus-mediated shoot formation. Pl. Cell Rep. 23: 50–8.

Lefebvre, V., Palloix, A. (1996). Both epistatic and additive effects of QTLs are involved in polygenic induced resistance to disease: A case study, the interaction pepper-Phytophthora capsici Leonian. Theor. Appl. Genet. 93, 503–511.

Lefebvre, V., Pflieger, S., Thabuis, A., Caranta, C., Blattes, A., Chauvet, J.C., Daubeze, A.M., Palloix, A. (2002). Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. Genome 45, 839–854.

Lefebvre, V., Daubèze, A.M., Rouppe van der Voort, J., Peleman, J., Bardin, M., Palloix, A. (2003). QTLs for resistance to powdery mildew in pepper under natural and artificial infections. Theor. Appl. Genet. 107, 661–666.

LGC, **biosearch technologies.** https://www.biosearchtech.com/.

Li, H., Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14):1754-1760.

Li, H., Handsaker, B., Wysoker, A. *et al*. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25(16):2078-2079.

Li, N., Yin, Y., Wang, F., Yao, M. (2018). Construction of a high-density genetic map and identification of qtls for cucumber mosaic virus resistance in pepper (Capsicum annuum l.) using specific length amplified fragment sequencing (slaf-seq). Breeding Science, 68 (2), pp. 233-241.

Li, X., Zhu, C., Wang, J., and Yu, J. (2012). Com.puter simulation in plant breeding. Adv. Agron. 116, 219–264. doi: 10.1016/B978-0-12-394277-7.00 006-3.

Lippman, Z.B., Zamir, D. (2007). Heterosis: revisiting the magic. Trends in Genetics 23, 60–66.

Liu, W., Kang, J., Jeong, H., Choi, H., Yang, H., Kim, K., Kang, B. (2014). Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to phytophthora capsici in pepper. Theoretical and Applied Genetics, 127(11), 2503-2513. doi:10.1007/s00122-014-2394-8.

Livingstone, K.D., Lackney, V.K., Blauth, J., Wijk, V.R., Jahn, M.K. (1999). Genome mapping in Capsicum and the evolution of genome structure in the Solanaceae. Genetics. 152: 1183-1202.

Lorenz, A.J., Chao, S., Asoro, F.G., Heffner, E.L., Hayashi, T., Iwata, H., Smith, K.P., Sorrells, M.K., Jannink, J.L. (2011). Genomic selection in plant breeding: knowledge and prospects. Adv. Agron., 2011, 110, 77-123.

Lozada, D.N., Bosland, P.W., Barchenger, D.W., Haghshenas-Jaryani, M., Sanogo, S., Walker, S. (2022). Chile Pepper (Capsicum) Breeding and Improvement in the "Multi-Omics" Era. Front. Plant Sci., Sec. Plant Breeding. https://doi.org/10.3389/fpls.2022.879182.

Lu, F.H., Kwoon, S.H., Yoon, M.I., Kim, K.T., Cho, M.C., Yoon, M.K., Park, Y.J. (2012). SNP marker integration and QTL analysis of 12 agronomic and morphological traits in F8 RILs of pepper (Capsicum annuum L.). Mol. Cells 34, 25–34.

Lusser, M., Parisi, C., Plan, D., Rodríguez-Cerezo, E. (2011). New plant breeding techniques, state-of-the-art and prospects for commercial development. JRC Scientific and Technical Reports. Publications Office of the European Union, Luxembourg. pp. 26–193.

Lusser, M., Parisi, C., Plan, D., Rodríguez-Cerezo, E. (2012). Deployment of new biotechnologies in plant breeding. Nature Biotechnology, 30, 231–239.

Lynch, M, W. B. (1998). Genetics and analysis of quantitative traits. (1998).

Macedo, M.A., Rojas, M.R., Gilbertson, R.L. (2019). First report of a resistance-breaking strain of Tomato spotted wilt orthotospovirus infecting sweet pepper with the Tsw resistance gene in California, U.S.A. Plant Dis. 103, 1048.

Maga, J.A. (1975). Capsicum. Crit. Rev. Food Sci. Nutr. 6: 177-199.

Mallard, S., Cantet, M., Massire, A., Bachellez, A., Ewert, S., Lefebvre, V. (2013). A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to Phytophthora capsici: A valuable locus for pepper breeding. Mol. Breed. 32, 349–364.

Mallard, S., Cantet, M., Massire, A., Bachellez, A., Ewert, S., Lefebvre, V. (2013). A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to phytophthora capsici: A valuable locus for pepper breeding. Molecular Breeding, 32(2), 349-364. doi:10.1007/s11032-013-9875-3.

Matsunaga, H., Saito, T., Hirai, M., Nunome, T., Yoshida, T. (2003). DNA markers linked to Pepper mild mottle virus (PMMoV) resistant locus (L4) in Capsicum. J. Jpn. Soc. Hortic. Sci. 72, 218–220.

Matthews, D., Kearney, J.F., Cromie, A., Hely, F.S., Amer, P.R. (2019). Genetic benefts of genomic selection breeding programmes considering foreign sire contributions. Genettic Selection Evolution 51:40 https://doi.org/10.1186/s12711-019-0483-5.

McLeod, M. J., Sheldon, I., Hardy Eshbaugh, W. (1982). Early evolution of chili peppers. Econ. Bot. 36, 361–368.

Melchinger, A.E., Lee, M., Lamkey, K.R., Woodman, W.L. (1990). Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. Crop Science 30, 1033–1040.

Messina, C. D., Technow, F., Tang, T., Totir, R., Gho, C., Cooper, M. (2018). Leveraging biological insight and environmental variation to improve phenotypic prediction: integrating crop growth models (CGM) with whole genome prediction (WGP). Eur. J. Agron. 100, 151–162. doi: 10.1016/j.eja.2018.01.007.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157, 1819-1829.

Millet, E. J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., CabreraBosquet, L., Lacube, S., *et al*. (2019). Genomic prediction of maize yield across European environmental conditions. Nat. Genet. 51, 952–956. doi: 10.1038/s41588-019-0414-y.

Minamiyama, Y., Tsuro, M., Kubo, T., Hirai, M. (2007). Qtl analysis for resistance to phytophthora in pepper using a high density SSR-based map. Breeding Science 57:129-134.

Mochizuki, T., Ohki, S. T. (2012). Cucumber mosaic virus: Viral genes as virulence determinants. Mol. Plant Pathol. 13, 217–225.

Molly, J., Ilan P., Katrin H., Elaine R., Kevin D.L., Rebecca C.G, Ester A., Moshe L., James M. (2000). Genetic Mapping of the Tsw Locus for Resistance to the Tospovirus Tomato spotted wilt virus in Capsicum spp. and Its Relationship to the Sw-5 Gene for Resistance to the Same Pathogen in Tomato Molly. Mol. Plant-Microbe Interact. v. 13, 673–682.

Mondini, L., Noorani, A. P. (2009). Assessing plant genetic diversity by molecular tools. Diversity v. 1.

Monteverde, E., Gutierrez, L., Blanco, P., Pérez de Vida, F., Rosas,J. E., Bonnecarrère, V., *et al*. (2019). Integrating molecular markers and environmental covariates to interpret genotype by environment interaction in rice (Oryza sativa L.) Grown in Subtropical Areas. G3 Genes|Genomes|Genetics 9, 1519–1531. doi: 10.1534/g3.119.400064.

Morais-Júnior, O. P., Duarte, J. B., Breseghello, F., Coelho, A. S. G., and Magalhães, A. M. (2018). Single-step reaction norm models for genomic prediction in multienvironment recurrent selection trials. Crop Sci. 58, 592–607. doi: 10.2135/cropsci2017.06.0366.

Morgante, M., Salamini, F. (2003). From plant genomics to breeding practice. Curr. Opin. Biotechnol., 2003, 14, 214-219.

Morris, M., Edmeades, G., Pehu, E. (2006). The global need for plant breeding capacity: what roles for the public and private sectors? Hortscience 41, 30–39. doi: 10.21273/HORTSCI.41.1.30.

Moulin, M., Rodrigues, R.. Bento, C., Gonçalves, L., Santos, J., Sudré, C., Viana, A, (2015). Genetic dissection of agronomic traits in Capsicum baccatum var. pendulum. Genet. Mol. Res., 14, 2122–2132.

Moulin, M., Rodrigues R., Ramos, H., Bento, C., Sudré, C., Gonçalves, L.,Viana, A. (2015). Construction of an integrated genetic map for Capsicum baccatum L. Genet. Mol. Res., 14, 6683–6694.

Moury, B., Palloix, A., Selassie, K.G., Marchoux, G. (1997). Hypersensitive resistance to Tomato spotted wilt virus in three Capsicum chinense accessions is controlled by a single gene and is overcome by virulent strains. Euphytica 94, 45–52.

Moury, B., Selassie, K.G., Marchoux, G., Daubeze, A.M., Palloix, A. (1998). High temperature e_ects on hypersensitive resistance to Tomato spotted wilt Tospovirus (TSWV) in pepper (Capsicum chinense Jacq.). Eur. J. Plant Pathol.104, 489–498.

Moury, B., Pflieger, S., Blattes, A., Lefebvre, V., Palloix, A. (2000). A CAPS marker to assist selection of tomato spotted wilt virus (TSWV) resistance in pepper. Genome 43, 137–142.

Mundt, C. C. (2015). Durable resistance: a key to sustainable management of pathogens and pests. Infect. Genet. Evol. v.27, 446–455.

NCBI - PubMed database, https://www.ncbi.nlm.nih.gov.

Nelson, R., Wiesner-Hanks, T., Wisser, R., Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. Nat. Rev. Genet. 19, 21–33.

Nigam, D., Garcia-Ruiz, H. (2020). Variation profile of the orthotospovirus genome. Pathogens 9, 1–28.

Niks, R. E., Qi, X., Marcel, T. C. (2015). Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions, and mechanisms. Annu. Rev. Phytopathol. v.53, 445– 470.

Nyquist, W.E. (1991). Estimation of heritability and prediction of selection response in plant populations. Critical Review of Plant Science 10, 235–322.

Nono-Womdim, R., Palloix, A., Gèbre-Selassie, K., Marchoux, G. (1993). Partial resistance of bell pepper to Cucumber mosaic virus movement within plants: Field evaluation of its e_ciency in southern France. J. Phytopathol 37, 125–132.

Ogundiwin, E.A., Berke, T.F., Massoudi, M., Black, L.L., Huestis, G., Choi, D., Lee, S., Prince, J.P. (2005). Construction of 2 intraspecific linkage maps and identification of resistance QTLs for Phytophthora capsici root-rot and foliar-blight diseases of pepper (Capsicum annuum L.). Genome 48, 698–711.

Oligo Explorer software, https://oligo-explorer.software.informer.com/1.5/.

Özkaynak, E., Devran, Z., Kahveci, E., Doğanlar, S., Başköylü, B., Doğan, F., İşleyen, M (2014). Pyramiding Multiple Genes for Resistance to PVY, TSWV and PMMoV in Pepper Using Molecular Markers. Europ.J.Hort.Sci., 79 (4). S. 233–239, 2014, ISSN 1611-4426.

Păcurar, D.I., Păcurar, M.L., Street, N., Bussell, J.D., Pop, T.I., Gutierrez, L., Bellini, C. (2012). A collection of INDEL markers for map-based cloning in seven Arabidopsis accessions. J Exp Bot. 2012 Apr; 63(7):2491-501. doi: 10.1093/jxb/err422. Epub 2012 Jan 25.

Paleontological Statistic Software, https://www.nhm.uio.no/english/research/resources/past/.

Panstruga, R. (2005). Serpentine plant MLO proteins as entry portals for powdery mildew fungi. Biochem Soc Trans 33:389–392.

Parisi, M., Di Dato, F., Minutolo, M., Festa, G., Alioto, D. (2015). Screening Capsicum spp. for tolerance to a resistance-breaking strain of Tomato spotted wilt virus by artificial inoculation. Plant Pathol. 2015, (Suppl. 97).

Parisi, M., Alioto, D., Tripodi, P. (2020). Overview of Biotic Stresses in Pepper (Capsicum spp.): Sources of Genetic Resistance, Molecular Breeding and Genomics, Molecular Breeding and Genomics. International Journal of Molecular Sciences. April 2020, DOI: 10.3390/ijms21072587.

Pavan, S., Zheng, Z., Van den Berg, P., Lotti, C., De Giovanni, C., Borisova, M., Lindhout, P., De Jong, H., Ricciardi, L., Visser, R., Bai, Y. (2008). Map- vs homology-based cloning for the recessive gene Mlo-2 conferring resistance to tomato powdery mildew. Euphytica 162:91–98.

Pickersgill, B. (1997). Genetic resources and breeding of Capsicum spp. Euphytica, 96(1), 129-133. 10.1023/A:1002913228101.

Pierre, M., Noel, L., Lahaye, T., Ballvora, A., Veuskens, J., Ganal, M., Bonas, U. (2000). High-resolution genetic mapping of the pepper resistance locus Bs3 governing recognition of the Xanthomonas campestris pv vesicatoria AvrBs3 protein. Theor. Appl. Genet. 101, 255–263.

Peleman, J.D., Van der Voort, J.R. (2003). Breeding by design. Trends Plant Sci., 8, 330-334.

Pérez-de-Castro, A.M., Vilanova, S., Cañizares, J., Pascual, L., Blanca, J.M., Díez, M.J,. Prohens J., Picó, B. (2012). Application of Genomic Tools in Plant Breeding. Current Genomics, 2012, 13, 179-195.

Perkins, J. M., and Jinks, J. L. (1968). Environmental and genotype-environmental components of variability. 3. Multiple lines and crosses. Heredity (Edinb). 23, 339–356. doi: 10.1038/hdy.1968.48.

Phillips, R.L. (2006). Genetic tools from nature and the nature of genetic tools. Crop Science 46, 2245–2252.

Phillips, R.L. (2008). Can genome sequencing of model plants be helpful for crop improvement? Proceedings of 5th International Crop Science Congress, 13–18 April 2008, Jeju, Korea. International Crop Science Society, Madison, Wisconsin.

Piffanelli, P., Ramsay, L., Waugh, R., Benabdelmouna, A., D'Hont, A., Hollricher, K., Jorgensen, J.H., Schulze-Lefert, P., Panstruga, R. (2004). A barley cultivation-associated polymorphism conveys resistance to powdery mildew. Nature 430:887–891.

Pinar, H., Mutlu, N., Ozasland, O., Argun, D., Keles, D., Canhilal R. (2016). Reliability assessment of molecular marker linked to resistance gene againts Meloidogyne spp. in diverse pepper genotype. Egyptian Journal of Biological Pest Control; Giza Vol. 26, Fasc. 3: 515-521.

Pinto, C.M.F., Santos, I.C., De Araujo, F.F., De Silva, T.P. (2016). Pepper importance and growth (Capsicum spp.), Production and breeding of chilli peppers (Capsicum spp.). p. 1-25, Springer Nature Switzerland AG, Switzerland.

Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C., Nelson, R. J. (2009). Shades of gray: the world of quantitative disease resistance. Trends Plant Sci. 14, 21–29.

Poland, J.A., Rife, T.W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. The Plant Genome Review & Interpretation. Open Access. https://doi.org/10.3835/plantgenome2012.05.0005.

Ponnam, N., Kumari, M., Acharya, G.C.,Anand, C., Lakshmana, R. (2019). Genetics and molecular markers for resistance to major soil borne pathogens in chilli (Capsicum annuum L.). Research Journal of Biotechnology Vol. 14.

Pook, T., Schlather, M., and Simianer, H. (2020). MoBPS-modular breeding program simulator. G3 10, 1915–1918. doi: 10.1534/g3.120.40 1193.

Prince, J., Pochard, E., Tanksley, D. (1993). Costruction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome, volume 36, n 3.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., Cheng, J., Zhao, S., Xu, M., Luo, Y., Yang, Y., Wu, Z., Mao, L., Wu, H., Changying, L.H., Zhou, H., Lin, H., González-Morales, S., Trejo-Saavedra, D.L., Tian, H., Tang, X., Zhao, M., Huang, Z., Zhou, A., Yao, X., Cui, J., Li, W., Chen, Z., Feng, Y., Niu, Y., Bi, S., Yang, X., Li, W., Cai, H., Luo, X., Montes-Hernández, S., Leyva-González, M.A., Xiong, Z., He, X., Bai, L., Tan, S., Tang, X., Liu, D., Liu, J., Zhang, S., Chen, M., Zhang, L., Zhang, L., Zhang, Y., Liao, W., Zhang, Y., Wang, M., Lv, X., Wen, B., Liu, H., Luan, H., Zhang, Y., Yang, S., Wang, X., Xu, J., Li, X., Li, S., Wang, J., Palloix, A., Bosland, P.W., Li, Y., Krogh, A., Rivera-Bustamante, R.F., Herrera-Estrella, L., Yin, Y., Yu, J., Hu, K., Zhang, Z. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum

domestication and specialization. Proceedings of the National Academy of Sciences of the United States of America, 111 (14), pp. 5135-5140. DOI: 10.1073/pnas.1400975111.

Quirin, E.A., Ogundiwin, E.A., Prince, J.P., Mazourek, M., Briggs, M.O., Chlanda, T.S., Kim, K.T., Falise, M., Kang, B.C., Jahn, M.M. (2005). Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto.5.2, a major QTL for resistance to Phytophthora capsici in pepper. Theoretical and Applied Genetics, 110 (4), pp. 605-612.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. Marker Development and Gene Cloning for Pepper Disease Resistance. Curr. Opin. Plant Biol. 5: 94-100.

Rao, G.U., Chaim, A.B., Borovsky, E., Paran, I. (2003). Mapping of yield related QTLs in pepper in an interspecific cross of Capsicum annuum and C. frutescens. Theo. Appl. Genet. 106: 1457-1466.

Rast, A.T.B. (1988). Pepper tobamoviruses and pathotypes used in resistance breeding. Capsicum Newsl. 7, 20–23.

Ravishankar, G.A., Sarma, K.S., Venkataraman, L.V., Kadyan, A.K. (1988). Effect of nutritional stress on capsaicin production in immobilized cell cultures of Capsicum annuum. Curr. Sci. 57: 381-383.

Rajesh, R. W., Madhukar, S. W. (2018). Identification of sequence-characterized amplified regions (SCARs) markers linking resistance to powdery mildew in chilli pepper (Capsicum annuum L.). African J. Agric. Res. 13, 2771–2779.

Reif, J.C., Melchinger, A.E., Xia, X.C., Warburton, M.L., Hoisington, D.A., Vasal, S.K., Srinivasan, G., Bohn, M. and Frisch, M. (2003). Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. Crop Science 43, 1275–1282.

Riley, D.G., Joseph, S.V., Srinivasan, R., Diffie, S. (2011). Thrips vectors of Tospoviruses. J. Integr. Pest Manag. 2, 1–10.

Rincent, R., Malosetti, M., Ababaei, B., Touzy, G., Mini, A., Bogard, M., *et al*. (2019). Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. Theor. Appl. Genet. 132, 3399–3411. doi: 10.1007/s00122-019-03432-y.

Robert, P., Le Gouis, J., and Rincent, R. (2020). Combining crop growth modeling with trait-assisted prediction improved the prediction of genotype by environment interactions. Front. Plant Sci. 11, 1–11. doi: 10.3389/fpls.2020.00827.

Römer, P., Jordan, T., Lahaye. T. (2010). Identification and application of a DNA-based marker that is diagnostic for the pepper (Capsicum annuum) bacterial spot resistance gene Bs3. Plant Breeding, 129 (6), pp. 737-740.

Ros, C., Lacasa, C.M., Martínez, V., Bielza, P., Lacasa, A. (2014). Response of pepper rootstocks to co-infection of Meloidogyne incognita and Phytophthora spp. Eur. J. Hortic. Sci. 2014, 79, 22–28.

Rostoks N., Ramsay L., MacKenzie K., Cardle L., Bhat P.R., Roose M.L., Svensson J.T., Stein N., Varshney R.K., Marshall D.F., Graner A., Close T.J. (2006). Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. PNAS, volume 103, n 49, December 5, 2006, 103 (49) 18656-18661.

Roy, D., Kharkwal, M.C. (2004). Breeding for Wider Adaptability. In: Jain, H.K., Kharkwal, M.C. Plant Breeding. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-1040-5_24.

Rubio, M., Caranta, C., Palloix, A. (2008). Functional markers for selection of potyvirus resistance alleles at the pvr2-eIF4E locus in pepper using tetra-primer ARMS-PCR. Genome 51, 767–777.

Ruffel, S., Dussault, M., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C., Caranta, C. (2002). A natural recessive resistance gene against potato virus Yin pepper corresponds to the eukaryotic initiation factor4E (eIF4E). The Plant Journal 32, 1067–1075.

Ruffel, S., Gallois J.L., Moury, B., Robaglia, C., Palloix, A., Caranta, C. (2006). Simultaneous mutations in translation initiation factors eIF4E and eIF(iso)4E are required to prevent pepper veinal mottle virus infection of pepper. J. Gen. Virol. 87:2089-2098.

Ruiz H., Pavon J. (1790). "Flora Peruviana, et Chilensis", 2: 30-31" cited in Heiser e Smith (1958).

Rutkoski, J.; Singh, R.P.; Huerta-Espino, J.; Bhavani, S.; Poland, J.; Jannink, J.L.; Sorrells, M.E. (2015). Genetic Gain from Phenotypic and Genomic Selection for Quantitative Resistance to Stem Rust of Wheat. Plant Genome 2015, 8, 2. [CrossRef] [PubMed].

Sant, V.J., Patankar, A.G., Sarode, N.D., Mhase, L.B., Sainani, M.N., Deshmukh, R.B., Ranjekar, P.K. and Gupta, V.S. (1999). Potential of DNA markers in detecting divergence and in analyzing heterosis in Indian elite chickpea cultivars. Theoretical and Applied Genetics 98, 1217–1225.

Sahin, F., Miller, S.A. (1998). Resistance in Capsicum pubescens to Xanthomonas campestris pv. vesicatoria pepper race 6. Plant Dis. 1998, 82, 794–799.

Sarath Babu, B., Pandravada, S.R., Pasada Rao, R.D.V.J., Anitha, K., Chakrabarty, S.K., Varaprasad, K.S. (2011). Global sources of pepper genetic resources against arthropods, nematodes and pathogens. Crop Prot. 2011, 30, 389–400.

Scopus database, https://www.elsevier.com.

Sehgal, D. (2016). Advances in Molecular Breeding of Pearl Millet. 397–419. doi:10.1007/978-3-319-27090-6_15.

Semagn K, Bjørnstad A, N. (2006). An overview of molecular marker methods for plants. Afr. J. Biotechnol. 19-35.

Servin, B., Martin, O.C., Mézard, M. and Hospital, F. (2004). Toward a theory of marker-assisted gene pyramiding. Genetics 168, 513–523.

Shanti, M. L., George, M. L. C., Cruz, C. M. V., Bernardo, M. A., Nelson, R. J., Leung, H., Reddy, J. N., Sridhar, R. 570 B. C. Y. Collard, D. J. Mackill (2001). Marker-assisted selection in plant breeding. Identification of resistance genes effective against rice bacterial blight pathogen in eastern India. Plant Dis. 85, 506–512. (doi:10.1094/PDIS.2001.85.5.506).

Simonne, A.H., Simonne, E.H., Eitenmiller, R.R., Mills, H.A., Green, N.R. (1997). Ascorbic acid and provitamin. A contents in unusually colored bell peppers (Capsicum annuum L.). J. Food Compo. Anal. 10: 299–311.

Singh, S., Sidhu, J. S., Huang, N., Vikal, Y., Li, Z., Brar, D. S., Dhaliwal, H. S., Khush, G. S. (2001). Pyramiding three bacterial blight resistance genes (xa5, xa13 and Xa21) using marker-assisted selection into indica rice cultivar PR106. Theor. Appl. Genet. 102, 1011–1015. (doi:10.1007/ s001220000495).

Smith, H.F. (1936). A discriminant function for plant selection. Annals of Eugenics 7, 240–250.

Smith, M.E., Coffman, W.R. and Barker, T.C. (1990). Environmental effects on selection under high and low input conditions. In: Kang, M.S. (ed.) Genotype-By-Environment Interactions and Plant Breeding. Louisiana State University Agriculture Center, Baton Rouge, Louisiana, pp. 261–272.

Smith P.G., Villalon B., Villa P.L. (1987). "Horticultural classification of pepper grown in the United States." HortScience, 22: 11-13.

Sol Genomics Network, https://solgenomics.net.

Soler, S., Debreczeni, D.E., Vidal, E., Aramburu, J., López, C., Galipienso, L., Rubio, L. (2015). A new Capsicum baccatum accession shows tolerance to wild-type and resistance-breaking isolates of Tomato spotted wilt virus. Ann. Appl. Biol. 167, 343–353.

Sprague, G.F. and Tatum, L.A. (1942). General vs. specific combining ability in single crosses of corn. Journal of American Society of Agronomy 34, 923–932.

Stam, P. (1995). Marker-assisted breeding. In: Van Ooijen, J.W. and Jansen, J. (eds) Biometrics in Plant Breeding: Applications of Molecular Markers. Proceedings of the 9th Meeting of EUCARPIA Section on Biometrics in Plant Breeding (1994). Centre for Plant Breeding and Reproduction Research, Wageningen, Netherlands, pp. 32–44.

Stoskopf, N. C., Tomes, D. T., Christie, B. R. (1993). Plant breeding: theory and practice. San Francisco, CA; Oxford: Westview Press Inc.

Stoytcheva, M. (2011). Pesticides in the Modern World: Risks and Benefits.

Stuber, C.W. (1999). Biochemistry, molecular biology and physiology of heterosis. In: Coors, J.G. and Pandey, S. (eds) The Genetics and Exploitation of Heterosis in Crops. American Society of Agronomy (ASA) and Crop Science Society of America (CSSA), Madison, Wisconsin, pp. 173–184.

Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T. and Lander, E.S. (1992). Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132, 823–839.

Sugita, T., Yamaguchi, K., Sugimura, Y., Nagata, R., Yuji, K., Kinoshita, T., Todoroki, A. (2004). Development of SCAR markers linked to L3 gene in Capsicum. Breeding Science, 54 (2), pp. 111-115.

Sugita, T., Yamaguchi, K., Kinoshita, T., Yuji, K., Sugimura, Y., Nagata, R., Kawasaki, S., Todoroki, A. (2006). QTL analysis for resistance to Phytophthora blight (Phytophthora capsici Leon.) using an intraspecific doubled-haploid population of Capsicum annuum. Breed. Sci. 56, 137–145.

Sushil, C., Mechuselie, K., Ramchiary, N. (2016). Advances in Molecular Breeding of Capsicum Species. Biotechnological Tools for Genetic Resources Pages 233–274.

Suzuki, K., Kuroda, T., Miura, Y., Murai, J. (2003). Screening and field trials of virus resistant sources in Capsicum spp. Plant Dis. 87, 779–783.

Tamisier, L., Szadkowski, M., Nemouchi, G., Lefebvre, V., Szadkowski, E., Duboscq, R., Santoni, S., Sarah, G., Sauvage, C., Palloix, A. (2020). Genome-wide association mapping of QTLs implied in potato virus Y population sizes in pepper: Evidence for widespread resistance QTL pyramiding. Mol. Plant Pathol. 21,3–16.

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP. (1988). Conservation of gene repertoire but not gene order in pepper and tomato. Proc. Natl. Acad. Sci. USA. 85: 6419-6423.

Tester, M., Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. Science, 2010, 327, 818-822.

Thabuis, A., Palloix, A., Pflieger, S., Daubeze, A.M., Caranta, C., Lefebvre, V. (2003). Comparative mapping of Phytophthora resistance loci in pepper germplasm: Evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. Theor. Appl. Genet. 106, 1473–1485.

Thomson, M.J. (2014). High-throughput SNP genotyping to accelerate crop improvement. Plant Breed. Biotech. 2:195-212.

Tripodi, P., Kumar, S. (2019). The Capsicum Crop: An Introduction. 1–8. doi:10.1007/978-3-319-97217-6_1.

Truong, H.T.H., Kim, K., Kim, S., Cho, M., Kim, H., Woo, J. (2011). Development of Gene-based Markers for the Bs2 Bacterial Spot Resistance Gene for Marker-assisted Selection in Pepper (Capsicum spp.). Hort. Environ. Biotechnol. 52(1):65-73. 2011.DOI 10.1007/s13580-011-0142-4.

Truong, H.T.H., Kim, K.T., Kim, D.W., Kim, S., Chae, Y., Park, J.H., Oh, D.G., Cho, M.C. (2012). Identification of isolate-specific resistance QTLs to Phytophthora root rot using an intraspecific recombinant inbred line population of pepper (Capsicum annuum). Plant Pathol. 61, 48–56.

Truong, H.T., Ramos, A.M., Yalcin, F., De Ruiter, M., Van der Poel, H.J., Huvenaars, K.H.J, Hogers, R.C.J, Van Enckevort, L.J.C., Janssen, A., Van Orsouw, N.J., Van Eijk, M.J.T. (2012). Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. PLoS One. 2012;7(5):e37565. doi: 10.1371/journal.pone.0037565. Epub 2012 May 25.

Truong, H.T.H., Kim, J.H., Cho, M.C., Chae, S.Y., Lee, H.E. (2013). Identification and development of molecular markers linked to Phytophthora root rot resistance in pepper (Capsicum annuum L.). European Journal of Plant Pathology, 135 (2), pp. 289-297.

Turina, M., Kormelink, R., Resende, R.O. (2016). Resistance to Tospoviruses in Vegetable Crops: Epidemiological and molecular aspects. Annu. Rev. Phytopathol. 54, 347–371.

Uncu, A.T., Celik, I., Devran, Z., Ozkaynak, E., Frary, A., Frary, A., Doganlar, S. (2015). Development of a SNP-based CAPS assay for the Me1 gene conferring resistance to root knot nematode in pepper. Euphytica, 206 (2), pp. 393-399.

Vallejos, C.E., Jones, V., Stall, R.E., Jones, J.B., Minsavage, G.V., Schultz, D.C., Rodrigues, R., Olsen, L.E., Mazourek, M. (2010). Characterization of two recessive genes controlling resistance to all races of bacterial spot in peppers. Theor. Appl. Genet. 121, 37–46.

Van Berloo, R. and Stam, P. (2001). Simultaneous marker-assisted selection for multiple traits in autogamous crops. Theoretical and Applied Genetics 102, 1107–1112.

Van Dun, C.M.P., Dirks, R.H.G. (2006). Rijk Zwaan Zaadteelten Zaadhandel BV Near Reverse Breeding, 2006. WO⁄2006⁄094773.

Varshney, R.K., Nayak, S.N., May, G.D., Jackson, S.A. (2009). Next generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol. 27: 522-530.

Venkatesh, J., An, J., Kang, W.H., Jahn, M., Kang, B.C. (2018). Fine Mapping of the Dominant Potyvirus Resistance Gene Pvr7 Reveals a Relationship with Pvr4 in Capsicum annuum. Phytopathology, 108(1), 142–148. doi:10.1094/phyto-07-17-0231-r .

Visscher, P. M., Haley, C. S., Thompson, R. (1996). Markerassisted introgression in backcross breeding programs. Genetics 144, 1923–1932.

Vos P., Hogers R., Bleeker M. (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. Nucleic Acids Res. 23, 4407–4414.

Wai, K.P., Siddique, M.I., Mo, H.S., Yoo, H.J., Byeon, S.E., Jegal, Y., Mekuriaw, A.A., Kim, B.S. (2015). Pathotypes of bacterial spot pathogen infecting Capsicum peppers in Korea. Plant Pathol. J. 31, 428–432.

Wang, K., Li, M., Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 38(16):e164.

Wang, J., Chapman, S.C., Bonnett, D.G., Rebetzke, G.J., Crouch, J. (2007). Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. Crop Science 47, 582–588.

Wang, P., Wang, L., Guo, J., Yang, W., Shen, H. (2016). Molecular mapping of a gene conferring resistance to Phytophthora capsici Leonian race 2 in pepper line PI201234 (Capsicum annuum L.). Molecular Breeding, 36 (6), art. no. 66.

Wang, X., Fazari, A., Cao, Y., Zhang, Z., Palloix, A., Mao, S., Wang, L. (2018). Fine mapping of the root-knot nematode resistance gene Me1 in pepper (Capsicum annuum L.) and development of markers tightly linked to Me1. Molecular Breeding, 38(4). doi:10.1007/s11032-018-0793-2.

Weiland J. (2003). Cleaved amplified polymorphic sequence (CAPS) marker associated with root-knot nem- atode resistance in sugarbeet.

Wiesner-Hanks, T., Nelson, R. (2016). Multiple disease resistance in plants. Annu. Rev. Phytophathol. 54: 229-252.

Xu, X., Chao, J., Cheng, X., Wang, R., Sun, B., Wang, H., Luo, S., Xu, X., Wu, T., Li, Y. (2016). Mapping of a novel race specific resistance gene to phytophthora root rot of pepper (Capsicum Annuum) using bulked segregant analysis combined with specific length amplified fragment sequencing strategy. PLoS ONE, 11 (3), art. no. e0151401.

Xu, Y., McCouch, S.R., Shen, Z. (1998). Transgressive segregation of tiller angle in rice caused by complementary action of genes. Crop Science 38, 12–19.

Xu, Y., Lobos, K.B. and Clare, K.M. (2002). Development of SSR markers for rice molecular breeding. In: Proceedings of Twenty-Ninth Rice Technical Working Group Meeting, 24–27 February 2002, Little Rock, Arkansas. Rice Technical Working Group, Little Rock, Arkansas, p. 49.

Xu, Yunbi. (2010). Molecular plant breeding / Yunbi Xu. p. cm. ISBN 978-1-84593-392-0 (alk. paper). ISBN: 978 1 84593 392 0 Typeset by SPi, Pondicherry, India.

Yabe, S., Iwata, H., Jannink, J. L. (2017). A simple package to script and simulate breeding schemes: the breeding scheme language. Crop Sci. 57, 1347–1354. doi: 10.2135/cropsci2016.06.0538.

Yang, H.B., Liu, W., Kang, W.H., Kim, J.H., Cho, H. J., Yoo, J.H., Kang, B.C. (2011). Development and validation of L allele-specific markers in Capsicum. Molecular Breeding, 30(2), 819–829. doi:10.1007/s11032-011-9666-7

Yang, H.B., Liu, W.Y., Kang, W.H., Jahn, M., Kang, B.C. (2009). Development of SNP markers linked to the L locus in Capsicum spp. by a comparative genetic analysis. Molecular Breeding, 24 (4), pp. 433-446.

Yang, H.B., Liu, W.Y., Kang, W.H., Kim, J.H., Cho, H., Yoo, J.H. (2012). Development and validation of L allele-specific markers in Capsicum. Mol. Breed. 30, 819–829.

Yang, L., Fu, S., Khan, A. (2013). Molecular cloning and development of RAPD-SCAR markers for Dimocarpus longan variety authentication. Springer Plus (2013).

Yeam, I., Kang, B., Lindeman, W., Frantz, J. D., Faber, N., Jahn, M. M. (2005). Allele-specific CAPS markers based on point mutations in resistance alleles at the pvr1 locus encoding eIF4E in capsicum. Theoretical and Applied Genetics, 112(1), 178-186. doi:10.1007/s00122-005-0120-2.

Zhang, W. ; Fang, Y., Shen, H. (2012). Mapping of a southern root-knot nematode resistance gene in pepper and marker-assisted selection. Journal of China Agricultural University Vol.17 No.2 pp.102-107 ref.14.

ACKNOWLEDGEMENTS

Firstly I would like to express my sincere gratitude to Prof. Francia for the opportunity to accomplish this doctoral project and for leading me in this pathway.

I would also offer a special thank to Dr. Justyna Anna Milc for your valuable advice, continuous flanking and patience during my PhD study.

A deeply grateful goes to whole ESASEM company, foremost to the President Mr. Gianni Gatto for having encouraged me and for believing without any hesitation in this formative step. I thank him for his foresight.

Finally, during these complicated years, I want to extend my total gratitude to my family especially my wife Cristina and my children Vinicio and Ottavio for unwavering support me in this decision to undertake a further academic and professional growth aware of the commitment required.

Thank you!