Review

Improving Pathogen Resistance by Exploiting Plant Susceptibility Genes in Coffee (Coffeea spp.)

Lei Cui 1,2,†, Katharina Hanika 2,†, Richard G. F. Visser 2,* and Yuling Bai 2,*

1 College of Agriculture, Shanxi Agricultural University, Taiyuan 030031, China; lei.cui@wur.nl
2 Plant Breeding, Wageningen University & Research, P.O. Box 386, 6700 AJ Wageningen, The Netherlands; katharina.hanika@wur.nl
* Correspondence: richard.visser@wur.nl (R.G.F.V.); bai.yuling@wur.nl (Y.B.)
† These authors contributed equally to this work.

Received: 7 November 2020; Accepted: 3 December 2020; Published: 8 December 2020

Abstract: Coffee (Coffeea spp.) is an economically important crop widely cultivated in (sub)tropical countries worldwide. Commercial coffee production relies mainly on two related species, namely C. arabica and C. canephora. Due to their perennial growth habit, cultivation practices, and narrow genetic diversity, coffees are constantly exposed to many diseases and pests. Coffee leaf rust (Hemileia vastatrix Berk. et Br.), coffee berry disease (Colletotrichum kahawae Bridge and Waller), and coffee wilt disease (Gibberella xylarioides Heim and Saccas/Fusarium xylarioides) are the top fungal diseases affecting C. arabica and C. canephora production areas worldwide. In many regions, chemical-based control measures are widely used and are the only way to control the diseases. Developing resistant cultivars is one of the prerequisites for increasing sustainable market demand and agriculture. However, desired and required resistance traits are not always available in the gene pool. Furthermore, from other crops it is clear that dominant resistance genes introduced into varieties are not durable because of pathogen variability and the emergence of new races of the different pathogens. Utilization of altered susceptibility genes (S genes) offers a novel and alternative strategy for the breeding of durable and broad-spectrum resistance. The S gene encodes a host factor that facilitates a compatible interaction with the pathogen, and impairment of S genes leads to loss-of-susceptibility. In this review, guidelines for effective identification, characterization, and utilization of dysfunctional S genes are proposed to aid breeding activities in order to introduce durable resistance in Coffeea spp.

Several candidate S genes likely contributing to the susceptibility of Colletotrichum spp., Fusarium spp., and Meloidogyne spp. are discussed. With the rapid development of genetic engineering techniques, including CRISPR-associated systems, we now have the potential to accelerate the application of S genes to achieve durable resistance in coffee.

Keywords: Coffeea arabica; C. canephora; disease resistance; susceptibility genes; plant breeding

1. Introduction

Coffee (Coffeea spp.), after palm oil, is the second-largest worldwide agricultural traded commodity [1]. It is grown in more than 80 countries spanning over 10.6 million hectares in the (sub)tropical regions, especially in Latin America, Asia, and Africa [2–4]. The top five countries of coffee production are Brazil, Vietnam, Colombia, Indonesia, and Ethiopia, which together produce more than 65% of the global total [3,5]. The Coffeea genus includes 103 species [1,6]. Commercial coffee production relies mainly on two species, C. arabica and C. canephora, which account for 65% and 35% of world coffee production, respectively [1]. C. arabica coffees are characterized by low bitterness, aromatic properties and low caffeine content, which are preferred by consumers. On the other hand, C. canephora coffees have a stronger bitterness and a higher caffeine content [1]. The costs of C. arabica
coffee production are often much higher than those of the C. canephora coffee, mainly due to more stringent demands of crop cultivation and soil conditions as well as management of multiple pests and diseases [7].

To date, C. arabica coffees have a rather narrow genetic diversity and worldwide production is based on traditional old cultivars or their derived offspring [7,8]. Traditional cultivars all have their strengths and weaknesses. Some are able to produce beans of excellent cup quality, but with relatively low yield, while some keep a high yielding potential but with more variable cup quality. Nevertheless, all varieties are very susceptible to various coffee diseases and pests, including the most destructive and commonly presented diseases, coffee leaf rust (CLR), coffee berry disease (CBD), and coffee wilt disease (CWD) [9].

There are several ways to control diseases and pests in coffee production, ranging from chemical and biological control, to good farm management practices. However, the interest of this review lies in the use of host genetic resistance with a focus on novel breeding strategies that are based on impaired susceptibility genes (S genes).

2. Current Status of C. arabica and C. canephora Genomes

Despite the economic importance of Coffea species, relatively little genomic information has been made available over the years. It was not until very recently that high-quality genome sequences and genetic maps of C. canephora and C. arabica have become available [10,11]. Allo-tetraploid C. arabica (2n = 4 x = 44 chromosomes), derived from diploid progenitor species C. canephora and C. eugenioides, has a relatively large genome size of 1.2-Gb containing 70,830 predicted genes. At the whole genome level, genetic variation was analyzed among different varieties of C. arabica, providing more insight into coffee quality traits and adaptation to diverse environments [11]. The genome of diploid C. canephora (2n = 2 x = 22 chromosomes, outcrossing and highly heterozygous) is 710-Mb in size with 25,574 annotated protein-coding genes [10]. A clear feature associated with this genome is the expansion of specific gene families including N-methyltransferases (involved in caffeine production), nucleotide binding site (NBS) defense-related genes, and alkaloid and flavonoid enzymes [10]. The NBS type of resistance (R) gene family accounts for approximately 2.2% of the annotated genes in the C. canephora genome. This proportion is higher than that of many other sequenced angiosperm genomes [10].

3. Common Diseases in C. arabica and C. canephora

Being a perennial crop, the coffee plant hosts a wide variety of pathogens and pests. Coffee cultivation is confined to the (sub)tropical countries with warm temperature and high relative humidity. Additionally, coffee plants are grown at high plant density and often have unbalanced nutrition. These factors favor the incidence and spread of many diseases [12,13]. Some coffee diseases cause serious constraints on coffee yield and deterioration of quality in the trade market [12]. The common diseases in coffee production include bacteria, fungi, nematodes, and viruses as well as other pests (Table 1) [9,12–32]. Among them, CLR, CBD, and CWD are the top three economically important diseases [12,33].

CLR, caused by Hemileia vastatrix Berk. et Br., can be found in nearly all regions where C. arabica and C. canephora coffees are grown and is considered the main coffee disease worldwide [13]. As the name suggests, this fungal disease occurs only on leaves (young and old) of all developmental stages. Unlike most rusts entering the host through the epidermis, this pathogenic fungus penetrates host plants through the stomata on the underside of the leaves, forming a penetration hypha. This hypha produces two lateral branches and each differentiates into a haustorial mother cell that gives rise to a haustorium, which primarily infects the stomatal subsidiary cells. This is a unique morphological feature of H. vastatrix [34,35]. Early in the infection cycle, patches of orange colored pustules form on the adaxial side of the leaves, and as these spots gradually expand in diameter, masses of urediniospores appear on the abaxial surfaces. The infected leaves drop prematurely. This disease is able to reduce coffee yield by up to 35 % [13,33].
Table 1. List of pathogens and pests commonly present in *Coffee arabica* and/or *C. canephora*.

<table>
<thead>
<tr>
<th>Class</th>
<th>Disease Common Name</th>
<th>Pathogen</th>
<th>Coffea Host</th>
<th>Country/Region of Occurrence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterium</td>
<td>Bacterial blight</td>
<td><em>Pseudomonas syringae</em></td>
<td><em>C. arabica</em></td>
<td>Brazil</td>
<td>[14,15]</td>
</tr>
<tr>
<td></td>
<td>Bacterial leaf scorch</td>
<td><em>Xylella fastidiosa</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Brazil</td>
<td>[16]</td>
</tr>
<tr>
<td>Fungus</td>
<td>American leaf spot</td>
<td><em>Mycena citricolor</em></td>
<td><em>C. arabica</em></td>
<td>Central America, Colombia</td>
<td>[9,17]</td>
</tr>
<tr>
<td></td>
<td>Black Rot</td>
<td><em>Corticium koleroga</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Ethiopia, India, Mexico</td>
<td>[9,18,19]</td>
</tr>
<tr>
<td></td>
<td>Coffee berry disease</td>
<td><em>Colletotrichum kahawae</em></td>
<td><em>C. arabica</em></td>
<td>Africa</td>
<td>[20]</td>
</tr>
<tr>
<td></td>
<td>Coffee leaf rust</td>
<td><em>Hemileia vastatrix</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Worldwide</td>
<td>[13,20]</td>
</tr>
<tr>
<td></td>
<td>Coffee wilt disease</td>
<td><em>Gibberella xylarioides</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Central and East African countries</td>
<td>[12,21]</td>
</tr>
<tr>
<td>Nematode</td>
<td>Burrowing nematode</td>
<td><em>Radopholus similis</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Vietnam</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td>Reniform nematode</td>
<td><em>Rotylenchulus reniformis</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Vietnam</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td>Root-knot nematodes</td>
<td><em>Meloidogyne spp.</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>America, Africa, India</td>
<td>[9,24]</td>
</tr>
<tr>
<td></td>
<td>Root-lesion nematodes</td>
<td><em>Pratylenchus spp.</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>America, Africa, India, Southeast Asia, Turkey</td>
<td>[9,25]</td>
</tr>
<tr>
<td>Insect</td>
<td>Antestia bug</td>
<td><em>Antestiopsis spp.</em></td>
<td><em>C. arabica</em></td>
<td>Africa</td>
<td>[26]</td>
</tr>
<tr>
<td></td>
<td>Black twig borer</td>
<td><em>Xylosandrus compactus</em></td>
<td><em>C. arabica</em></td>
<td>Vietnam, Indonesia, India, Africa, Brazil</td>
<td>[27]</td>
</tr>
<tr>
<td></td>
<td>Coffee berry borer</td>
<td><em>Hypothenemus hampei</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Worldwide</td>
<td>[19,28]</td>
</tr>
<tr>
<td></td>
<td>Leaf miner</td>
<td><em>Leucoptera coffeella</em></td>
<td><em>C. arabica, C. canephora</em></td>
<td>Africa, South and Central America Asia</td>
<td>[29,30]</td>
</tr>
<tr>
<td></td>
<td>White stem borer</td>
<td><em>Xylotrechus quadripes</em></td>
<td><em>C. arabica</em></td>
<td>Asia</td>
<td>[31]</td>
</tr>
<tr>
<td>Virus</td>
<td>Coffee ringspot</td>
<td><em>Coffee ringspot virus</em></td>
<td><em>C. arabica</em></td>
<td>Brazil, Costa Rica</td>
<td>[32]</td>
</tr>
</tbody>
</table>
CBD, caused by Colletotrichum kahawae Bridge and Waller, is an anthracnose and a major constraint to C. arabica coffee production. It is restricted to C. arabica in Africa. It is ranked as a quarantine pathogen in the main C. arabica growing countries including America and Asia [36,37]. In the genus Colletotrichum, C. kahawae is highly aggressive and specialized in infecting green coffee berries to cause CBD. Other Colletotrichum spp. are able to colonize ripe coffee berries, but they are incapable of causing diseases. The fungal spores germinate and enter the host via a penetration peg produced beneath an appressorium by piercing the cuticle and wall of the epidermal cells [38]. During berry development, infection takes place, producing black sunken anthracnose lesions on the green berry. Coffee berry losses of up to 80% were reported due to anthracnose outbreak [20,33,36].

CWD, caused by Gibberella xylarioides Heim and Saccas/Fusarium xylarioides, is also referred to as Fusarium wilt. Currently, CWD has a devastating effect on coffee production and has reached epidemic proportions in eastern and central African countries. It occurs not only in C. canephora but also in C. arabica, however, so far only restricted to Ethiopia [12,21]. It is a vascular fungal disease that affects different organs (i.e., leaf, stem, and berry) of the trees at all stages of growth [21]. The fungus enters the host through the soil or vascular wounds [39]. The typical symptom is yellowing and wilting of the infected seedlings and trees, as well as discoloration on the stems. CWD is able to rapidly kill infected mature trees, resulting in a severe yield loss of over 50%. Berries on infected trees turn red prematurely, which exerts a negative effect on coffee quality [12,21].

Although these diseases can be partially managed by integrated approaches, such as biological control, use of plant extracts and mineral nutrition, chemical-based control is still widely used and is considered as an effective measure [13]. Application of chemical measures accounts for 30%–40% of field managing cost, which makes coffee growing, in particular for smallholder farmers, unprofitable [35,40]. In many breeding programs worldwide, development of disease resistant cultivars is the major objective, which reduces the costs of using pesticides and is both environmentally and human friendly [40].

4. Breeding for Disease Resistance

Breeding resistant cultivars is the most effective control measure for major coffee diseases and has been explored in the past mainly by conventional breeding methods aimed at introgressing resistances identified in C. arabica, C. canephora, and C. liberica into elite cultivars via backcrossing [40]. Breeding programs have the main objective of combining optimum yield potential with high beverage quality and wide climatic adaptation [20].

Host resistance to CBD has been identified in the C. arabica cultivars which are governed by three genes; two are dominant and the third one is recessive [41]. Up until now, the genetics of CBD resistance are not yet fully understood. Multiple CBD resistant cultivars have been released in the 1980s. The resistance in these cultivars appears to be durable, since no breakdown of resistance has been reported over the past 40 years [7,37]. These varieties include “Rume Sudan” harboring the dominantly inherited R gene and Híbrido De Timor, as well as its derived progenies Caturra and Catimor containing the dominant T gene. In C. arabica cv. Ruiru 11 and MGS Catiguá 3, a major resistance gene was designated as Ck-1 (synonymous to the T gene). SSR (simple sequence repeats) marker CBD-Sat235 co-segregates with the Ck-1 gene and CBD-Sat207 is located at an estimated distance of 17.2-cM. [40,42,43]. Molecular markers associated with Ck-1 are located on chromosome 1, which have been used in breeding programs to speed up the introgression efficiency (Figure 1) [44–47]. Three quantitative trait loci (QTL), qCBD 1-1 on chromosome 1, qCBD 2-1, and qCBD 2-2 on chromosome 2 were identified as responsible for CBD resistance in the C. arabica cultivar Rume Sudan [48]. SNP (single nucleotide polymorphism) markers for these genomic loci are located on coffee Chromosomes 1 (SNP_100025973) and 2 (SNP_100034991) (Figure 1) [41,42]. These markers can be applied for marker assisted selection in C. arabica coffee breeding programs. The recessively inherited resistance was conditioned by the k gene which was identified in C. arabica cultivars Rume Sudan and K7 [49]. The genetic basis of k-gene mediated CBD resistance and its associated molecular markers has not yet been exploited. Conventional breeding heavily relies on backcross approaches aiming to
restore the desirable yield and quality attributes in the original superior genotype simultaneously in combination with resistance. This selection process led to the loss of the k-gene, due to its recessive nature [41].

The Hibrido de Timor (HDT) populations, natural hybrids between C. arabica and C. canephora, provided effective sources of resistance against CLR. These HDT hybrids have served as the basis for breeding programs worldwide, resulting in numerous CLR resistant cultivars [35]. In the HDT populations, nine genetic loci (SH1-SH9) have been detected contributing to the CLR resistance. Among them, SH1, SH2, SH4 and SH5 originated from C. arabica. Other loci, except SH3 which was identified in C. liberica, were found in C. canephora. The importance of HDT populations as resistance sources relies on their durability, as exemplified by the long-lasting effectiveness of some of these resistance factors in the field which have been in use for more than 30 years [35]. The SH3 gene has shown to provide durable protection under field conditions. The markers Sat244, SP-M16- SH3, and BA-48-21OR co-segregated with the SH3 gene and are positioned on C. canephora chromosome 3 (Figure 1) [43,45]. In contrast, the host resistance to CLR acquired by introgression of the other SH genes, either monogenic independently or polygenic jointly, has shown to be non-durable in multiple countries, particularly in areas where climatic conditions are favorable to the pathogen. Due to pathogen variability and the emergence of new races, CLR resistance has been reported to be broken and the list of countries with CLR outbreaks has expanded rapidly in recent years including Brazil, Colombia, China, India, and Thailand [7,13]. Therefore, there is an urgent need to search in the Coffea gene pool in order to identify novel resistant sources against CLR. However, such an effort is partially restricted due to the fact that C. arabica coffee with its autogamous nature is known to have a very narrow genetic base [8,33]. Partial CLR resistance identified from C. arabica line DI.200 was able to reduce rust incidence and defoliation. Two SSR markers (Sat225 and Sat229) located on chromosomes 11 and 5, respectively, showed significant association with this partial resistance (Figure 1) [46]. Using the same C. arabica line, one major QTL designed as QCLR_4 was identified for disease severity index under natural field conditions. This QTL was positioned within the interval between markers SSR_SFCan022c and SSR_05RM1 on chromosome 11 (Figure 1) [44]. Molecular markers at 8.69-cM and 25.10-cM were identified as being linked to a dominant resistance gene responsible for the resistant response of C. arabica genotype HDT UFV 427-15 [50]. Various other sources of resistance to the disease have been identified and are currently used in breeding resistant cultivars [7].

Host resistance to CWD is achieved through resistant cultivars of C. canephora. However, resistance has been broken due to emergence of new races [12]. Durable resistance seems to be present in Ethiopian germplasm of C. arabica and breeding efforts in Ethiopia have resulted in the development of resistant C. arabica cultivars, such as Fayate [7].
Figure 1. Physical positions at the Megabase (Mb) scale of previously described resistant genes, quantitative trait loci (QTLs), resistance-linked molecular markers, and putative susceptibility genes to coffee berry disease, coffee wilt disease, coffee leaf rust, and root-knot nematodes on the chromosomes of Coffea canephora according to the C. canephora genome database, The Coffee Genome Hub. QTL for resistance are displayed as bars and the names given by the authors are maintained [44]. Molecular markers are depicted as associated disease and original names given by the authors [41,43,45–49]. MCBD, MCLR, and MRKN denote molecular markers (M) associated with resistance to coffee berry disease, coffee leaf rust, and root-knot nematodes. Susceptibility candidate genes are indicated as red solid lines. QTL and markers derived from C. canephora are shown in green. The blue highlights indicate that the resistances are from C. liberica. Resistances originating from C. arabica are designated in pink color.
5. Breeding for Durable and Broad-Spectrum Resistance—Loss of Susceptibility as a Novel Breeding Strategy

A novel strategy to achieve durable (and broad-spectrum) resistance involves loss-of-function of S genes in the host plant [51]. S genes encode host proteins that are exploited by the pathogen to establish a compatible interaction with the host, and thus facilitate infection [51–55]. While, from the pathogen’s perspective, this plant gene aids the establishment of disease, for the host this gene presents a weak link in its defense. Therefore, mutation or loss of function of S genes can limit the ability of the pathogen to cause disease. Besides the intrinsic function S genes have for the host, these genes also have a plethora of functions in plant–pathogen interaction that can be categorized into three classes based on the phase of their action [53]. The first class comprises gene products that are involved in the early interactions between host and pathogen to provide early pathogenicity establishment. An example is the *Medicago truncatula* mutant *irg1* (*Inhibitor of Rust Germ Tube Differentiation1*) in which germ tube growth of *Puccinia enamulata* is inhibited due to the loss of epicuticular waxes [56]. The second class of S genes concerns those genes that are involved in negative regulation of immunity. This class contains mutants with a constitutive expression of the defense hormone salicylic acid (SA), such as the constitutive expresser of PR 5 (cpr5) mutant in *Arabidopsis*, which is resistant to *Pseudomonas syringae* and *Peronospora parasitica* [57]. The third class of S genes encompasses genes that allow pathogen proliferation at late infection stages. One well-studied example is the efflux sugar transporter SWEET11 which underlies the recessive *xa13* resistance in rice against *Xanthomonas oryzae*. This sugar transporter transports sugars into the apoplastic space, thereby providing nutrients to the pathogen [58,59]. In general, it should be acknowledged that S genes have a wide variety of functions in the interaction with pathogens and act at different stages of the infection.

From the above-mentioned examples, it already becomes clear that S genes play a role in disease susceptibility to a wide range of pathogens, concerning biotrophs such as discussed in the cases above as well as necrotrophs. For example, downregulation or knock-out of the expansin-like A2 (*EXLA2*) gene that encodes a cell wall-modifying enzyme confers resistance against the necrotrophic fungi *Botrytis cinerea* as well as *Alternaria brassicola* [60]. Besides playing roles in interactions with fungi, bacteria and oomycetes, S genes also play roles in interactions with viruses that rely on host factors for their replication [61]. For example, simultaneous mutations in the two *Arabidopsis* genes *Tobamovirus Multiplication 1* and 3 (*TOM1* and *TOM3*), which encode transmembrane proteins required for tobamovirus replication, leads to undetectable viral titers [62]. Another essential and well-studied host factor required by different viruses is the eukaryotic translation initiation factor (eIFs), which provided resistance when mutated [63]. To a far smaller extent, S genes were also characterized in interactions with nematodes. A mutant in the *heavy metal associated isoprenylated plant protein 27* (*HIPP27*) provides resistance against the cyst nematode *Heterodera schachtii* [64]. Lastly, studying S genes in the interaction with insects has been proposed, for example for aphids in which omics studies may lay the foundation to identify insect-related S genes [65].

In order to utilize S genes in resistance breeding, their function in disease susceptibility needs to be impaired [51]. Thus, S genes can be characterized as dominant genes while the resistance that is based on their impairment inherits recessively. The concept of S gene-mediated resistance breeding is based on impairing an S gene from the plant and thereby adding resistance.

6. Your Guidelines—How to Find Putative Candidate S Genes in *Coffeea spp.*

In the past few years, ample body of evidence suggests that resistance can be achieved by altering plant S genes. The core component of using impaired S genes for any given plant–pathogen interaction is the identification of these genes in the first place. Generally, two main strategies can be pursued to identify S genes: forward and reverse genetics. In an attempt to shed light on how S genes are generally identified, 168 S genes mentioned in an extensive review by van Schie and Takken [53] were categorized by their method of identification through either forward or reverse genetics. This search
revealed that 60 S genes (35.7%) were identified via a forward genetics screening, while 108 genes (64.3%) were identified using reverse genetics.

A forward genetics approach entails the screening of a mutant population to identify plants that show loss of susceptibility to a given pathogen. Such assays are only feasible if a large number of mutants can be screened simultaneously and if the phenotyping allows rapid identification of resistant mutants. Hence, the majority of studies that used forward genetics to identify S genes were performed in Arabidopsis. Examples are the pmr mutants which were isolated from a screening of 26,000 Arabidopsis mutants with Erysiphe cichoracearum [66]. Other studies, especially in crops, made use of forward genetics by mapping naturally occurring resistance.

The identification of S genes via reverse genetics can follow three main strategies. Frequently used are expression analyses followed by functional characterization. Such analyses are based on the fact that pathogens induce transcriptomic changes in the host [67,68], and transcriptomic studies can cover a wide spectrum of research questions. For example, a study focused on transcriptional changes induced by fungal trichothecene toxins that revealed the involvement of a homologue of the putative human transcription repressor NF-X1 (AtNFXLI), which was further shown to be a negative regulator in defense to P. syringae in Arabidopsis [69,70]. The second reverse genetics strategy is based on the fact that some S genes are effector targets [54]. These targets can be identified with a yeast two-hybrid assay as demonstrated for the viral genome-linked protein (VPg) [71]. The host target of VPg is a DEAD-box RNA helicase from Arabidopsis (AIRH8), and AIRH8 mutants were found to be resistant to potyviruses. The third reverse genetics strategy focuses on the identification of orthologues of known S genes in other plant species. This approach relies on the key feature that the role of S genes was found to be conserved between plant species, providing the possibility to impair the S gene in different plant species. One of the best-known S genes is MLO (mildew resistance locus O), which is required for powdery mildew susceptibility [72]. mlo and its orthologs are highly conserved in presence and function of powdery mildew incompatibility in many plant species including Arabidopsis, tomato, pepper, tobacco, pea, wheat, strawberry, apple, grapevine, cucumber, melon, rose, and petunia [53,73–79]. Another example is the suppressor of salicylate insensitivity of npr1-5 (ssi2) mutant of Arabidopsis that showed enhanced resistance to P. parasitica, presently known as Hyaloperonospora arabidopsidis, due to an impairment in a stearoyl acyl carrier protein fatty acid desaturase (SACPD) [80]. In soybean, silencing of an SACPD orthologue resulted in reduced bacterial titers of P. syringae and reduced lesion size in response to P. sojae [81]. In addition, in rice, downregulation of an SACPD orthologue lowered the number and size of lesions caused by Magnaporthe grisea and X. oryzae [82].

In coffee, application of the forward genetics approach is difficult and not feasible at this moment and thus in the current review, the strategy focuses on the candidate gene approach to identify putative S genes by reverse genetic analysis. The first reverse genetics strategy lies on a model-to-crop translation by studying Coffea orthologues of known S genes from model species such as Arabidopsis. Numerous S genes have been identified in many model/plant species for a wide range of pathogens [53]. As inspired by the fact that S genes are conserved across plant species, mutated S gene orthologs can probably give rise to resistance in Coffea spp. Guidelines of such model-to-crop translations can be derived in order to introduce the potential S gene ortholog(s)-based resistance in Coffea spp. (Figure 2). Firstly, it involves the identification of functionally proven S genes known from other model/plant species according to the list of van Schie and Takken [53]. In Arabidopsis, tomato, maize, cucumber, and Medicago, several S genes were characterized as contributing susceptibility to anthracnose diseases and Fusarium wilt diseases caused by Colletotrichum spp. and Fusarium spp., respectively. Although this review focuses extensively on the economically important fungal diseases, it is noteworthy that three S genes known from Arabidopsis were proven to promote susceptibility to root-knot nematode disease caused by Meloidogyne spp. These pathogens constitute ubiquitous threats also to Coffea spp., causing CBD, CWD, and root-knot nematode disease [56,83–93]. Thus, these genes are considered as putative S genes candidates in Coffea spp. (Table 2).
How to find and use candidate susceptibility genes in *Coffea* spp.?

**Step 1 | Identification of the candidate S genes by literature mining**
- Numerous S genes contributing susceptibility according to the list of van Schie and Takken (2014)
- S genes are conserved across plant species

**Step 2 | Retrieval of S gene orthologs in *C. arabica* and *C. canephora***
Blast the functionally characterized S gene homolog to the available genomic database

Perform a phylogenetic analysis with the newly identified homologs

Select the phylogenetically closest S gene ortholog(s) with extra attention to:
1. The upregulated ortholog(s) upon pathogen challenge
2. Ortholog(s) co-localizing with (recessive) QTLs for resistance

**Step 3 | Functional validation of the candidate S genes**
Knocking down and/or knocking out of its expression through either a transient or stable way

Are deleterious pleiotropic effects visible?
- YES
- NO

Can the pleiotropic effects be avoided?
- YES
- NO

Targeted genome editing technologies or not to introduce into breeding practice

**Step 4 | Discovery of naturally mutated alleles of the S genes**
Are there natural mutations available in the germplasm?
- YES
- NO

Introgression into elite cultivars
Induce variation using mutagen

---

Figure 2. Guidelines on how to find and use susceptibility genes in *Coffea* spp.
Table 2. Possible susceptibility gene orthologs (according to van Schie and Takken [53]) in *Coffeea* spp.

<table>
<thead>
<tr>
<th>Coffee Disease</th>
<th>Pathogen</th>
<th>Susceptibility Gene</th>
<th>Accession Number</th>
<th>Gene Product</th>
<th>Plant Species</th>
<th>Disease Pathogen</th>
<th>Reference</th>
<th>Ortholog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coffee berry disease</td>
<td>Colletotrichum kahawae</td>
<td>IRG1/PAI1M</td>
<td>GL298201167</td>
<td>Transcription factor</td>
<td>Medicago</td>
<td>Anthracnose</td>
<td>Colletotrichum trifolii</td>
<td>[56]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DMR6</td>
<td>A5g24530</td>
<td>Oxidoreductase</td>
<td>Arabidopsis</td>
<td>Anthracnose</td>
<td>Colletotrichum higginsianum</td>
<td>[83–85]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SGR</td>
<td>MH493893</td>
<td>STAYGREEN</td>
<td>Cucumber</td>
<td>Anthracnose</td>
<td>Colletotrichum orbiculare</td>
<td>[86,87]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOX3</td>
<td>Zm00001d033623</td>
<td>Lipoygenase</td>
<td>Maize</td>
<td>Anthracnose</td>
<td>Colletotrichum graminicola</td>
<td>[88]</td>
</tr>
<tr>
<td>Coffee wilt disease</td>
<td>Fusarium xylarioides</td>
<td>PFT1</td>
<td>At1g25540</td>
<td>Mediator25 subunit</td>
<td>Arabidopsis</td>
<td>Fusarium wilt</td>
<td>Fusarium oxysporum</td>
<td>[89]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSY1R</td>
<td>At1g72300</td>
<td>Phytosulfokine receptors</td>
<td>Arabidopsis</td>
<td>Fusarium wilt</td>
<td>Fusarium oxysporum</td>
<td>[90]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XSP10</td>
<td>HM590582</td>
<td>Lipid transfer protein</td>
<td>Tomato</td>
<td>Fusarium wilt</td>
<td>Fusarium oxysporum</td>
<td>[91]</td>
</tr>
<tr>
<td>Root-knot nematodes</td>
<td>Meloidogyne spp.</td>
<td>KMD3</td>
<td>At2g44130</td>
<td>F-box/Kelch repeat protein</td>
<td>Arabidopsis</td>
<td>Root-knot nematode</td>
<td>Meloidogyne incognita</td>
<td>[92]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCS2A1</td>
<td>At4g22910</td>
<td>Cdb1-type activators</td>
<td>Arabidopsis</td>
<td>Root-knot nematode</td>
<td>Meloidogyne incognita</td>
<td>[93]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCS2B</td>
<td>A5g13840</td>
<td>F-box/Kelch repeat protein</td>
<td>Arabidopsis</td>
<td>Root-knot nematode</td>
<td>Meloidogyne incognita</td>
<td>[92]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
S genes involved in host and anthracnose (*Colletotrichum* spp.) compatible interactions were shown in *Medicago*, *Arabidopsis*, cucumber, and maize [56,83–88]. *IRG1* encodes a Cys(2)His(2) zinc finger transcription factor, PALM1 in *Medicago truncatula*. An *irg1* mutant was shown to significantly reduce *C. trifolii* spore germination and appressoria formation [56]. The *Arabidopsis Downy Mildew Resistant 6 (DMR6)* encodes a 2-oxoglutarate (2OG)-Fe(II) oxygenase with unknown function [83,84]. The *dmr6* mutant carrying a recessive mutation provides resistance against the anthracnose fungus *C. higginsianum* [84,85]. A loss-of-susceptibility mutation in the *STAYGREEN (CsSGR)* gene is responsible for the anthracnose (*C. orbiculare*) resistance in cucumber [86]. The *Csgr*-mediated resistance is achieved through the inhibition of the normal function of the chlorophyll degradation pathway and reactive oxygen species production [87]. The host oxylipin metabolism pathway mediated by a specific plant lipoxygenase 9-LOX isoform, LOX3, is required for fungal spore production and disease development. In maize *lox3* mutants, conidia production of anthracnose fungus *C. graminicola* and disease severity were significantly reduced [88].

Host *S* genes that are required for susceptibility to the causal agent of the Fusarium wilt disease, *F. oxysporum*, have been identified in two model species, *Arabidopsis* and tomato [89–91]. The *PHYTOCHROME AND FLOWERING TIME1 (PFT1)* gene, which encodes the MEDIATOR25 subunit of Mediator, acts as a positive regulator of jasmonate (JA) signaling and is required for JA-dependent defense responses against necrotrophic fungal pathogens. *PFT1* confers susceptibility to *F. oxysporum*, a hemi-biotrophic fungal pathogen; however, the increased *F. oxysporum* resistance was observed in the *Arabidopsis pft1* mutant due to the reduced expression of the JA-associated defense gene [89]. In *Arabidopsis*, *PSY1R*, which is the putative receptor for an endogenous tyrosine-sulfated peptide growth regulator *PSY1*, promotes susceptibility to Fusarium wilt disease. The *psy1r* mutant was able to upregulate robust defense responses [90]. The xylem sap protein XSP10 is required for full susceptibility of tomato to Fusarium wilt disease. *XSP10*-silenced tomato plants were observed to show reduced disease symptom severity [91].

*Arabidopsis* compatibility factors have also been revealed to be involved in the susceptibility of the root-knot nematode, *Meloidogyne incognita* [92,93]. Overexpression of the *Arabidopsis* F-box/Kelch-Repeat protein induces hypersusceptible responses with enhanced nematode attraction and infestation. In contrast, knocking down the F-box gene drastically decreases the number of parasitic nematodes inside the roots, thus the susceptibility to *M. incognita* [92]. Two of the key regulators controlling plant endocycle are *CCS52A* (*Cdhl/Fztr/Srw1*-type) and *CCS52B*. The expression of *CCS52A1* and *CCS52B* strongly induced nematode feeding site development as galls. Down-regulation of *CCS52* significantly restrained nematode development [93].

The next step requires the retrieval of *S* gene orthologs in *C. arabica* and *C. canephora*. To achieve this aim, the genome sequences of *C. arabica* and *C. canephora* and the BLAST tool are important resources. Functional characterized *S* gene protein sequences should be used as a query in a BLASTP program against the *Coffeea canephora* genome database [94] to search for homologous sequences, which will be used for constructing a phylogenetic tree to reveal the closely related homologues per candidate *S* gene. To illustrate this step, we performed the BLAST analysis on the genes listed in Table 2. Then, the retrieved *C. canephora* amino acid sequences and queried amino acid sequences were subjected to phylogenetic analyses using standard parameters, and the *C. canephora* sequences displaying a high level of homology were selected and considered as *S* genes orthologues (Table 2). The same procedure was applied to identify *C. arabica* orthologues using a tblastn analysis at the *Coffeea arabica*–Phytozome v12.1 website (Table 2) [11]. In case multiple candidates are identified based on sequence relatedness, transcript abundance of each gene can be assessed upon pathogen infection. It is known that many *S* genes are induced upon pathogen infection, for example, the *DMR6* gene is induced upon infection with *Hyaloperonospora parasitica* [84] and also the *mlo* gene is induced in barley upon infection with *Blumeria graminis f. sp. hordei* [95]. Hence, the selection of induced genes is frequently used to identify *S* genes.
The second reverse genetics strategy combines QTL mapping studies with candidate gene approaches. The presence of an S gene candidate in a known QTL region could be used to further investigate whether the resistance is caused by an impaired S gene. This approach was used in a study on powdery mildew resistance in cucumber. A cucumber Mlo gene, CsaMlo8, was found to co-localize with a known QTL for resistance to powdery mildew [76]. At present, there are relatively few QTL reports for disease resistance in coffee [44,48]. Among them, several QTLs were inherited in a dominant manner [48]. For many of the other QTLs, it is not clear from the studies whether the resistances inherited dominantly or recessively. With markers that are closely linked to (recessive) resistance genes/QTLs [45–47], their physical locations can be defined on chromosomes in order to check whether there is overlap with known S gene candidates. A coffee orthologue of the known Arabidopsis S gene KMD3 (F-box/Kelch-repeat protein) was found to co-localize with a known molecular marker, SSRCafe-39, which was associated with high level of reproduction factor or high susceptibility to root-knot nematodes in a segregating population derived from a cross between resistant C. canephora var. robusta and a susceptible C. arabica cultivar [47,92] (Figure 1). This gene might be the causal gene to promote nematode susceptibility. However, a detailed mapping study or a functional analysis of KMD3 is required to get an insight into whether a loss of function mutation in KMD3 may cause resistance.

7. Turning the Weak Link into the Winning Weapon—How to Make Use of S Genes in Coffee Resistance Breeding?

Once a candidate gene has been identified, the next step is to verify that it actually acts as an S gene (Figure 2). This is of particular interest for the application of the disrupted S gene-based resistance in breeding practices. Through transgenic approaches, the function of a candidate S gene can be verified by knocking down and/or knocking out its expression with either a transient or a stable assay. Depending on the reverse genetics approach chosen for the selection of S gene candidates, different strategies for functional characterization can be adopted. For example, a transient assay such as virus-induced gene silencing (VIGS) can be used to rapidly screen candidate genes [96]. However, VIGS is known to be patchy throughout the plant and to be prone to environmental influences [97–99]. To circumvent these issues, stable transformation with CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats) or RNA interference (RNAi) can be used as an alternative. Genome editing can be utilized to introduce deletions or single nucleotide polymorphisms in the coding sequence or the promoter region of the S gene candidates [100]. For high throughput assays to screen for S gene candidates, it is possible to design CRISPR-Cas9 constructs which target multiple genes simultaneously. Moreover, S gene candidates can also be functionally characterized using RNAi [101]. Silencing might circumvent potential pleiotropy accompanied by targeted knock-outs as the residual expression of the target could still be sufficient to fulfill the function for the host.

Following the guidelines at this point, the correct S gene has been found and functionally validated. The next step is its utilization in plant breeding (Figure 2). Although conventional breeding methods are mainly used for coffee improvement, several major limitations are associated with such breeding efforts such as the long generation time of the coffee tree, differences in ploidy level between tetraploid C. arabica and other diploid species, and the lack of genetic variation to some (a)biotic stresses [9,102]. Genetic modification technologies are, thus, considered as an extension of traditional breeding efforts. However, the deployment of transgenic plants in agriculture is in many parts of the world at least still under debate. The non-transgenic strategy is important to identify naturally occurring resistant genotypes by searching for natural variation within the gene pool. Non-functional S gene alleles present in wild germplasm can be exploited in breeding programs. An example for a naturally occurring S gene allele is a wild tomato accession from Ecuador that was shown to be resistant to powdery mildew which was found to be caused by an impaired mlo allele (SlMlo1) [103]. Nowadays, increasing amounts of sequence information have become available, for example whole genome sequences of 13 different C. arabica cultivars. Allele mining is, thus, a promising approach to identify natural mutated S genes alleles [11].
Identification of recessive alleles is not restricted by the naturally occurring traits only. In practice, if the naturally occurring recessive allele is absent in the gene pool, genetic variation can be created by artificial ways. Both chemical and physical mutagenesis are used for this purpose. Among them, the use of chemical mutagens like EMS (ethyl methanesulfonate) is one of the most popular methods which induces large numbers of point mutations that spread the genome [104]. Subsequently, mutations in a specific target sequence can be identified using “Targeting Induced Local Lesions IN Genomes” (TILLING) [105]. The successful examples of TILLING include the identification of virus resistance in pepper due to nucleotide changes in the aforementioned elfs [106] or the identification of mlo-mediated resistance against powdery mildew in polyploid wheat [107]. A more targeted approach for impairing S genes or screening mutant populations is the use of genome editing. Genome editing entails the alteration of an organism’s DNA at a specific location in the genome by deleting, adding or modifying DNA. Such techniques, including zinc finger nucleases (ZFNs) [108], transcription activator-like effector nucleases (TALENs) [109], and the CRISPR/Cas system [110], have gained increasing attention over the last few years and have revolutionized the field of biology. The CRISPR-associated system is particularly useful in polyploid outcrossing crops such as C. arabica for which homozygous tetra-allelic mutants are difficult to obtain. Recently, proof of concept was established targeting the C. canephora genome by the CRISPR/Cas9 method, which gave rise to stable transformants containing biallelic homozygous mutations [111].

8. Two Sides of the Same Coin—Advantages and Disadvantages of Using Impaired S Genes

So far, numerous host genes have been identified and defined as S genes [53]. The increasing popularity of the impairment of S genes for crop protection and resistance breeding can be explained by several reasons, additionally to the above mentioned one—the fact that they are functionally conserved between plant species. First of all, there is the aspect of durability. Overcoming resistance mediated by an impaired S gene is thought to be more challenging for a pathogen than to overcome R gene-mediated resistance. The well-studied S gene for powdery mildew susceptibility is MLO. Naturally occurring mlo-based resistance was discovered in barley [72] and was characterized as a membrane-anchored protein [112]. Barley cultivars mounting mlo-based resistance have maintained a high powdery mildew resistance level during the past decades ever since their first introduction into the market in the 1980’s [113,114]. Overcoming mutated S gene mediated resistance implies that the pathogen can acquire an alternative means of infecting the host. This is most unlikely to happen, making this kind of resistance more durable. Secondly, impaired S genes are known to not only provide resistance to one strain or race of a given pathogen, but to many if not all [51]. This makes their resistance less specific than resistance mediated by R genes. This is again exemplified by the mlo mutant in barley, for which resistance has not yet been overcome by any powdery mildew strain tested so far [72,113]. Thirdly, impaired S genes were also found to provide resistance to multiple pathogens and therefore can lead to broad-spectrum resistance. An example is the aforementioned cucumber STAYGREEN gene, for which a naturally occurring single nucleotide polymorphism confers resistance to three diseases, namely downy mildew, bacterial angular leaf spot and fungal anthracnose [87].

As with many methods, besides the benefits, the use of impaired S genes also has its limitations which need to be overcome. Application of a dysfunctional S gene into a breeding program can be problematic if there are pleiotropic effects associated with the impairment of S genes [115]. Such pleiotropy can have influences on plant growth or fertility and hence are undesired for breeding. In many cases, it concerns S genes in class two which are frequently associated with constitutively elevated salicylic acid levels. Abundant levels of this phytohormone might lead to enhanced disease resistance, subsequently it can also result in severe fitness costs. One example is the defense no death 1 (dnd1) mutant in Arabidopsis for resistance to P. syringae [116]. The dnd1 mutant accumulated less bacteria than inoculated wild type plants, but at the same time the mutants were severely dwarfed. Another aspect that needs to be considered is that mutants with a disturbed hormonal balance might, on the one hand, gain resistance to one pathogen, but on the other hand, loose resistance to another [117].
For example, the symptoms to multiple avr genotypes 4 (sma4) mutant is resistant to the necrotrophic pathogen B. cinerea, while it became more susceptible to P. syringae than wild type plants [118]. Additionally, impairment of S genes can negatively influence other traits, such as sensitivity to abiotic stresses. The EXLA2 gene confers resistance against B. cinerea as well as A. brassicola; however, it was more sensitive to salt and cold stress than wild type plants [60]. In many cases, the occurrence of pleiotropic effects can be uncoupled from the resistance [119]. The possible side effects of dysfunctional S genes are dependent on the type of gene that is impaired, as well as differences between plant species. For example, it was found that silencing of DND1 in tomato leads to dwarfing and autonecrosis, while silencing in potato resulted only in mild autonecrosis [120].

More consideration is needed regarding the occurrence of pleiotropic effects. If a deleterious phenotype is visible, it is important to consider the benefits and limitations as well as whether or not to introduce the mutated gene into breeding practice. Loss of the susceptibility function needs to be adjusted so that pleiotropic effects are uncoupled from the resistance, for instance by introducing missense mutations using targeted genome-editing technologies.

Current developments in plant breeding are moving towards using different forms of resistance simultaneously [121]. This can be achieved by pyramiding various genes that cover a diverse range of resistance mechanisms. In sustainable crop breeding programs, the application of impaired S genes should therefore be integrated with R genes or quantitative resistance in order to combat diseases.

Author Contributions: Conceptualization, L.C. and Y.B.; investigation, L.C. and K.H.; writing—original draft preparation, L.C. and K.H.; writing—review and editing, L.C., K.H., R.G.F.V. and Y.B. All authors have read and agreed to the published version of the manuscript.

Funding: L.C. was funded by a fellowship from the China Scholarship Council. K.H. was supported by Foundation Topconsortium voor Kennis en Innovatie (TKI) Starting Materials, project number 1409-026.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lécotier, A.; Besse, P.; Charrier, A.; Tchakaloff, T.N.; Noirot, M. Unraveling the origin of Coffea arabica ‘Bourbon pointu’ from La Réunion: A historical and scientific perspective. Euphytica 2009, 168, 1–10. [CrossRef]
5. Ovalle-Rivera, O.; Läderach, P.; Bunn, C.; Obersteiner, M.; Schroth, G. Projected shifts in Coffea arabica suitability among major global producing regions due to climate change. PLoS ONE 2015, 10, e0124155. [CrossRef]
9. Van der Vossen, H.; Bertrand, B.; Charrier, A. Next generation variety development for sustainable production of arabica coffee (Coffea arabica L.): A review. Euphytica 2015, 204, 243–256. [CrossRef]
11. The UC Davis Coffee Genome Project. The Coffee (C. arabica) genome UCD v0.5. Available online: https://phytozome-next.jgi.doe.gov/info/Carabica_v0_5 (accessed on 20 July 2020).


21. Rutherford, M.A. Current knowledge of coffee wilt disease, a major constraint to coffee production in Africa. *Phytopathology* 2006, 96, 663–666. [CrossRef]


27. Greco, E.B.; Wright, M.G. Ecology, biology, and management of *Xylosandrus compactus* (Coleoptera: Curculionidae: Scolytinae) with emphasis on coffee in Hawaii. *J. Integr. Pest Manag.* 2015, 6, 7. [CrossRef]


34. Arneson, P.A. *Coffee Rust in the Plant Health Instructor*; The American Phytopathological Society: Eagen, MN, USA, 2000. [CrossRef]


61. Garcia-Ruiz, H. Susceptibility genes to plant viruses. *Viruses* 2018, 10, 484. [CrossRef] [PubMed]


71. Huang, T.S.; Wei, T.; Lalliberté, J.F.; Wang, A. A host RNA helicase-like protein, AtRH8, interacts with the potyviral genome-linked protein, VPg, associates with the virus accumulation complex, and is essential for infection. *Plant Physiol.* 2010, 152, 255–266. [CrossRef] [PubMed]


92. Curtis, R.H.; Powers, S.J.; Napier, J.; Matthes, M.C. The Arabidopsis F-box/Kelch-repeat protein At2g44130 is upregulated in giant cells and promotes nematode susceptibility. Mol. Plant Microbe Interact. 2013, 26, 36–43. [CrossRef]
97. Liu, Y.; Schiff, M.; Dinesh-Kumar, S.P. Virus-induced gene silencing in tomato. Plant J. 2002, 31, 777–786. [CrossRef]


**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).