Biotechnology Research and Innovation (2020) 3, 80–94

Biotechnology Research & Innovation

http://www.journals.elsevier.com/biotechnology-research-and-innovation/

REVIEW ARTICLE

Plant NLR receptor proteins and their potential in the development of durable genetic resistance to biotic stresses

Amanda Cristina de Araújo, Fernando Campos De Assis Fonseca, Michelle Guitton Cotta, Gabriel Sergio Costa Alves, Robert Neil Gerard Miller*

Universidade de Brasília, Instituto de Ciências Biológicas, Departamento de Biologia Celular, Campus Universitário Darcy Ribeiro, Asa Norte, 70910-900, Brasília, DF, Brazil

Received 18 November 2019; accepted 31 January 2020
Available online 19 February 2020

KEYWORDS
Innate immunity;
NLR resistance genes;
RenSeq

Abstract In order to meet global food security demands in the next decades, considerable changes are required for sustainable agriculture in the context of plant disease, with sufficient food production depending on the development of durable genetically disease resistant crops. For this, further advances are required in our understanding of the plant innate immune system and how plants respond to invading pathogenic micro-organisms. Over the past 20 years, considerable research has been conducted into the characterization and cloning of plant nucleotide-binding, leucine-rich repeat (NLR) immune receptors. These intracellular receptors can recognize directly or indirectly pathogen effector proteins, resulting in effector-triggered immunity (ETI). Elucidation, however, of the diversity of NLR resistance gene families and the molecular basis of NLR-driven effector recognition and defense signaling is incomplete. Here, we present a summary of the understanding of NLR structure, function, genomic organization and diversity in plants. Recent advances in target enrichment approaches for NLR characterization and function validation are highlighted in the context of NLR engineering possibilities for accelerated durable genetic resistance to biotic stresses.

Introduction

Climate change and rapid expected global population growth are important challenges of the 21st century. Agricultural production will require optimization across numerous fronts to guarantee food security, with reduction of the current 30% losses of global crop production due to biotic stresses at the forefront of priorities. Disease control strategies have been heavily reliant upon agrochemical usage, associated with negative environmental impact and increasing incidence of resistance to fungicides developing in pathogen populations (Lucas, Hawkins, & Fraaije, 2015). Genetic resistance, by contrast, offers a long-term approach for disease control, appropriate for global sustainable agriculture. Although plant breeding has been widely employed in the 20th and

* Corresponding author.
E-mail: robertmiller@unb.br (R.N. Miller).

https://doi.org/10.1016/j.biori.2020.01.002

21st centuries for introgression of disease resistance genes (*R* genes) into commercial cultivars, the numbers of cultivated crop species and cultivars containing appropriate *R* genes for disease resistance are limited (Miller, Alves, & Van Sluys, 2017). In this context, advances in fundamental understanding of the plant immune system will have far reaching implications for genetic resistance development, appropriate for effective and durable disease control and global sustainable agriculture.

Here, we provide an overview of current understanding of plant nucleotide-binding, leucine-rich repeat (NLR) immune receptors, in terms of structure, function and genomic organization and diversity in plants. Recent advances in enrichment approaches for accelerated NLR characterization and function validation are also discussed, appropriate for genetic engineering development for genetic resistance to biotic stresses.

Plant immunity models

Considerable advance has been made into our understanding of the plant innate immune system. How plants can withstand pathogen challenge is now known to involve a series of receptor proteins for pathogen effector protein recognition. In one specific defense network, termed pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), plant host transmembrane pattern recognition receptors (PRRs) (Dangl & Jones, 2001; Monaghan & Zipfel, 2012) recognize conserved and essential pathogen (or microbial)-associated molecular patterns (PAMPs/MAMPs) (Boutrot & Zipfel, 2017; Jones & Dangl, 2006). Such microbial PAMPs include molecules such as bacterial flagellin, elongation factor Tu (EF-Tu), lipopolysaccharide (LPS), peptidoglycan, bacterial RNA and fungal chitin (Felix, Duran, Volko, & Boller, 1999; Kunze et al., 2004; Meyers et al., 1999; Thomma, Nürnberger, & Joosten, 2011; Zipfel, 2014). PRR plant immune receptor proteins typically contain an extracellular domain with ligand binding capacity, a transmembrane domain, and an intracellular kinase domain (Zipfel, 2014). PAMP recognition and activation of PRRs triggers intracellular signaling and plant defense gene expression to impede pathogen advance (Chisholm, Coaker, Day, & Staskawicz, 2006). Functioning at the pathogen species level or above, PTI is responsible for non-host resistance, explaining how plants are resistant to many potential plant pathogens. Specific evolving pathogen effector proteins, or avirulence (Avr) proteins, when secreted by adapted pathogens into the host cell cytoplasm, however, can affect intracellular signaling and suppress PTI, resulting in effector-triggered susceptibility (ETS) and the disease development (Boller & Felix, 2009; Jones & Dangl, 2006).

In a second plant immunity defense network, intracellular resistance (R) protein receptors recognize directly specific pathogen effectors or indirectly their interaction effects on host cell components, activating a response known as effector-triggered Immunity (ETI), qualitative or monogenic resistance (Ellis, Dodds, & Pryor, 2000; Jones & Dangl, 2006). ETI is responsible for a more intense host response, involving calcium ion signaling, reactive oxygen species production, accumulation of pathogenesis-related (PR) proteins (Chisholm et al., 2006; Gururani et al., 2012; Hammond-Kosack & Jones, 1996) and changes in levels of plant hormones salicylic acid (SA) and jasmonic acid (JA) (Creelman & Mullet, 1995). Such responses can also involve the hypersensitive response, which consists of a programmed and localized death at the infection site (Dangl & Jones, 2001; Heath, 2000), limiting pathogen advance. Subsequent systemic acquired resistance (SAR) can also occur, whereby a broad spectrum and generalized plant response increases resistance to subsequent pathogen challenge (Dong, 2001; Spoel & Dong, 2012).

NLR plant disease resistance genes

Plant disease resistance (*R)* genes that encode intracellular resistance protein receptors confer resistance to diverse pathogen effectors, including those from bacteria, fungi, oomycetes, nematodes and viruses. To date, over 300 cell surface and intracellular *R* genes have been cloned and characterized, with resistance function confirmed *in planta*, for not only model species such as *Arabidopsis thaliana*, but also important crop species such as *Oryza sativa*, *Zea mays*, *Triticum aestivum*, *Hordeum vulgare*, *Solanum tuberosum* and *Solanum lycopersicum* (see reviews by Sekhwal et al., 2015 and Kourelis & van der Hoorn, 2018 for references therein). Despite a diverse range of recognized pathogen effectors, cloned *R* genes share significant homologies in amino acid sequences and structural motifs. Initial speculation of common protein−protein interactions as components of receptor systems and conserved roles in defense signaling responses (Dangl, 1994) was later confirmed, with effector protein recognition occurring either directly (Dodds & Rathjen, 2010) or indirectly (Dangl & Jones, 2001; Van Der Hoorn & Kamoun, 2008).

The most abundant *R* genes characterized to date comprise a single family that encode intracellular multidomain receptor proteins carrying a stereotypical nucleotide binding site (NBS) and leucine-rich repeat (LRR) domain. These are known as NLR immune receptors, as their function is analogous to that observed in intracellular NOD-like receptors in the animal innate immune system (Bentham, Burdett, Anderson, Williams, & Kobe, 2017). The NLR protein family may have originated in green algae (Andolfo et al., 2019; Gao et al., 2018) and was well-defined early in the land plant lineage (Yue, Meyers, Chen, Tian, & Yang, 2012). NLRs are one of the largest multi-gene families known in plants, with genomes harbouring hundreds or even thousands of such genes (Steuernagel et al., 2018).

The central NBS protein domain in NLRs is highly conserved, whilst the C-terminal LRR domain is variable in length, comprising 10−40 short LRR (Cannon et al., 2002). The NBS domain is a component of the larger NB-ARC (nucleotide-binding adaptor shared by apoptotic protease activating factor 1 (APAF1), R proteins and cell death 4 (CED-4) domain (Van Der Biezen & Jones, 1998a), and contains characteristic motifs widely conserved in plants (Tameling et al., 2002). The central nucleotide binding NB-ARC domains binds ADP or ATP molecules, which induces a protein conformation change and the switch from ''OFF'' to an ''ON'' state (Steele, Hughes, & Banfield, 2019; Van Ghelder et al., 2019). The NB-ARC domain, which is often used to perform phylogenetic studies contain highly

conserved motifs involved in intra- and extra molecular interactions (Moffett, Farnham, Peart, & Baulcombe, 2002; Van Ooijen et al., 2008). These include the motifs P-loop (Walker A / kinase 1), RNBS-A, kinase 2 (Walker B), kinase 3a, RNBS-B, RNBS-C, GLPL and RNBS-D. This *R* gene class can be further subdivided into sub-classes, based on different N-terminus architecture, namely TIR-NBS-LRR, or Toll/Interleucina-1 / TNLs, and non-TIR-NBS-LRR, or CNLs. For this second subclass, member proteins can be further divided into those containing either a predicted helical coiled-coil (CC) motif (Meyers, Kozik, Griego, Kuang, & Michelmore, 2003; Maekawa, Kufer, & Schulze-Lefert, 2011; Meyers et al., 1999) or a zinc finger or RPW8 domain (CCr-NLRs / RNLs) (Ameline-Torregrosa et al., 2008; Bent et al., 1994; Meyers, Morgante, & Michelmore, 2002; Milligan et al., 1998; Zhong & Cheng, 2016). The first sub-class contain a protein domain with homology to the *Drosophila* TOLL and human interleukin-1 receptor (TIR), with five α-helices covering a five-strand β-sheet (Ve, Williams, & Kobe, 2015). The TNL proteins, whilst being the most common sub-class of NLR proteins (Zhang et al., 2016), appear to be restricted to dicots, with suggested evolution after the divergence of monocots and dicots (Kim et al., 2012; Meyers et al., 2003). By contrast, CNL proteins occurs in both monocots and dicots (Meyers et al., 1999).

Although structural analysis of NLR proteins and their specific binding to effector proteins has still to be determined, as members of the STAND (signal transduction ATPases with numerous domains) superfamily, these intracellular receptor proteins are known to control signal transduction, working as molecular switches. These occur either in open and activated or closed and inactivated states, as a result of nucleotide binding (Takken & Tameling, 2009; Takken, Albrecht, & Tameling, 2006; Van Ghelder et al., 2019).

The NBS domain appears to control this protein conformation, with binding to the nucleotide ATP enabling an active conformation, and binding to ADP an inactive conformation (Bonardi et al., 2011; Takken et al., 2006). In the absence of pathogen effectors, by contrast, intramolecular binding of domains or extramolecular binding with additional proteins will result in a closed and inactive conformation. Mutation in the NBS P-loop motif (GxxxxGKT/S) that binds to nucleotides has been shown to result in loss of function (Williams et al., 2011), supporting its' role in conformation changes during direct interaction with effector proteins.

The N-terminal NLR domains appear to be involved in regulation of signal transduction pathways for downstream defense responses, following recognition of pathogen presence (DeYoung & Innes, 2006; Ellis & Jones, 1998; Van Der Biezen & Jones, 1998b). Analysis of deletion in the CC or TIR domains has shown that these domains are required for signaling and subsequent cell death responses (e.g. Adachi et al., 2019; Collier, Hamel, & Moffett, 2011; Swiderski & Innes, 2001). Recently, TIR domains of plant NLRs have been shown to be enzymes that cleave NAD^+ for activation of immune responses (Wan et al., 2019). Oligomerization of the N-terminal region has also been proposed to be necessary for signal transduction (Monteiro & Nishimura, 2018), with Wang et al. (2019) demonstrating through cryo-electron microscopy the oligomerization of an *A. thaliana* CNL (ZAR1) into an active wheel-like pentameric NLR receptor complex or ''resistosome'' during the plant immune response.

The C-terminal LRR domain comprises a tandem repeat of amino acids. The consensus sequence of LxxLxLxxNxL (where L is a leucine residue, x any amino acid, and N either an asparagine, threonine, serine or cysteine (Bella, Hindle, McEwan, & Lovell, 2008; Kajava, 1998; Stange, Matus, Domínguez, Perez-Acle, & Arce-Johnson, 2008) is required for the structural arrangement of the domain (Dubey & Singh, 2018). The tertiary protein structure of this domain is in the form of repeated horseshoe superhelices for each tandem amino acid repeat, with an inner surface of parallel ß-strand loop units assumed to be a site for interaction with other proteins, and an outer surface of repeating α-helix-loops connected with ß-strands loop units. Under diversifying selection, mutation in the LRR exposed amino acid residues enables co-evolution for specific recognition and physical interaction with diverse pathogen effector molecules (Dodds et al., 2006; Ellis et al., 2000; Jia, McAdams, Bryan, Hershey, & Valent, 2000; Jones & Jones, 1997; Kobe & Deisenhofer, 1994; Krasileva, Dahlbeck, & Staskawicz, 2010; Ravensdale et al., 2012; Shen et al., 2003; Wang et al., 2007). Removal of this domain has not only been shown to result in a loss of effector recognition (Rairdan & Moffett, 2006; Van Ooijen et al., 2008) but also to increase auto necrosis in NLRs, indicating an additional role for the LRR domain in NLR auto-inhibition (Ade, DeYoung, Golstein, & Innes, 2007; Bentham et al., 2017; Faustin et al., 2007).

Evidence for the involvement of additional NLR domains in effector recognition also indicates that effector binding may occur in a competitive manner, destabilizing intramolecular interactions in the NLR protein and activating a specific NLR form that results in signaling responses (Bernoux et al., 2016; Ravensdale et al., 2012). Mutation in N-terminal CC domains, for example, has also been shown to influence effector recognition (Giannakopoulou et al., 2015), with identical LRRs but different TIR domains in genes at the L locus in flax conferring different pathogen recognition specificities (Ellis, Lawrence, Luck, & Dodds, 1999).

Direct and indirect effector recognition

In the gene-for-gene model for *R* gene-mediated resistance, originally proposed by Flor (1942) as a binary model for a single plant resistance gene interacting with a single pathogen virulence gene, R protein recognition of pathogen effectors is now recognized to occur through both direct and indirect interactions. In indirect interactions, R proteins have been proposed to effectively monitor or guard pathogen effector-induced alterations in the integrity of other conserved host molecules that function as targets of pathogen effector proteins, resulting in defense responses (Dangl & Jones, 2001; Jones & Dangl, 2006; Lozano-Torres et al., 2012; Swiderski & Innes, 2001; Van Der Biezen & Jones, 1998b; Van Esse et al., 2008). In such indirect interactions, the guard model involves a second host protein, known as a guardee, which is a defense protein operative target of a pathogen effector. In this model, initially proposed to explain why the protein kinase Pto requires the NLR protein Prf for defense activation, an R protein will monitor, or guard, changes occurring

in the guardee following interaction with an effector protein, resulting in activation of defense responses (Dangl & Jones, 2001; Van Der Biezen & Jones, 1998b). In a second format, the decoy model also requires an additional host protein that performs a role exclusively as a decoy. In contrast to the guard model, the decoy will play no specific role in plant defense, simply serving to mimic a true defense component effector target and attract the pathogen effector (Jones & Dangl, 2006; Shao et al., 2016; Van Der Hoorn & Kamoun, 2008). Effector-induced modification of the decoy will then be perceived by an R protein, with subsequent activation of the plant immune response. This modification of the guard model was proposed to explain the role of Pto as a decoy for the effector protein AvrPto in tomato. As a mimic of other PRRs, but without a direct defense function in this model, interaction with the effector is monitored by Prf with detected modification triggerings defense activation (Zong, Xiang, Zou, Chai, & Zhou, 2008). Such indirect mechanisms involving guardees or decoys have been proposed to enable host recognition of different pathogens, without the need for different genomically-encoded receptor R proteins for every possible pathogen (Dangl & Jones, 2001).

Within the context of decoys, *R* genes can also encode NLRs that contain additional integrated domains (NLR-IDs), which may play sensor roles in pathogen effector recognition (Baggs, Dagdas, & Krasileva, 2017; Kroj, Chanclud, Michel-Romiti, Grand, & Morel, 2016; Sarris, Cevik, Dagdas, Jones, & Krasileva, 2016). Integration has been described in numerous locations within NLRs, including N- and C-terminal regions, as well as between NLR domains (Cesari, Bernoux, Moncuquet, Kroj, & Dodds, 2014; De La Concepcion et al., 2019; Kroj et al., 2016; Maqbool et al., 2015; Sarris et al., 2016). Over 100 atypical domains within NLR-encoding genes have now been identified across plant genomes, with common NLR-IDs including transcription factors WRKY and BED, protein kinases and DUF protein domains (Kroj et al., 2016; Monteiro & Nishimura, 2018; Sarris et al., 2016). NLR-IDs have recently been discovered to function with other genomically linked NLRs. Working in pairs, a traditional NLR can serve as a signaling molecule, guarding (or helping) a receptor NLR-ID that interacts with a pathogen effector as a sensor or decoy (Cesari et al., 2013). In the latter, integrated motifs for effector interaction show evolutionary links to the pathogen's typical host targets during infection, and therefore act as baits (Cesari et al., 2014; Nishimura, Monteiro, & Dangl, 2015). A number of such NLR-IDs have recently been examined in detail, with each co-expressed, under the control of a single promoter, with an opposite orientation NLR lacking the ID, enabling heterodimerization of the NLR protein pair and subsequent signal transduction. Effector recognition by NLR-IDs may occur either following a direct binding with the pathogen effector, or as a result of an enzymatic activity of the effector on the ID (Kourelis & van der Hoorn, 2018). Examples of such NLR-IDs have been described in rice RGA5 and PiK NLRs. Both R genes harbor an HMA ID, which is related to the copper binding protein ATX1 in *Saccharomyces cerevisiae* (Cesari et al., 2013). Integration into each *R* gene was likely to have occurred independently, with the HMA domain at the C-terminal region in RGA5 and at the N-terminal region between the CC and NB-ARC domains in PiK. In RGA5, the ID has been shown to interact with *Magnaporthe oryzae* effectors Avr-Pia and Avr-CO39, which, together with RGA4, enables signaling and defense responses (Cesari et al., 2013). Similarly, in PiK, HMA ID alleles can interact with *M. oryzae* Avr-Pik effector alleles (Maqbool et al., 2015), with immune signaling triggered through the involvement of Pik-2 (Ashikawa et al., 2008).

In the context of the connections that can occur between NLRs during the immune response to pathogen presence, which is now recognized to involve NLRs working not only in pairs but also as networks (Adachi, Derevnina, & Kamoun, 2019; Wu, Derevnina, & Kamoun, 2018), certain NLRs act as functional requisites or 'helpers' for other 'sensor' NLRs, helping to initiate signal transduction and defense responses to pathogen invasion. Examples of such functionally linked helpers include *NGR1*, which is a CCr RNL required in disease resistance mediated by the *N* gene (a TNL) in tobacco, conferring resistance to *Tobacco mosaic virus* (Whitham et al., 1994). Similarly, for CNLs, *NRC* family helpers have also been reported (Wu et al., 2017). Whilst the evolutionary relationships between NLR pairs or networks are largely unknown, the NRC network appears to have expanded from a single NLR pair to up to half of all NLRs in certain asterid species (Wu et al., 2018). Many of these, together with numerous NRC-dependent NLRs, control immune responses to various pathogens and pests. The ADR1 family of NLRs in *Arabidopsis* have also been shown to work as helpers to sensor NLRs (Bonardi et al., 2011), enabling signal transduction after specific NLR sensor activation. Recent analysis has further investigated such a role for sensor NLRs, which, upon pathogen effector perception, act in releasing inhibition of autoactivity in helper NLRs, whose activity is normally inhibited in the absence of pathogen presence (Wu et al., 2018).

NLR structure and genomic organization

The structure of NLR-encoding genes is often complex, with not only the stereotypical domain encoding regions, but also with integrated domains. NLR numbers in plant genomes vary in orders of magnitude across plant species (e.g. Porter et al., 2009; D'hont et al., 2012; Gu, Si, Zhao, Yang, & Zhang, 2015), with higher numbers in woody plant species potentially reflecting evolution through duplication events due to greater pathogen exposure levels over time and the infrequency of meiosis (Baggs et al., 2017). NLRs occur not only as isolated genes but also commonly as homogeneous and heterogeneous NLR clusters of different sizes at specific genomic loci, within a distance of 200 kb of each other (Hulbert, Webb, Smith, & Sun, 2001)). Contiguous NLR-encoding genes in homogeneous clusters (i.e. sharing a common ancestor) can be nearly identical as a result of tandem duplications of paralogs (Meyers et al., 2003), whereas non-paralogous NLR-encoding genes in heterogeneous clusters of more distantly related genes, by contrast, will differ as a result of segmental and ectopic duplications of gene blocks and smaller groups of genes, as well as through random rearrangements (Borrelli et al., 2018; Meyers et al., 2003). This genomic organization likely favours gene evolution, with different *R* genes within a cluster potentially conferring resistance to different pathogen strains or to different pathogen species (Michelmore & Meyers, 1998; Van Der Vossen et al., 2000). NLR-encoding genes that work in pairs are also often organized in a head to head configura-

tion, likely under the control of the same promoter elements regulating expression (Saucet et al., 2015; Stein et al., 2018; Wu et al., 2018).

In addition to functional NLR-encoding genes, many NLR sequences in plant genomes are non-functional, with such pseudogenes harbouring indels, frameshift mutations or premature stop codons. The high rate of pseudogenization is considered to reflect the rapid coevolution of NLR loci with pathogen effectors. Evolutionary mechanisms occuring in NLR-encoding gene families include the abovementioned duplication events, unequal crossing-over, ectopic recombination and gene conversion and mutation events (Chen, Han, Jiang, Tian, & Yang, 2010; Guo et al., 2011; Michelmore & Meyers, 1998). Transposable element insertions are also overrepresented at NLR loci (Richter & Ronald, 2000), with evidence for Long Terminal Repeat retrotransposon (LTR-Rs)–mediated retroduplication also driving novel variation and evolution in these genes (Kim et al., 2017).

Truncated NLRs

In addition to complete TNL and CNL NLRs, as well as extended NLRs with IDs, truncated versions lacking specific domains are also abundant in plant genomes, including TIR-X (TX), TIR-NBS-X (TNX), CC-NBS-X, RPW8-X and RPW8-NBS-X (Baggs et al., 2017). Function determination in plant immunity has been limited to date, although a hypothesis of roles in heteroduplex formation with full length NLRs has been proposed and validated (Monteiro & Nishimura, 2018). *Chilling sensitive 1* (*CHS1)*, for example, a TNX-encoding gene in *Arabidopsis*, has been shown to require the neighbouring complete TNL *SOC3* for modulation of temperature-dependent autoimmunity (Zhang, Dodds, & Bernoux, 2017, Zhang et al., 2017).

Discovery approaches for NLR-encoding genes

Through PCR cloning

Prior to the development of next generation sequencing technologies, the isolation of subsets of plant NLRs across different plant species was achieved using degenerate primers for targeting known conserved domain motifs previously identified in model species, such as the NBS P-loop, GLPL and RNBS motifs (Kanazin, Marek, & Shoemaker, 1996; Leister et al., 1998). Considerable work on PCR cloning of target Resistance Gene Analogs (RGAs) has been reported during the last 25 years across a wide range of dicot species, including *Arabidopsis thaliana* (Aarts et al., 1998; Meyers et al., 1999; Speulman, Bouchez, Holub, & Beynon, 1998), cotton (He et al., 2004), soybean (Peñuela, Danesh, & Young, 2002), peanut (Bertioli et al., 2003), potato (Leister, Ballvora, Salamini, & Gebhardt, 1996), and perennials such as apple (Baldi et al., 2004). For a review of the extension of this approach see Sharma, Das, Kumar, & Lodha, 2009. In the case of monocotyledons, the method has also been widely applied, for example in rice (Leister & Katagiri, 2000; Mago, Nair, & Mohan, 1999), maize (Collins et al., 1998), wheat (Bozkurt, Hakki, & Akkaya, 2007) and banana (Azhar & Heslop-Harrison, 2008; Miller et al., 2008).

WGS and homology-based approaches for NLR annotation

With the rapid development of second and third generation sequencing technologies, whole genome sequencing and annotation has enabled more exhaustive analysis of NLR repertoires, with data available today for over 100 plant species (Monteiro & Nishimura, 2018). Although ever increasing numbers of NLRs are now therefore being predicted, automated gene callers are generally insufficient for accurate annotation of the entire NLR complement in a plant genome. Up to 50% of NLRs can be missed during such annotation, with many genes often split and annotated incorrectly (Steuernagel, Jupe, Witek, Jones, & Wulff, 2015). Manual identification of NLR-encoding genes in plant genome sequences has been achieved through a laborious application of a number of computational homology-based methods. These generally have taken into consideration the conserved motifs, domains and recognized structural features of these genes. Amongst others, these have been broadly based on BLAST similarity searches (Altschul et al., 1997), pfam (http://pfam.xfam.org/), and InterProScan, (http://www.ebi.ac.uk/Tools/pfa/iprscan5/), as well as through sequence alignment with reference sequences from curated databases, such as the plant resistance gene database PRGdb (Sanseverino et al., 2009). Domain sequence homolog searches using programs based on application of Hidden Markov Models (HMM), such as HMMER (Eddy & Wheeler, 2007), motif discovery using MEME (Bailey et al., 2009), transmembrane and signal peptide prediction using Phobius and TMHMM (Käll, Krogh, & Sonnhammer, 2004; Krogh, Larsson, Von Heijne, & Sonnhammer, 2001)) have also been frequently adopted in bioinformatic pipelines for NLR analysis. Secondary structure prediction for folds for CC domain prediction have also included tools such as nCOILS (Lupas, Van Dyke, & Stock, 1991), MultiCoil (Wolf, Kim, & Berger, 1997), MARCOIL (Delorenzi & Speed, 2002) and Paircoil2 (McDonnell, Jiang, Keating, & Berger, 2005).

Given the large numbers of repeat sequences present in plant genomes in general, annotation projects have often avoided the counting of transposon-related genes, applying filters to mask repetitive regions through cross reference with databases of repeat sequences such as those in Repbase (Bao, Kojima, & Kohany, 2015). Recently, Bayer, Edwards, & Batley (2018) identified repeat sequences in this database that may in fact be transposons fused to *R* genes. A consequence of such repeat masking using reference sequences may be an incorrect annotation and loss of *R* genes in genome annotations. The employment of different repeat masker approaches over the years has been hypothesised by this group to explain in part the large variance observed in *R*-gene content in closely related species, for example as observed in *Brassica* sp.

Until recently, no single and standardized platform was available for accurate and exhaustive NLR identification from DNA sequence data. As such, different annotation criteria have limited accurate comparison of NLR data from different studies. NLR-parser (github.com/steuernb/NLR-Parser) was developed as a Java application to automate accurate annotation of sequence data according to 20 known possible conserved amino acid

motifs described in the plant NLR-encoding gene family (Jupe et al., 2012). With WGS, scaffolds or sequence contig data employed as input data, the Motif alignment and search tool (MAST) output within the MEME suite is then refined to identify sequences harbouring characteristic NLR motifs (Steuernagel et al., 2015). This pipeline for classifying NLR-encoding genes is based on the identification not only these short target motifs of up to 50 amino acids, but their presence in tuples or triplets, as a means of avoiding false positive NLRs due to individual motifs occurring randomly across the genome. In the case of multiple NLRs within a single sequence, however, NLR-parser cannot identify the border between NLRs, with classification of a single NLR in such sequences resulting in errors in gene identification and an underestimation of the true number of NLRs in a genome. In a modified pipeline, NLR annotator, such problems are avoided through the analysis of genomic DNA in an overlapping fragment format. Fragments are analysed with MAST for target motifs across all six reading frames. Motif combinations are then determined and likely NLR-positive fragments converted back to nucleotide sequences. Data from all fragments are then reintegrated, with definitive motif positions and combinations in the core NB-ARC domain used as a seed for elongation into the N-terminal and C-terminal encoding regions for accurate identification of the NLR loci (Steuernagel et al., 2018).

RenSeq

Although whole genome sequencing of plant genomes since 2000 has advanced analysis of genome diversity, gene discovery and determination of function, high quality annotation of genome data has been a major challenge (Yandell & Ence, 2012). Considering the complexity in NLR-encoding gene structure and genomic organization, the assembly of NLR loci, including their intra- and intergenic regions, based on Sanger or short read sequencing approaches, with a resolution for accurately distinguishing recently duplicated genes, has often been prone to inaccuracy (Giolai et al., 2016; Marone, Russo, Laidò, De Leonardis, & Mastrangelo, 2013).

Target enrichment enables a selected subset of a genome to be captured and then sequenced, with reduced genome size and complexity enabling increased sequencing depth and accurate annotation. With considerable advances in the last 15 years since the first target-enrichment technologies for the selective capture and sequencing of genomic regions of interest (Bashiardes et al., 2005), strategies for NLR characterization in genomes have recently also been developed. Such approaches are necessary for both exhaustive sequencing and accurate annotation of the large NLR-encoding gene family complement in crop plant genomes. R-gene enrichment sequencing (RenSeq) was developed as a method for reduced complexity selective sequencing of NLR sequences in genomic DNA (Jupe et al., 2013). This approach relies upon the use of short (60 to 120-mer) oligo biotinylated overlapping RNA probes, or baits, designed as complementary for target NLR loci that enable bait capture and enrichment of homologous target DNA sequences for subsequent high-throughput sequencing (Fig. 1). Partial reference gene models from a single genotype of a plant species have been shown to be sufficient for bait design, with resequencing of target captured DNA fragments enabling considerable improvement in reannotation of NLR-encoding gene complements from previously assembled genomes, as well as novel NLR-encoding gene discovery in previously uncharacterized related resistant wild species or resistant genotypes within a species (Table 1). Initially developed and tested using available short read sequencing technologies (Jupe et al., 2012, 2013), accurate identification of NLR-encoding gene models in *Solanum tuberosum* increased numbers from 438 to 755, highlighting the efficiency in RenSeq as a target enrichment approach for this gene family. Although improvement in characterization of the NLR complement can be achieved by this approach, short read sequencing technologies such as Illumina continue to limit accurate assembly. To resolve these limitations, RenSeq has also been adapted successfully for long read single molecule sequencing of DNA fragments, up to 7 kb in size (Giolai et al., 2016). Using single-molecule real-time (SMRT) sequencing by PacBio, accurate long read sequencing is achieved as a result of the multiple sequencing of template captured DNA that the technology provides. Witek et al. (2016) applied SMRT Renseq using PacBio to isolate *Rpo-amr3i* in a wild relative of potato, an NLR-encoding gene conferring resistance to the important potato late blight pathogen *Phytophthora infestans*. Longer insert sizes employed in RenSeq have also been successfully employed in the sequencing of intergenic sequences for NLR promoter and terminator sequences, with 7 kb libraries efficient in the case of PacBio SMRT sequencing of NLR *R* gene libraries in *Solanum verrucosum* (Giolai et al., 2016).

PanNLRome

In addition to enabling accurate assembly of sequence data for NLR loci, RenSeq can be applied to exhaustively characterize the intraspecific NLR diversity in ecotypes of a particular plant species (panNLRome). Gene polymorphism, even at this taxonomic level, is known to be considerable as a result of the co-evolution with pathogen effectors, manifest as variation in NLR presence/absence, copy number (Leister et al., 1998; Michelmore & Meyers, 1998), and NLR sequence and integrated domain polymorphisms, including haplotype differences at NLR loci (Baurens et al., 2010). In order to exhaustively characterize the NLR diversity in a particular plant species, Van De Weyer et al. (2019) reported, in a first proof of concept study, a near complete panNLRome for *A. thaliana*, employing SMRT RenSeq across a curated diversity panel of 64 accessions. Employing RNA baits designed for over 700 potential NLR-encoding gene targets across the Brassicaceae, each accession harboured between 167 and 251 NLR-encoding genes, with the panNLRome comprising in excess of 13,000 NLR-encoding genes in total, grouped into the classes TNL, CNL, RNL and NL. With a total of 36 integrated domains observed, NLR domain architecture also varied considerably, with a total of 97 distinct architectures identified across all accessions, compared with only 22 annotated in the Col-0 reference genome. Analysis of genome organization also revealed common NLR pairing, with up to 17 NLR pairs per accession. Tajima's D values of paired NLRs provided evidence for co-evolution, with mutation in one gene leading to compensatory changes in the other pair member (Van De Weyer et al., 2019).

Table 1 Characterization of NLR-encoding genes across different plant species following whole genome sequencing or RenSeq approaches.

Species	Genome Size (Mbp)	Gene total	NLR number (WGS)	NLR number (RenSeq)	NLR sub-class			NLR annotation Method	NGS Platform	Reference
					CNL	TNL	Partial			
Solanum tuberosum	844	39,031	438	755	584	157	–	NP, MS, B	Illumina	(Jupe et al., 2013; Xu et al., 2011)
S. lycopersicum	900	34,727	260	326	195	26	102	NP, MS, B, BW, BE, P, ME, CW	Illumina	(Andolfo et al., 2014; Jupe et al., 2013)
S. pimpinellifolium	811	25,970	229	355	110	14	231	NP, MS, B, BW, BE, P, ME, CW	Illumina	(Andolfo et al., 2014; Jupe et al., 2013)
S. pennelli	942	20,076	486[a]	220	164	39	17	NP, MS, MU, B, BW, H	Illumina	(Bolger et al., 2014; Stam, Scheikl, & Tellier, 2016)
S. americanum	ND	ND	ND	646	330	71	245	NP, MS, B, BW	Illumina	(Jupe et al., 2013; Steuernagel et al., 2016)
Fragaria × ananassa	813.4	108,087	325	975	192	432	351	H, P, CL, BG, MU, B	PacBio Illumina	(Barbey et al., 2019; Edger et al., 2019; Zhong, Zhang, & Cheng, 2018)
F. vesca	240	34,809	144	367	54	170	143	H, P, CL, BG, MU, B	Illumina	(Barbey et al., 2019; Edger et al., 2019; Zhong et al., 2018)
F. iinumae	221	27 560	155	387	134	152	101	H, P, CL, BG, MU, B	Illumina	(Zhong et al., 2018; Barbey et al., 2019; Hirakawa et al., 2014)
Arabidopsis thaliana	125	25,498	201	464	269	102	93	H, P, CL, BG, MU, B, P, IPS, ME, CO, PC	PacBio	(Cao et al., 2011; Van De Weyer et al., 2019)

B, BLAST; CW, ClustalW; H, HMM; IPS, InterProScan; ME, MEGA; MU, MUSCLE; P, Pfam; CL, CLC Bio; BG, BLAST2GO; MS, MAST; NP, NLR PARSER; BW, BWA; BE, Bowtie; CO, Coils; PC, Paircoil2; WGS, whole genome sequence.

[a] Analyses considered all proteins with any domain associated with canonical NLRs (NB-ARC, LRR, CC, TIR or other domain).

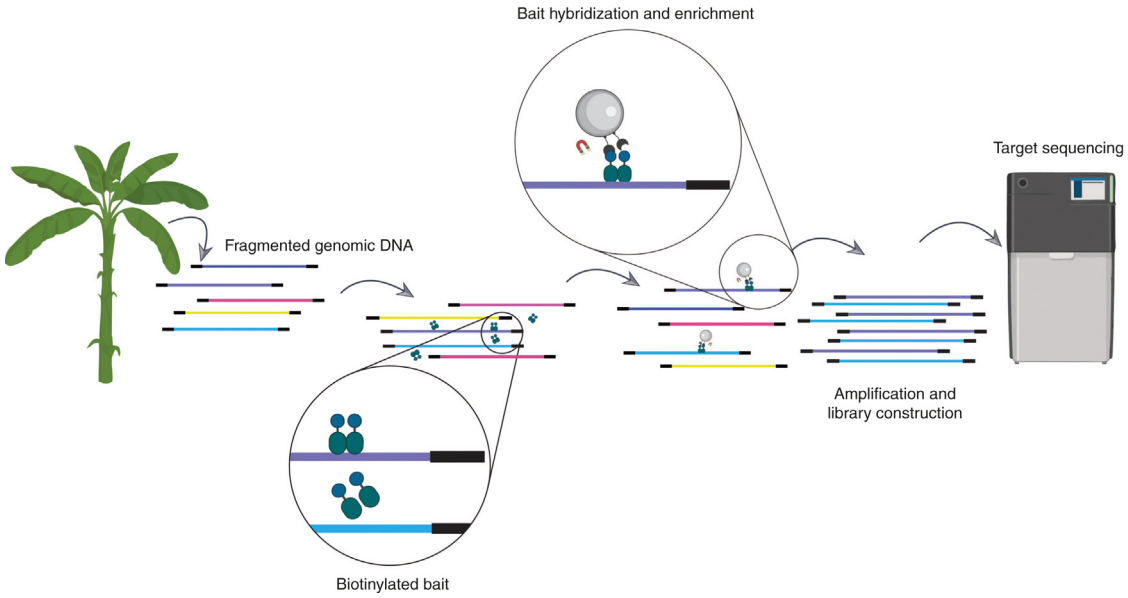

Figure 1 Overview of the RenSeq method combined with single molecule real-time (SMRT) sequencing for capture, enrichment and full-length sequencing of NLR-encoding genes.

Such exhaustive investigation enables core NLR-encoding gene complement identification, analysis of how the genes expand or contract according to pathogen selective pressures (Monteiro & Nishimura, 2018), as well as the characterization of novel NLRs and their architecture, associated with intraspecific diversity in plant populations from specific geographic regions such as diversity centres of origin.

Applications of RenSeq for *R* gene function validation

MutRenSeq

Mutagenesis Resistance gene enrichment and Sequencing (MutRenSeq) was recently developed to combine RenSeq with mutagenesis as an approach for identification of NLR resistance genes. Here, resistant wild-type plants are firstly subjected to mutagenesis in order to generate loss-of-resistance mutants. Target capture and sequencing of both resistant wild types and these loss-of-function mutants then enables identification of mutations in NLR-encoding genes that are responsible for the loss of resistance. This approach has recently been developed and applied in the identification of the *R* genes *Sr22* and *Sr45* that confer resistance the stem rust wheat pathogen *Puccinia graminis* f. sp. *tritici* (Steuernagel et al., 2016), two genes which previous positional cloning approaches were unable to isolate.

AgRenSeq

AgRenSeq is a further recent development of RenSeq for NLR resistance gene identification. Genome-wide association mapping (GWAS) has been widely employed as a genetic approach for linking genotype to phenotype, with trait correlation identification exploiting the accumulated recombination events present in natural populations. As this approach requires a reference genome, however, accurate identification of sequence variants which have significantly diverged from the reference, or absent in the reference, such as *R* genes, can limit this approach. One way to overcome such limitations can be through analysis of trait associations using variable length K-mer sub-sequences, as observed, for example, with bacterial genomes (Lees et al., 2016). Successfully combining association genetics with RenSeq, *R* genes have now been isolated from a diversity panel of plant accessions. In a proof of concept study, Arora et al. (2019) evaluated AgRenSeq against a diversity panel of *Aegilops tauschii* ssp. *strangulata* accessions. As a wild species and progenitor of the D genome in *Triticum aestivum*, this species harbours resistance to wheat stem rust and has served as a source of resistance alleles that have been introgressed into commercial wheat. In this approach, bait capture and high-throughput sequencing of target NLR sequences was conducted on a disease resistance-phenotyped diversity panel, with positive NLR *k*-mer sequences filtered according to the correlation between their presence and level of resistance in the panel. The study successfully identified a total of four *Sr R* genes, namely *Sr33*, *Sr45* (which was previously identified by MutRenSeq (Steuernagel et al., 2016), *Sr46* and *SrTA1662*. This speed cloning approach can efficiently exploit the natural sequence variation present in wild relative species for *R* gene isolation and subsequent introgression into elite materials, without dependence upon the development of mapping populations or mutagenized materials.

dRenSeq

dRenSeq is a further application of RenSeq as a diagnostic tool that enables the presence or absence of known reference NLR-encoding *R* genes to be screened for based on stringent mapping conditions in resistant, susceptible and

bulk materials for plant accessions contrasting in disease resistance (Van Weymers et al., 2016). Through a standard RenSeq target enrichment and Illumina MiSeq sequencing approach, followed by stringent Bowtie2 (Langmead & Salzberg, 2012) mapping to known *R* genes, this group confirmed the presence of the potato late blight resistance gene *Rpi-vnt1.1* in *Solanum okadae* accessions, highlighting the applicability of the approach for characterization of germplasm collections harbouring wild accessions.

Engineering plant immunity with NLRs

Traditionally, plant genetic improvement has focused on the introgression of resistance alleles into elite varieties. Such an approach, however, can allow mutations that occur in effectors in rapidly evolving pathogens to overcome such monogenic resistance in typical monoculture cropping systems.

Research on the plant immune system in plants during the last 30 years has resulted in the identification of a relatively limited number of specific NLRs responsible for effector recognition and ETI-based disease resistance, with 191 functional NLRs or NLR-IDs now cloned from all analysed plant taxa to date (Goodstein et al., 2011; Kourelis & van der Hoorn, 2018).

RenSeq, as a recently developed technology, is now rapidly expanding the repertoire of potential NLR-encoding *R* genes to different biotic stresses. With this more exhaustive identification of the NLR complement across plant taxa, the identification of cognate pathogen effector proteins is now a major bottleneck, requiring high-throughput protein-protein interaction analysis through *in planta* co-expression (Zhang, Dodds et al., 2017) or yeast two-hybrid approaches. With such data, crop improvement programs look set to be able to develop breeding lines harbouring not only single NLR-encoding *R* gene receptors, potentially including those transferred even between distantly related taxa, but also stacked NLRs from the resistance gene pool, respecting potentially necessary sensor and helper pair formats, to confer resistance to single and multiple pathogens. Appropriate deployment in the field in different combinations will likely increase durability of resistance in comparison to single loci (Bentham et al., 2017; Dangl, Horvath, & Staskawicz, 2013; Steuernagel et al., 2016). Restricted taxonomic functionality may, however, be a factor limiting utilization of *R* genes between distant plant taxa (Tai et al., 1999) especially in the context of compatibility limitations between sensor and helper NLRs from different genetic backgrounds. Appropriate regulation of inserted NLRs is also essential to avoid autoactivity and fitness reduction. Nonetheless, examples of successful NLR transfer between distantly related species have been observed, with, for example, the CNL *RXO1* from maize conferring resistance to *Xanthomonas oryzae* pv. *oryzicola* in rice (Zhao et al., 2005), and the CNL MLA1 from barley into Arabidopsis (Lu et al., 2016). Examples of stacked NLR-encoding genes have included Pi2 and Pi54 in rice, conferring resistance to *Magnaporthe oryzae* (Ellur et al., 2016), as well as Rpi-sto1, Rpi-vnt1.1 and Rpi-blb3 in potato, for control of Late blight (Zhu, Li, Vossen, Visser, & Jacobsen, 2012).

Genetic modification has so far been employed on only a limited number of individual NLR-encoding genes, with focus on understanding molecular mechanisms underlying NLR protein response to pathogen effectors, as well as a means of engineering altered specificity towards pathogen effectors. Random mutagenesis on the LRR domain of the potato *Rx* CNL gene, for example, has been successfully employed to increase recognition of multiple *Potato Virus X* strains (Farnham & Baulcombe, 2006). Although mutations in the NBS domain can result in autoactivity (Takken & Goverse, 2012), gain-of-function mutations in the NBS domain have been reported for the tomato *NRC1* gene, a gene required in intracellular signaling in resistance response to *Cladosporium fulvum* and *Verticillium dahliae* (Sueldo et al., 2015). Segretin et al. (2014) expanded effector recognition in the potato NLR-encoding gene R3a to AVR3a variants in *Phytophthora infestans* through the same overall approach, modifying only a few amino acids in the LRR domain. Further application has been conducted on I2, an ortholog of R3a in tomato that confers resistance to *Fusarium oxysporum* f. sp. *lycopersici* (Giannakopoulou et al., 2015). Remarkably, a mutant in the N-terminal CC domain, $I2^{I141N}$, displayed an increased response to AVR3a, with partial resistance to *P. infestans* and an expanded response to different *F. oxysporum* f. sp. *lycopersici* effectors. This work highlights the potential in engineering synthetic NLRs, based on known NLR *R* genes, for enhanced effector recognition and towards broadening recognition of diverse pathogens. Given the importance now recognized of NLR-IDs in effector interaction, these also represent attractive targets for engineering. De La Concepcion et al. (2019) recently employed a structure-guided protein engineering approach for NLRs, with mutation within the Pikp−HMA ID expanding the recognition profile of the rice NLR Pikp to variants of the rice blast pathogen effector AVR-Pik.

Given the recently developed site specific mutagenesis genome editing methods, such as Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) / Cas9 (CRISPR-associated protein 9), which has now been widely employed in targeted gene modification in plants (Gaj, Gersbach, & Barbas, 2013; Horvath & Barrangou, 2010; Puchta, 2017), both validation of function of candidate *R* gene alleles as well as rational genome engineering for resistance are now feasible with this approach. Work focusing on precise modification of specific nucleotide sequences conferring resistance and specificity in NLR-encoding genes, together with the deletion of larger genomic regions (Wu et al., 2018), such as promoters or tandemly repeated genes, as well as truncation of ORFs for gain of function in NLRs (e.g. Weaver, Swiderski, Li, & Jones, 2006), will undoubtedly advance crop improvement for disease resistance.

Conclusions

Given the current constraints to sustainable agricultural production, with increasing crop losses due to pre-harvest diseases and climate change, considerable advances are required in crop improvement approaches for enabling durable disease resistance. In this context, NLRs provide genetic solutions. Having been widely applied in crop breeding programs, with introgression generally enabling

monogenic resistance, a more complete understanding of their structure and function in plant immunity is opening up possibilities for broader, polygenic stable disease resistance development. Whilst significant advances have recently been achieved into targeted enrichment for their exhaustive characterization in plant genomes, together with the recognition of distinct roles in pathogen sensing or intracellular signaling, how different NLR protein pairs or networks actually recognize pathogen effector proteins and cooperate at the molecular level still remains largely unresolved. Once such mechanisms are recognized, synthetic biology approaches focusing on gene design, domain motif editing and subsequent introgression into elite materials will contribute significantly towards the development of broad and stable resistance to pathogens.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

This work was financially supported by CAPES (grant number 88887.310583/2018-00), INCT (grant number 465480/2014-4) and FAPDF (grant number 0193.001725/2017). RNGM was supported by a fellowship from CNPq (grant number 305418/2017-3).

References

Aarts, M. G. M., Hekkert, B. T. L., Holub, E. B., Beynon, J. L., Stiekema, W. J., & Pereira, A. (1998). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, *11*(4), 251–258.

Adachi, H., Contreras, M., Harant, A., Wu, C. H., Derevnina, L., Sakai, T., et al. (2019). An N-terminal motif in NLR immune receptors is functionally conserved across distantly related plant species. *Elife*, *8*, e49956.

Adachi, H., Derevnina, L., & Kamoun, S. (2019). NLR singletons, pairs, and networks: Evolution, assembly, and regulation of the intracellular immunoreceptor circuitry of plants. *Current Opinion in Plant Biology*, *50*, 121–131.

Ade, J., DeYoung, B. J., Golstein, C., & Innes, R. W. (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(7), 2531–2536.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402.

Ameline-Torregrosa, C., Wang, B. B., O'Bleness, M. S., Deshpande, S., Zhu, H., Roe, B., et al. (2008). Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiology*, *146*(1), 5–21.

Andolfo, G., Di Donato, A., Chiaiese, P., De Natale, A., Pollio, A., Jones, J. D. G., et al. (2019). Alien domains shaped the modular structure of plant NLR proteins. *Genome Biology and Evolution*, *11*(12), 3466–3477.

Andolfo, G., Jupe, F., Witek, K., Etherington, G. J., Ercolano, M. R., & Jones, J. D. G. (2014). Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC Plant Biology*, *14*(1), 1–12.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., et al. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology*, *37*(2), 139–143.

Ashikawa, I., Hayashi, N., Yamane, H., Kanamori, H., Wu, J., Matsumoto, T., et al. (2008). Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. *Genetics*, *180*(4), 2267–2276.

Azhar, M., & Heslop-Harrison, J. S. (2008). Genomes, diversity and resistance gene analogues in *Musa* species. *Cytogenetic and Genome Research*, *121*(1), 59–66.

Baggs, E., Dagdas, G., & Krasileva, K. V. (2017). NLR diversity, helpers and integrated domains: Making sense of the NLR IDentity. *Current Opinion in Plant Biology*, *38*, 59–67.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Research*, *37*(suppl_2), W202–W208.

Baldi, P., Patocchi, A., Zini, E., Toller, C., Velasco, R., & Komjanc, M. (2004). Cloning and linkage mapping of resistance gene homologues in apple. *Theoretical and Applied Genetics*, *109*(1), 231–239.

Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, *6*(1), 11.

Barbey, C. R., Lee, S., Verma, S., Bird, K. A., Yocca, A. E., Edger, P. P., et al. (2019). Disease resistance genetics and genomics in octoploid strawberry. *G3 (Bethesda, Md)*, *9*(10), 3315–3332.

Bashiardes, S., Veile, R., Helms, C., Mardis, E. R., Bowcock, A. M., & Lovett, M. (2005). Direct genomic selection. *Nature Methods*, *2*(1), 63.

Baurens, F. C., Bocs, S., Rouard, M., Matsumoto, T., Miller, R. N. G., Rodier-Goud, M., et al. (2010). Mechanisms of haplotype divergence at the RGA08 nucleotide-binding leucine-rich repeat gene locus in wild banana (*Musa balbisiana*). *BMC Plant Biology*, *10*(1), 149.

Bayer, P. E., Edwards, D., & Batley, J. (2018). Bias in resistance gene prediction due to repeat masking. *Nature Plants*, *4*, 762–765.

Bella, J., Hindle, K. L., McEwan, P. A., & Lovell, S. C. (2008). The leucine-rich repeat structure. *Cellular and Molecular Life Sciences*, *65*, 2307–2333.

Bent, A. F., Kunkel, B. N., Dahlbeck, D., Brown, K. L., Schmidt, R., Giraudat, J., et al. (1994). RPS2 of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. *Science*, *265*(5180), 1856–1860.

Bentham, A., Burdett, H., Anderson, P. A., Williams, S. J., & Kobe, B. (2017). Animal NLRs provide structural insights into plant NLR function. *Annals of Botany*, *119*, 689–702.

Bernoux, M., Burdett, H., Williams, S. J., Zhang, X., Chen, C., Newell, K., et al. (2016). Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *The Plant Cell*, *28*(1), 146–159.

Bertioli, D. J., Leal-Bertioli, S. C. M., Lion, M. B., Santos, V. L., Pappas, G., Cannon, S. B., et al. (2003). A large scale analysis of resistance gene homologues in Arachis. *Molecular Genetics and Genomics*, *270*(1), 34–45.

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., et al. (2014). The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics*, *46*(9), 1034–1038.

Boller, T., & Felix, G. (2009). A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology*, *60*(1), 379–406.

Bonardi, V., Tang, S., Stallmann, A., Roberts, M., Cherkis, K., & Dangl, J. L. (2011). Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of

pathogen effectors. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(39), 16463–16468.
Borrelli, G. M., Mazzucotelli, E., Marone, D., Crosatti, C., Michelotti, V., Valè, G., et al. (2018). Regulation and evolution of NLR genes: A close interconnection for plant immunity. *International Journal of Molecular Sciences*, *19*(6), 1662.
Boutrot, F., & Zipfel, C. (2017). Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annual Review of Phytopathology*, *55*(1), 257–286.
Bozkurt, O., Hakki, E. E., & Akkaya, M. S. (2007). Isolation and sequence analysis of wheat NBS-LRR type disease resistance gene analogs using degenerate PCR primers. *Biochemical Genetics*, *45*(5–6), 469–486.
Cannon, S. B., Zhu, H., Baumgarten, A. M., Spangler, R., May, G., Cook, D. R., et al. (2002). Diversity, distribution, and ancient taxonomic relationships within the TIR and Non-TIR NBS-LRR resistance gene subfamilies. *Journal of Molecular Evolution*, *54*(4), 548–562.
Cao, J., Schneeberger, K., Ossowski, S., Günther, T., Bender, S., Fitz, J., et al. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics*, *43*(10), 956–965.
Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., & Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: The ''integrated decoy'' hypothesis. *Frontiers in Plant Science*, *5*, 606.
Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., et al. (2013). The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. *The Plant Cell*, *25*(4), 1463–1481.
Chen, Q., Han, Z., Jiang, H., Tian, D., & Yang, S. (2010). Strong positive selection drives rapid diversification of R-Genes in *Arabidopsis* relatives. *Journal of Molecular Evolution*, *70*(2), 137–148.
Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell*, *124*, 803–814.
Collier, S. M., Hamel, L. P., & Moffett, P. (2011). Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Molecular Plant-Microbe Interactions*, *24*(8), 918–931.
Collins, N. C., Webb, C. A., Seah, S., Ellis, J. G., Hulbert, S. H., & Pryor, A. (1998). The isolation and mapping of disease resistance gene analogs in maize. *Molecular Plant-Microbe Interactions*, *11*(10), 968–978.
Creelman, R. A., & Mullet, J. E. (1995). Jasmonic acid distribution and action in plants: Regulation during development and response to biotic and abiotic stress. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(10), 4114–4119.
D'hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, *488*(7410), 213–217.
Dangl, J. L. (1994). The enigmatic avirulence genes of phytopathogenic bacteria. *Current Topics in Microbiology and Immunology*, *192*, 99–118.
Dangl, J. L., & Jones, J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, *411*(6839), 826–833.
Dangl, J. L., Horvath, D. M., & Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science*, *341*(6147), 746–751.
De La Concepcion, J. C., Franceschetti, M., MacLean, D., Terauchi, R., Kamoun, S., & Banfield, M. J. (2019). Protein engineering expands the effector recognition profile of a rice NLR immune receptor. *eLife*, *8*, e47713.
Delorenzi, M., & Speed, T. (2002). An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics*, *18*(4), 617–625.
DeYoung, B. J., & Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, *7*, 1243–1249.
Dodds, P. N., & Rathjen, J. P. (2010). Plant immunity: Towards an integrated view of plant–Pathogen interactions. *Nature Reviews Genetics*, *11*, 539–548. http://dx.doi.org/10.1038/nrg2812
Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A., et al. (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8888–8893.
Dong, X. (2001). Genetic dissection of systemic acquired resistance. *Current Opinion in Plant Biology*, *4*, 309–314.
Dubey, N., & Singh, K. (2018). Role of NBS-LRR proteins in plant defense. In *Molecular aspects of Plant-Pathogen interaction*. pp. 115–138.
Eddy, S. R., & Wheeler, T. (2007). *HMMER-biosequence analysis using profile hidden Markov models URL Http://Hmmer. Janelia. Org*.
Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nature Genetics*, *51*(3), 541–547.
Ellis, J., & Jones, D. (1998). Structure and function of proteins controlling strain-specific pathogen resistance in plants. *Current Opinion in Plant Biology*, *1*(4), 288–293.
Ellis, J. G., Lawrence, G. J., Luck, J. E., & Dodds, P. N. (1999). Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *The Plant Cell*, *11*(3), 495–506.
Ellis, J., Dodds, P., & Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, *3*(4), 278–284.
Ellur, R. K., Khanna, A., Yadav, A., Pathania, S., Rajashekara, H., Singh, V. K., et al. (2016). Improvement of Basmati rice varieties for resistance to blast and bacterial blight diseases using marker assisted backcross breeding. *Plant Science*, *242*, 330–341.
Farnham, G., & Baulcombe, D. C. (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proceedings of the National Academy of Sciences*, *103*(49), 18828–18833.
Faustin, B., Lartigue, L., Bruey, J. M., Luciano, F., Sergienko, E., Bailly-Maitre, B., et al. (2007). Reconstituted NALP1 inflammasome reveals two-step mechanism of Caspase-1 activation. *Molecular Cell*, *25*(5), 713–724.
Felix, G., Duran, J. D., Volko, S., & Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal*, *18*(3), 265–276.
Flor, H. H. (1942). Inheritance of pathogenicity in *Melampsora lini*. *Phytopathology*, *32*, 653–669.
Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, *31*(7), 397–405.
Gao, Y., Wang, W., Zhang, T., Gong, Z., Zhao, H., & Han, G. Z. (2018). Out of water: The origin and early diversification of plant R-genes. *Plant Physiology*, *177*(1), 82–89.
Giannakopoulou, A., Steele, J. F. C., Segretin, M. E., Bozkurt, T. O., Zhou, J., Robatzek, S., et al. (2015). Tomato I2 Immune Receptor Can Be Engineered to Confer Partial Resistance to the Oomycete *Phytophthora infestans*in Addition to the Fungus Fusarium oxysporum. *Molecular Plant-Microbe Interactions*, *28*(12), 1316–1329.
Giolai, M., Paajanen, P., Verweij, W., Percival-Alwyn, L., Baker, D., Witek, K., et al. (2016). Targeted capture and sequencing of gene-sized DNA molecules. *BioTechniques*, *61*(6), 315–322.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2011). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(D1), D1178–D1186.

Gu, L., Si, W., Zhao, L., Yang, S., & Zhang, X. (2015). Dynamic evolution of NBS–LRR genes in bread wheat and its progenitors. *Molecular Genetics and Genomics*, *290*(2), 727–738.

Guo, Y. L., Fitz, J., Schneeberger, K., Ossowski, S., Cao, J., & Weigel, D. (2011). Genome-wide comparison of nucleotide-binding site-leucine-rich repeat-encoding genes in Arabidopsis. *Plant Physiology*, *157*(2), 757–769.

Gururani, M. A., Venkatesh, J., Upadhyaya, C. P., Nookaraju, A., Pandey, S. K., & Park, S. W. (2012). Plant disease resistance genes: Current status and future directions. *Physiological and Molecular Plant Pathology*, *78*, 51–65.

Hammond-Kosack, K. E., & Jones, J. D. (1996). Resistance gene-dependent plant defense responses. *The Plant Cell*, *8*(10), 1773–1791.

He, X., Anderson, J. C., Del Pozo, O., Gu, Y. Q., Tang, X., & Martin, G. B. (2004). Silencing of subfamily I of protein phosphatase 2A catalytic subunits results in activation of plant defense responses and localized cell death. *The Plant Journal*, *38*(4), 563–577.

Heath, M. C. (2000). Hypersensitive response-related death. *Plant Molecular Biology*, *44*(3), 321–334.

Hirakawa, H., Shirasawa, K., Kosugi, S., Tashiro, K., Nakayama, S., Yamada, M., et al. (2014). Dissection of the octoploid strawberry genome by deep sequencing of the genomes of *Fragaria* species. *DNA Research*, *21*(2), 169–181.

Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science*, *327*(5962), 167–170.

Hulbert, S. H., Webb, C. A., Smith, S. M., & Sun, Q. (2001). Resistance gene complexes: Evolution and utilization. *Annual Review of Phytopathology*, *39*(1), 285–312.

Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P., & Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO Journal*, *19*(15), 4004–4014.

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329.

Jones, D. A., & Jones, J. D. G. (1997). The role of leucine-rich repeat proteins in plant defences. *Advances in Botanical Research*, *24*, 89–167.

Jupe, F., Pritchard, L., Etherington, G. J., MacKenzie, K., Cock, P. J. A., Wright, F., et al. (2012). Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics*, *13*(1), 75.

Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., et al. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal*, *76*(3), 530–544.

Kajava, A. V. (1998). Structural diversity of leucine-rich repeat proteins. *Journal of Molecular Biology*, *277*, 519–527.

Käll, L., Krogh, A., & Sonnhammer, E. L. L. (2004). A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*, *338*(5), 1027–1036.

Kanazin, V., Marek, L. F., & Shoemaker, R. C. (1996). Resistance gene analogs are conserved and clustered in soybean. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(21).

Kim, J., Lim, C. J., Lee, B. W., Choi, J. P., Oh, S. K., Ahmad, R., et al. (2012). A genome-wide comparison of NB-LRR type of resistance gene analogs (RGA) in the plant Kingdom. *Molecules and Cells*, *33*(4), 385–392.

Kim, S., Park, J., Yeom, S. I., Kim, Y. M., Seo, E., Kim, K. T., et al. (2017). New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. *Genome Biology*, *18*(1), 210.

Kobe, B., & Deisenhofer, J. (1994). The leucine-rich repeat: A versatile binding motif. *Trends in Biochemical Sciences*, *19*, 415–421.

Kourelis, J., & van der Hoorn, R. A. L. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for r protein function. *The Plant Cell*, *30*, 285–299.

Krasileva, K. V., Dahlbeck, D., & Staskawicz, B. J. (2010). Activation of an *Arabidopsis* resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *The Plant Cell*, *22*(7), 2444–2458.

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567–580.

Kroj, T., Chanclud, E., Michel-Romiti, C., Grand, X., & Morel, J. B. (2016). Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *The New Phytologist*, *210*(2), 618–626.

Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., & Felix, G. (2004). The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *The Plant Cell*, *16*(12), 3496–3507.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357.

Lees, J. A., Vehkala, M., Välimäki, N., Harris, S. R., Chewapreecha, C., Croucher, N. J., et al. (2016). Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. *Nature Communications*, *7*, 12797.

Leister, R. T., & Katagiri, F. (2000). A resistance gene product of the nucleotide binding site - Leucine rich repeats class can form a complex with bacterial avirulence proteins in vivo. *The Plant Journal*, *22*(4), 345–354.

Leister, D., Ballvora, A., Salamini, F., & Gebhardt, C. (1996). A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nature Genetics*, *14*(4), 421–429.

Leister, D., Kurth, J., Laurie, D. A., Yano, M., Sasaki, T., Devos, K., et al. (1998). Rapid reorganization of resistance gene homologues in cereal genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(1), 370–375.

Lozano-Torres, J. L., Wilbers, R. H. P., Gawronski, P., Boshoven, J. C., Finkers-Tomczak, A., Cordewener, J. H. G., et al. (2012). Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(25), 10119–10124.

Lu, X., Kracher, B., Saur, I. M. L., Bauer, S., Ellwood, S. R., Wise, R., et al. (2016). Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. *Proceedings of the National Academy of Sciences*, *113*(42), E6486–E6495.

Lucas, J. A., Hawkins, N. J., & Fraaije, B. A. (2015). The evolution of fungicide resistance. *Advances in Applied Microbiology*, *90*, 29–92.

Lupas, A., Van Dyke, M., & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, *252*(5009), 1162–1164.

Maekawa, T., Kufer, T. A., & Schulze-Lefert, P. (2011). NLR functions in plant and animal immune systems: So far and yet so close. *Nature Immunology*, *12*, 818–826.

Mago, R., Nair, S., & Mohan, M. (1999). Resistance gene analogues from rice: Cloning, sequencing and mapping. *Theoretical and Applied Genetics*, *99*(1–2), 50–57.

Maqbool, A., Saitoh, H., Franceschetti, M., Stevenson, C. E. M., Uemura, A., Kanzaki, H., et al. (2015). Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *Elife*, *4*, e08709.

Marone, D., Russo, M. A., Laidò, G., De Leonardis, A. M., & Mastrangelo, A. M. (2013). Plant nucleotide binding site–Leucine-rich repeat (NBS-LRR) genes: Active guardians in

host defense responses. *International Journal of Molecular Sciences*, *14*(4), 7302–7326.
McDonnell, A. V, Jiang, T., Keating, A. E., & Berger, B. (2005). Paircoil2: Improved prediction of coiled coils from sequence. *Bioinformatics*, *22*(3), 356–358.
Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W., & Young, N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal*, *20*(3), 317–332.
Meyers, B. C., Kozik, A., Griego, A., Kuang, H., & Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *The Plant Cell*, *15*(4), 809–834.
Meyers, B. C., Morgante, M., & Michelmore, R. W. (2002). TIR-X and TIR-NBS proteins: Two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. *The Plant Journal*, *32*(1), 77–92.
Michelmore, R. W., & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, *8*(11), 1113–1130.
Miller, R. N. G., Bertioli, D. J., Baurens, F. C., Santos, C., Alves, P. C., Martins, N. F., et al. (2008). Analysis of non-TIR NBS-LRR resistance gene analogs in *Musa acuminata* Colla: Isolation, RFLP marker development, and physical mapping. *BMC Plant Biology*, *15*, 1–15.
Miller, R. N. G., Alves, G. S. C., & Van Sluys, M. A. (2017). Plant immunity: Unravelling the complexity of plant responses to biotic stresses. *Annals of Botany*, *119*(5), 681–687.
Milligan, S. B., Bodeau, J., Yaghoobi, J., Kaloshian, I., Zabel, P., & Williamson, V. M. (1998). The root knot nematode resistance gene Mi from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. *The Plant Cell*, *10*(8), 1307–1319.
Moffett, P., Farnham, G., Peart, J., & Baulcombe, D. C. (2002). Interaction between domains of a plant NBS–LRR protein in disease resistance-related cell death. *The EMBO Journal*, *21*(17), 4511–4519.
Monaghan, J., & Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*, *15*, 349–357.
Monteiro, F., & Nishimura, M. T. (2018). Structural, functional, and genomic diversity of plant NLR proteins: An evolved resource for rational engineering of plant immunity. *Annual Review of Phytopathology*, *56*(1), 243–267.
Nishimura, M. T., Monteiro, F., & Dangl, J. L. (2015). Treasure your exceptions: Unusual domains in immune receptors reveal host virulence targets. *Cell*, *161*, 957–960.
Peñuela, S., Danesh, D., & Young, N. D. (2002). Targeted isolation, sequence analysis, and physical mapping of nonTIR NBS-LRR genes in soybean. *Theoretical and Applied Genetics*, *104*(2–3), 261–272.
Porter, B. W., Paidi, M., Ming, R., Alam, M., Nishijima, W. T., & Zhu, Y. J. (2009). Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. *Molecular Genetics and Genomics*, *281*(6), 609–626.
Puchta, H. (2017). Applying CRISPR/Cas for genome engineering in plants: The best is yet to come. *Current Opinion in Plant Biology*, *36*, 1–8.
Rairdan, G. J., & Moffett, P. (2006). Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *The Plant Cell*, *18*(8), 2082–2093.
Ravensdale, M., Bernoux, M., Ve, T., Kobe, B., Thrall, P. H., Ellis, J. G., et al. (2012). Intramolecular interaction influences binding of the Flax L5 and L6 resistance proteins to their AvrL567 ligands. *PLoS Pathogens*, *8*(11), e1003004.
Richter, T. E., & Ronald, P. C. (2000). The evolution of disease resistance genes. *Plant Molecular Biology*, *42*(1), 195–204.
Sanseverino, W., Roma, G., De Simone, M., Faino, L., Melito, S., Stupka, E., et al. (2009). PRGdb: A bioinformatics platform for plant resistance gene analysis. *Nucleic Acids Research*, *38*(suppl.1), D814–D821.
Sarris, P. F., Cevik, V., Dagdas, G., Jones, J. D. G., & Krasileva, K. V. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology*, *14*(1).
Saucet, S. B., Ma, Y., Sarris, P. F., Furzer, O. J., Sohn, K. H., & Jones, J. D. G. (2015). Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nature Communications*, *6*, 6338.
Segretin, M. E., Pais, M., Franceschetti, M., Chaparro-Garcia, A., Bos, J. I., Banfield, M. J., et al. (2014). Single amino acid mutations in the potato A, immune receptor R3a expand response to *Phytophthora* effectors. *Molecular Plant-Microbe Interactions*, *27*, 624–637.
Sekhwal, M. K., Li, P., Lam, I., Wang, X., Cloutier, S., & You, F. M. (2015). Disease resistance gene analogs (RGAs) in plants. *International Journal of Molecular Sciences*, *16*, 19248–19290.
Shao, Z. Q., Xue, J. Y., Wu, P., Zhang, Y. M., Wu, Y., Hang, Y. Y., et al. (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiology*, *170*(4), 2095–2109.
Sharma, T. R., Das, A., Kumar, S. P., & Lodha, M. L. (2009). Resistance gene analogues as a tool for rapid identification and cloning of disease resistance genes in plants - A review. *Journal of Plant Biochemistry and Biotechnology*, *18*, 01–11.
Shen, Q. H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K., & Schulze-Lefert, P. (2003). Recognition specificity and RAR1/SGT1 dependence in barley Mla disease resistance genes to the powdery mildew fungus. *The Plant Cell*, *15*(3), 732–744.
Speulman, E., Bouchez, D., Holub, E. B., & Beynon, J. L. (1998). Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis thaliana*. *The Plant Journal*, *14*(4), 467–474.
Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology*, *12*, 89–100.
Stam, R., Scheikl, D., & Tellier, A. (2016). Pooled enrichment sequencing identifies diversity and evolutionary pressures at NLR resistance genes within a wild tomato population. *Genome Biology and Evolution*, *8*(5), 1501–1515.
Stange, C., Matus, J. T., Domínguez, C., Perez-Acle, T., & Arce-Johnson, P. (2008). The N-homologue LRR domain adopts a folding which explains the TMV-Cg-induced HR-like response in sensitive tobacco plants. *Journal of Molecular Graphics & Modelling*, *26*(5), 850–860.
Steele, J. F. C., Hughes, R. K., & Banfield, M. J. (2019). Structural and biochemical studies of an NB-ARC domain from a plant NLR immune receptor. *PloS One*, *14*(8), e0221226.
Stein, J. C., Yu, Y., Copetti, D., Zwickl, D. J., Zhang, L., Zhang, C., et al. (2018). Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nature Genetics*, *50*(2), 285–296.
Steuernagel, B., Jupe, F., Witek, K., Jones, J. D. G., & Wulff, B. B. H. (2015). NLR-parser: Rapid annotation of plant NLR complements. *Bioinformatics*, *31*(10), 1665–1667.
Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology*, *34*(6), 652–655.
Steuernagel, B., Witek, K., Krattinger, S. G., Ramirez-Gonzalez, R. H., Schoonbeek, H., Yu, G., et al. (2018). Physical and transcriptional organisation of the bread wheat intracellular immune receptor repertoire. *BioRxiv*, 339424.
Sueldo, D. J., Shimels, M., Spiridon, L. N., Caldararu, O., Petrescu, A., Joosten, M. H. A. J., et al. (2015). Random mutage-

nesis of the nucleotide-binding domain of NRC 1 (NB-LRR required for Hypersensitive Response-Associated Cell Death-1), a downstream signalling nucleotide-binding, leucine-rich repeat (NB-LRR) protein, identifies gain-of-function mutations in the nucleotide-binding pocket. *The New Phytologist*, *208*(1), 210–223.

Swiderski, M. R., & Innes, R. W. (2001). The *Arabidopsis* PBS1 resistance gene encodes a member of a novel protein kinase subfamily. *The Plant Journal*, *26*(1), 101–112.

Tai, T. H., Dahlbeck, D., Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C., et al. (1999). Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proceedings of the National Academy of Sciences*, *96*(24), 14153–14158.

Takken, F. L. W., & Goverse, A. (2012). How to build a pathogen detector: Structural basis of NB-LRR function. *Current Opinion in Plant Biology*, *15*(4), 375–384.

Takken, F. L. W., & Tameling, W. I. L. (2009). To nibble at plant resistance proteins. *Science*, *324*, 744–746.

Takken, F. L., Albrecht, M., & Tameling, W. I. L. (2006). Resistance proteins: Molecular switches of plant defence. *Current Opinion in Plant Biology*, *9*, 383–390.

Tameling, W. I. L., Elzinga, S. D. J., Darmin, P. S., Vossen, J. H., Takken, F. L. W., Haring, M. A., et al. (2002). The tomato R gene products i-2 and Mi-1 are functional ATP binding proteins with ATPase activity. *The Plant Cell*, *14*(11), 2929–2939.

Thomma, B. P. H. J., Nürnberger, T., & Joosten, M. H. A. J. (2011). Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *The Plant Cell*, *23*(1), 4–15.

Van De Weyer, A. L., Monteiro, F., Furzer, O. J., Nishimura, M. T., Cevik, V., Witek, K, et al. (2019). A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell*, *178*(5), 1260–1272.

Van Der Biezen, E. A., & Jones, J. D. G. (1998a). The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology*, *8*(7), R226–R228.

Van Der Biezen, E. A., & Jones, J. D. G. (1998b). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, *23*(12), 454–456.

Van Der Hoorn, R. A. L., & Kamoun, S. (2008). From guard to decoy: A new model for perception of plant pathogen effectors. *The Plant Cell*, *20*(8), 2009–2017.

Van Der Vossen, E. A., Van Der Voort, J. N., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C., et al. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: A virus and a nematode. *The Plant Journal*, *23*(5), 567–576.

Van Esse, H. P., Van't Klooster, J. W., Bolton, M. D., Yadeta, K. A., Van Baarlen, P., Boeren, S., et al. (2008). The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense. *The Plant Cell*, *20*(7), 1948–1963.

Van Ghelder, C., Parent, G. J., Rigault, P., Prunier, J., Giguère, I., Caron, S., et al. (2019). The large repertoire of conifer NLR resistance genes includes drought responsive and highly diversified RNLs. *Scientific Reports*, *9*(1), 11614.

Van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. C., & Takken, F. L. W. (2008). Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*, *59*(6), 1383–1397.

Van Weymers, P. S. M., Baker, K., Chen, X., Harrower, B., Cooke, D. E. L., Gilroy, E. M., et al. (2016). Utilizing ''Omic'' technologies to identify and prioritize novel sources of resistance to the oomycete pathogen *Phytophthora infestans* in potato germplasm collections. *Frontiers in Plant Science*, *7*, 672.

Ve, T., Williams, S. J., & Kobe, B. (2015). Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis*, *20*(2), 250–261.

Wan, L., Essuman, K., Anderson, R. G., Sasaki, Y., Monteiro, F., Chung, E. H., et al. (2019). TIR domains of plant immune receptors are NAD+-cleaving enzymes that promote cell death. *Science*, *365*, 799–803.

Wang, C. I. A., Gunčar, G., Forwood, J. K., Teh, T., Catanzariti, A. M., Lawrence, G. J., et al. (2007). Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *The Plant Cell*, *19*(9), 2898–2912.

Wang, J., Hu, M., Wang, J., Qi, J., Han, Z., Wang, G., et al. (2019). Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science*, *364*(6435), eaav5870.

Weaver, L. M., Swiderski, M. R., Li, Y., & Jones, J. D. (2006). The *Arabidopsis thaliana* TIR-NB-LRR R-protein, RPP1A; protein localization and constitutive activation of defence by truncated alleles in tobacco and *Arabidopsis*. *The Plant Journal*, *47*(6), 829–840.

Whitham, S., Dinesh-Kumar, S. P., Choi, D., Hehl, R., Corr, C., & Baker, B. (1994). The product of the tobacco mosaic virus resistance gene N: Similarity to toll and the interleukin-1 receptor. *Cell*, *78*(6), 1101–1115.

Williams, S. J., Sornaraj, P., De Courcy-Ireland, E., Ian Menz, R., Kobe, B., Ellis, J. G., et al. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions*, *24*(8), 897–906.

Witek, K., Jupe, F., Witek, A. I., Baker, D., Clark, M. D., & Jones, J. D. G. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology*, *34*(6), 656–660.

Wolf, E., Kim, P. S., & Berger, B. (1997). MultiCoil: A program for predicting two-and three-stranded coiled coils. *Protein Science*, *6*(6), 1179–1189.

Wu, C. H., Derevnina, L., & Kamoun, S. (2018). Receptor networks underpin plant immunity. *Science*, *360*(6395), 1300–1301.

Wu, C. H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., et al. (2017). NLR network mediates immunity to diverse plant pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 8113–8118.

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, *475*(7355), 189–195.

Yandell, M., & Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics*, *13*(5), 329–342.

Yue, J., Meyers, B. C., Chen, J., Tian, D., & Yang, S. (2012). Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *The New Phytologist*, *193*(4), 1049–1063.

Zhang, Y., Wang, Y., Liu, J., Ding, Y., Wang, S., Zhang, X., et al. (2017). Temperature-dependent autoimmunity mediated by chs1 requires its neighboring TNL gene SOC3. *The New Phytologist*, *213*(3), 1330–1345.

Zhang, Y. M., Shao, Z. Q., Wang, Q., Hang, Y. Y., Xue, J. Y., Wang, B., et al. (2016). Uncovering the dynamic evolution of nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes in *Brassicaceae*. *Journal of Integrative Plant Biology*, *58*(2), 165–177.

Zhang, X., Dodds, P. N., & Bernoux, M. (2017). What Do We Know About NOD-Like Receptors in Plant Immunity? *Annual Review of Phytopathology*, *55*(1), 205–229.

Zhao, H., Li, C. C., Pardo, J., Chu, P. C., Liao, C. X., Huang, J., et al. (2005). A novel E3 ubiquitin ligase TRAC-1 positively regulates T cell activation. *The Journal of Immunology*, *174*(9), 5288–5297.

Zhong, Y., & Cheng, Z. M. (2016). A unique RPW8-encoding class of genes that originated in early land plants and evolved through domain fission, fusion, and duplication. *Scientific Reports*, *6*, 32923.

Zhong, Y., Zhang, X., & Cheng, Z. M. (2018). Lineage-specific duplications of NBS-LRR genes occurring before the divergence of six *Fragaria* species. *BMC Genomics*, *19*(1), 128.

Zhu, S., Li, Y., Vossen, J. H., Visser, R. G. F., & Jacobsen, E. (2012). Functional stacking of three resistance genes against *Phytophthora infestans* in potato. *Transgenic Research*, *21*(1), 89–99.

Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends in Immunology*, *35*, 345–351.

Zong, N., Xiang, T., Zou, Y., Chai, J., & Zhou, J. M. (2008). Blocking and triggering of plant immunity by *Pseudomonas syringae* effector AvrPto. *Plant Signaling & Behavior*, *3*(8), 583–585.