Short Title: Integrated domains in plant NLRs

- one-sentence summary: Integration of unusual domains in plant immune receptors is a widespread mechanism of NLR diversification enabling specific pathogen detection.

- author contributions: E. G, D. T., and L. D. equally contributed in the writing of the manuscript
Plant NLRs with integrated domains: Unity makes strength

Authors: Grund Elisabeth, Tremousaygue Dominique and Deslandes Laurent*

LIPM, Centre National de la Recherche Scientifique, Institut National de la Recherche Agronomique, INPT, Université de Toulouse, Castanet-Tolosan, France

*corresponding author : laurent.deslandes@inra.fr

1. Introduction

Unlike animals, plants lack an adaptive and circulating immune system. Thus, to detect pathogens and generate effective defense responses, plants rely on an elaborate innate immunity that involves different types of immune receptors (Cook et al., 2015). Conserved pathogen-associated molecular patterns (PAMPs) are recognized in the extracellular compartment of the host by cell surface-localized receptors. This event triggers the activation of basal immune responses called PAMP-triggered immunity (PTI). During evolution, pathogens have evolved sophisticated virulence strategies to overcome host defense responses. Host-adapted pathogens use an arsenal of virulence factors called effectors, that are delivered into the plant cell in order to subvert diverse cellular functions (effector-triggered susceptibility, ETS) through interference with PTI signaling (Jones and Dangl, 2006). However, effector activities can turn against the pathogen, as they often betray its presence within the cell. To recognize pathogen effectors, plants use a repertoire of intracellular immune receptors that belong to a superfamily of nucleotide-binding domain (NB) and leucine-rich repeat (LRR)-containing proteins (NLRs). NLRs can mediate specific recognition of pathogen effectors and initiate effector-triggered immunity (ETI). ETI involves transcriptional reprogramming overlapping with transcriptional regulations during PTI and often provokes localized host cell death at infection sites to limit pathogen spread (Jones and Dangl, 2006; Maekawa et al., 2011). Adapted pathogens can evade host recognition by reconfiguring their effector repertoires through various mechanisms including gain and loss of effector genes, modulation of expression, and rapid evolution of effectors by mutation (Arnold and Jackson, 2011; Lo Presti et al., 2015). Cycles of pathogen-induced ETS and plant-mediated ETI are therefore considered as major forces driving plant host-pathogen co-evolution (Jones and Dangl, 2006).
NLR proteins belong to STAND (Signal Transduction ATPases with Numerous Domains) P-loop NTPases. Canonical plant NLRs possess a multidomain architecture composed of a central nucleotide binding site (NBS) and a carboxy terminal leucine-rich-repeat (LRR) domain. The NBS is believed to function, through nucleotide-dependent conformational changes, as a molecular switch for NLR activation (Takken et al., 2006). Depending on the nature of their N-terminal domain, NLRs can be divided into two main classes: those having a Toll and interleukin-1 receptor (TIR) domain and those with a coiled-coil (CC) domain (CNLs) (Takken and Goverse, 2012). A third class, based on the presence of the N-terminal RPW8 domain, can also be defined. Over the last two decades, with the cloning of plant NLRs and their associated effectors, molecular characterization of the mechanisms employed by NLRs for specific effector recognition and signaling have been the subject of intensive research.

NLR-mediated effector recognition often involves host components that bind to and/or are modified by effectors (Figure 1). In the guard model, effector interference with a host target (co-factor or bait) is detected by the NLR, acting as a guard of modified-self (Dodds and Rathjen, 2010; Maekawa et al., 2011). Many identified guarded host proteins (also referred as “guardees”) represent hubs with key immune-related functions, including signaling or regulation of gene expression, and are therefore commonly targeted by various effectors (Mukhtar et al., 2011; Weßling et al., 2014). An effector “decoy” model has been proposed for a number of effector-sensing NLRs. In this model, the guarded host protein has no defense role but mimics an operational effector target and thus acts as a “decoy” that lures the pathogen effector and diverts it from its real target(s) (van der Hoorn and Kamoun, 2008; Lewis et al., 2013; Ntoukakis et al., 2014). Studies have shown that several NLRs can both detect pathogens and initiate downstream signaling, whereas other NLR proteins form heterogeneous protein complexes (Césari et al., 2014; Williams et al., 2014). In the core of these complexes are NLR pairs in which the two members are encoded by genes arranged in a head-to-head orientation with a common promoter region, which strongly suggests their co-regulation (Birker et al., 2009; Saucet et al., 2015). In several cases, the two proteins of NLR pairs form a heterocomplex receptor with each partner featuring specific attributes: one detects pathogen effectors (the sensor) while the other functions as an inducer of disease.
resistance (the transducer), and the signaling activity of the latter is repressed by the sensor (Césari et al., 2014; Williams et al., 2014). To explain how effectors are recognized by NLR pairs, an extension of the decoy model has been proposed with the “integrated decoy” hypothesis (Cesari et al., 2014). Indeed, sensor NLR partners were shown to contain, in addition to their conserved multidomain NLR architecture, unconventional domains that are able to physically interact with their corresponding effectors (Kanzaki et al., 2012; Cesari et al., 2013). Recent studies demonstrated that several of these integrated domains act as decoys of effector targets, enabling the sensor NLR to specifically detect pathogens (Le Roux et al., 2015; Maqbool et al., 2015; Ortiz et al., 2017) (Figure 1). Perturbations of the sensor NLR are perceived by the signaling partner, which activates immune signaling (Le Roux et al., 2015; Sarris et al., 2015; Ortiz et al., 2017). Comparative analyses of plant immune receptor architectures suggest that the integration of unusual domains (hereafter designated as “IDs” for integrated domains), which potentially serve as “baits” for pathogen effectors, is not restricted to paired NLRs and represent a widespread mechanism (Kroj et al., 2016; Sarris et al., 2016; Bailey et al., 2018). The identification of NLR-IDs signifies a breakthrough in plant NLR biology pathology, since it has profoundly changed our view of how plant NLRs can function and evolve.

In this review, we summarize current knowledge of NLR-IDs with detailed examples, discuss their genetic and functional diversity, and illustrate how the study of NLR function and mode of action has led to advances in plant disease control.

2. NLRs with integrated decoys: an ingenious pathogen detection mechanism

Recent independent studies have provided convincing evidence that IDs enable specific detection of pathogens by acting as molecular decoys that structurally mimic pathogen true virulence targets to monitor host immunosuppression attempts. How these integrated domains confer effector recognition and trigger activation of immune signaling are very intriguing questions. Well-characterized examples of NLR-ID fusions in paired NLRs include the Arabidopsis thaliana RRS1, which carries a WRKY domain (Le Roux et al., 2015), and the RGA5 and Pik-1 proteins from rice (Oryza sativa), both of which integrate an HMA (RATX1) domain (Ortiz et al., 2017). These examples are described in detail below.

The WRKY domain of RRS1
Experimental validation of the integrated decoy model was first provided for the Arabidopsis/Ralstonia solanacearum model. In 2001, Deslandes and colleagues cloned a resistance gene encoding RRS1-R (RESISTANCE TO RALSTONIA SOLANACEARUM) conferring broad-spectrum resistance to the soil-borne bacteria R. solanacearum, the causal agent of bacterial wilt (Deslandes et al., 1998; Deslandes et al., 2002). RRS1-R contains at its carboxyl terminus a ‘WRKY’ DNA-binding domain. This domain is conserved in defensive plant WRKY transcription factors that orchestrate biotic stress responses through the recognition of W-box cis-regulatory elements in gene promoters (Rushton et al., 2010). As being the first cloned NLR with an extra domain, RRS1-R was initially considered an anomaly in the field. Later, RRS1-R was shown to cooperate genetically and molecularly with a second TNL, namely RPS4 (RESISTANCE TO PSEUDOMONAS SYRINGAE) (Birker et al., 2009; Narusaka et al., 2009), to recognize effectors from different pathogens. These effectors included R. solanacearum PopP2, a member of the YopJ superfamily of acetyltransferase (Deslandes et al., 2003; Tasset et al., 2010); and AvrRps4, an unrelated effector from leaf-infecting Pseudomonas syringae pathovar pisi (Hinsch and Staskawicz, 1996; Sohn et al., 2012). Encoded by two co-regulated genes present in a head-to-head orientation, RRS1-R and RPS4 TNLs form a functional receptor recognition/signaling complex through homo- and heterodimerization involving their respective TIR domains (Williams et al., 2014). Two recent studies revealed that the RRS1-R/RPS4 NLR complex is activated through the targeting of the RRS1-R WRKY domain by PopP2 and AvrRps4 effectors (Le Roux et al., 2015; Sarris et al., 2015). Catalytically active PopP2 acetylates a key lysine residue (K1221) in the invariant heptad of the WRKY domain of RRS1-R, blocking its binding to W-box DNA sequences. Homology modeling predicts that K1221 acetylation disrupts WRKY domain electrostatic potential at the interface with DNA. In the absence of RRS1-R/RPS4 recognition, PopP2 uses this acetylation strategy to inhibit WRKY-DNA binding activities and transactivation functions needed for defence gene expression and basal resistance. RRS1-R WRKY domain represents therefore a decoy which betrays the defense-suppressing abilities of PopP2 and AvrRps4 on their host virulence targets: the defensive WRKY transcription factors. The direct fusion of a WRKY decoy domain within the RRS1-R/RPS4 NLR complex creates an efficient monitoring system for the indispensable virulence activities of different pathogens. Recently, Ma and colleagues demonstrated that, prior to effector
detection, the WRKY domain negatively regulates the RPS4-RRS1 complex through specific interactions with an adjacent domain in RRS1 (Ma et al., 2018). Binding of AvrRps4 to the WRKY domain disrupts these intramolecular interactions leading to the derepression of RRS1. Therefore, besides its effector sensing function, the integrated WRKY domain of RRS1 also has an important regulatory role in preventing inappropriate receptor activation in the absence of pathogens.

The HMA domain of RGA5 and Pik-1
The study of the RGA4/RGA5 receptor NLR pair in rice has enabled significant progress in deciphering the mode of action of paired NLRs. This NLR pair cooperates genetically and physically in the recognition of AVR-PiA and AVR1-CO39, two unrelated effectors of the rice blast fungus *Magnaporthe oryzae* (Okuyama et al., 2011; Cesari et al., 2013). In the absence of the pathogen, constitutive disease resistance and cell death mediated by RGA4 is repressed by RGA5 through the formation of heteroprotein complexes. The C-terminal part of RGA5 contains a Heavy Metal-Associated (HMA) domain, initially found in a cytoplasmic copper chaperone in *Saccharomyces cerevisiae*, which can directly interact with AVR-PiA and AVR1-CO39, enabling pathogen recognition. Physical association of the AVR-PiA effector with the HMA domain of RGA5 triggers cell death through RGA4 de-repression (Césari et al., 2014). Interestingly, recognition of the AVR-Pik effector of *M. oryzae* by the unrelated CC-NLR pair Pik-1/Pik-2 in rice is also triggered by direct binding to an HMA domain in Pik-1 that, contrary to RGA5, is integrated between its CC and NB regions (Ashikawa et al., 2008; Maqbool et al., 2015). The different location of the HMA domain in RGA5 and Pik-1 suggests that these domains have been fused to those two unrelated NLRs through independent events (Cesari et al., 2013). Although HMA domain-containing proteins have not been previously described as effector targets, the presence of a HMA domain in the rice Pi21 factor, which is required for full susceptibility to the rice blast fungus (Fukuoka et al., 2009; Zhang et al., 2016), supports the idea that the HMA domains of RGA5 and Pik-1 are decoys for AVR-PiA, AVR1-CO39, AVR-Pik, and functionally related effectors. Determination of the crystal structure of AVR-PikD complexed with a dimer of the Pikp-1 HMA domain revealed that key residues at the interaction interface are required for effector binding and recognition (Maqbool et al., 2015). In addition, variations at binding interfaces between AVR-Pik effector variants and HMA domains of Pik alleles were found to
determine recognition specificity. Such recent findings highlight how new receptor specificities arise from natural selection (De la Concepcion et al., 2018). How the binding of effectors to HMA domains can trigger activation of immune signaling remains unknown. It is hypothesized that effector binding promotes NLR domain rearrangements leading to immune complex activation (Césari et al., 2014). Binding of AVR-PiA to RGA5 HMA domain is also necessary for pathogen recognition, but protein-protein interaction analyses revealed that moderate affinity to mutated AVR-PiA proteins still confers recognition (Ortiz et al., 2017). Furthermore, additional sites in RGA5, outside the ID, are suspected to mediate interaction with the effector. Thus, the juxtaposition of integrated decoy domains with additional NLR sites interacting properties creates a highly resilient surveillance system. How the binding of effectors to HMA domains can trigger immune signaling remains unknown. It is hypothesized that effector binding promotes NLR domain rearrangements leading to immune complex activation (Césari et al., 2014).

The NOI/RIN4 domain of Pii-2

Pathogen detection by paired NLR-IDs is not restricted to the direct recognition model. Indeed, IDs might also function in indirect recognition by perceiving modifications of a host protein targeted by an effector. This concept is supported by a study of the unconventional NOI/RIN4 domain of the rice NLR-ID Pii-2 that is hypothesized to monitor, through direct binding, the integrity of the OsExo70-F3 host protein, a target of the M. oryzae effector AVR-Pii (Fujisaki et al., 2017).

3. NLR-IDs: a mechanism of NLR diversification

The search for protein domains associated with typical NLR domains in public databases made it possible to identify entire NLR-ID directories and to analyze their structure in many plants. Already present in mosses, NLR-IDs occasionally represent a large proportion of NLRs in terrestrial plants (Kroj et al., 2016; Sarris et al., 2016; Bailey et al., 2018; Stein et al., 2018) (Table1). Kroj et al. (2016) detected 162 hypothetical IDs across 33 genomes by looking for interpro domains using the GreenPhyl database. Although their analysis was not exhaustive due to potential misannotations of the applied databases, they identified a high diversity of IDs (90 different domains). Sarris et al. (2016) reported 265 unique ID fused to NLRs in 37 genomes of land plants (Table1). Bailey et al. (2018) identified NLR-IDs in 9
grass species. 31 types of different domains were mainly represented in these species. By focusing on 13 *Oryza* species, Stein et al. (2018) described a highly variable structure of genes coding for several hundreds of different NLR-ID proteins. They were able to detect a significant enrichment for these NLR-IDs within pairs of genes arranged in a head-to-head configuration in the genomes. The widespread distribution of NLR-IDs, despite their low abundance in some plant genomes (see Table 1), suggests a successful evolutionary mechanism of NLR diversification commonly used by plants to expand their pathogen recognition capabilities, allowing them to cope with highly and rapidly adaptable pathogen-derived molecules. Accordingly, the IDs identified in these studies, are derived from proteins which are extremely diverse in sequence and predicted molecular functions. The most frequent domains found in NLR-IDS include the WKRY and BED zinc finger (Znf-BED) DNA-binding domains, and the protein kinase domains. The decoy function of the WRKY domain in the RRS1-R NLR has been already validated (see above) (Le Roux et al., 2015; Sarris et al., 2015). The BED-zinc finger domain was originally identified in transposases and transcription factors (Hayward et al., 2013) but, contrary to RRS1-R, the targeting by pathogen effectors remains to be demonstrated.

However, Kroj and colleagues showed that the ZBED NLR protein from rice, containing three BED domains, is required for resistance to *M. oryzae* (Kroj et al., 2016). In response to the pathogen, ZBED over-expressing lines were more resistant, whereas a *zbed* null mutant showed increased susceptibility. These data strongly suggest that the BED domains in ZBED NLR proteins represents decoys that mimic host BED proteins targeted by *M. oryzae* effectors. The Xa1 NLR from rice, which confers resistance against isolates of the bacterial blight pathogen *Xanthomonas oryzae* by recognizing multiple TALEs (Yoshimura et al., 1998; Ji et al., 2016), contains a Znf-BED domain in its N-terminal part. The mechanism that allows Xa1 to recognize TALEs remains to be elucidated. It is tempting to speculate that Xa1 Znf-BED domain might also act as a decoy to lure TALEs that target host Znf-BED proteins, for the subversion of host gene expression (Zuluaga et al., 2017).

Also the functionality of NLR IDs with predicted catalytically active protein kinase domains need to be experimentally validated. However, their sensing abilities can be deduced from well described examples of kinases acting as decoys that physically interact with classical NLRs, e.g. the kinases Pto and PBS5 interacting with the NLRs Prf and RPS5, respectively (Miao et al., 2015). It is not clear why in primitive land
plants the fusion of Kinase domains or DUF676 to NBS-LRRs that lack CC or TIR
domains has been described (Gao et al., 2018), suggesting that these kind of IDs
could ensure the signaling function of these missing domains. Therefore, such IDs in
NLR proteins could fulfill either sensor or signaling functions, or both.
Interestingly, there are significant overlaps between IDs and protein domains
previously identified as interacting partners of effectors in interactome screens
(Mukhtar et al., 2011; Weßling et al., 2014; Sarris et al., 2016), including well-
characterized guards or decoys. Examples are the exocyst complex factor Exo70,
required for recognition of AvrPii by NLR Pii in rice (Fujisaki et al., 2015), and RIN4, a
target of multiple effectors that is guarded by RPS2 and RPM1 NLRs in Arabidopsis
(Mackey et al., 2002; Mackey et al., 2003; Kim et al., 2005). Such overlaps strongly
suggest that IDs could act as sensor/decoy by mimicking effector targets. Since
many IDs correspond to protein domains with unknown biological activity, they
represent promising candidates to uncover host components targeted by effectors
and whose participation in various layers of plant immunity has not been assigned
yet.
Whether all the putative IDs identified in the whole genome analyses also serve as
sensor/decoy in immune response remains to be demonstrated. Moreover, detailed
investigations on gene structure and function should help to reduce false positive
among computationally predicted NLR-IDs (Giannakopoulou et al., 2016), and shed
light on new resistance mechanisms.

4. Mechanisms involved in NLR-ID fusion events

Within NLR-IDs, the majority of IDs appear as singular N- or C-terminal domains.
However, in some cases the fusion of several domains in the same protein is
observed. For example, AtWRKY19 NLR in Arabidopsis integrates both an N-
terminus WRKY domain and a C-terminal kinase domain. In a minority of cases,
including rice Pik-1, integration has occurred between the N-terminal signaling
domain and the central nucleotide binding domain of the NLR. These observations
indicate that some NLRs can tolerate integration of sensor domains at various
positions in their modular architecture, while maintaining their signaling functions.
Bailey et al. recently investigated the evolutionary dynamic of NLR-IDs in the
genomes of nine grass species. They concluded that NLR-IDs in grasses were not
evenly distributed across the phylogeny, but a single species with up to 58% of
NLRs containing IDs was observed (Bailey et al., 2018). In this clade, they highlighted an amino acid sequence motif located immediately upstream of the fusion site, which could play an important role in the integration process. They proposed that DNA transposition and/or ectopic recombination is a major driving force behind domain integration in grasses; repeated independent integration events were observed, suggesting that integration occurred frequently and independently during evolution, giving rise to a high diversity of IDs. Similarly, Brabham et al. recently showed that orthologs of the RGH2 NLR from species across the grasses were subject to large variation in domain structure, including the presence/absence of an integrated Exo70 domain (Brabham et al., 2018). These trans-species polymorphisms provided an opportunity to follow the molecular evolution of the Exo70 gene families and to investigate the role of Exo70 as an integrated domain in the RGH2 NLR. This study showed that upon pathogen pressure, non-integrated Exo70 genes are under strong purifying selection, whereas they are under relaxed purifying selection when integrated into RGH2. Across the Oryza genomes, the presence of IDs in 17 different NLR subfamilies (from a total of 36) point to multiple and independent acquisition of IDs (Stein et al., 2018).

5. NLR-IDs: towards the elucidation of additional NLRs functions?

Besides the well-documented role of NLRs for innate immunity in both animals and plants, additional functions controlled by NLRs are currently discussed. In animals, NLRs play a role in developmental processes, such as spermatogenesis and fertility, suggesting a control of the reproductive system by NLR proteins (Meunier and Broz, 2017). In plants, inappropriate activation of NLRs caused by mutations or incompatible combinations also impact development (Chae et al., 2014; Chen et al., 2016; Atanasov et al., 2018; Chakraborty et al., 2018). The integrated decoy model predicts an important role of IDs in pathogen detection. Beyond their function of effector sensor, some IDs might have retained the biological activities of the proteins from which they are derived. Thus, additional functions controlled by NLRs in plants could be revealed by looking at IDs.

In this regard, the integration of a BED domain, one of the most frequently found ID in plant species, gained attention. This domain appears to be shared by transposases and by proteins that perform critical cellular functions (Aravind, 2000; Hayward et al., 2013). The BED domain is a domain of the Type I Transposase of Activator from
maize) transposase has been shown to suppress the DNA TAM3 transposon activity
in *Anthirrhinum majus* by re-localizing the Tam3 transposase out of the nucleus
(Zhou et al., 2017). More globally, BED-related IDs could be sensors of cellular
homeostasis perturbation by environmental stress which rely in part on DNA
transposition control (Negi et al., 2016).

In the case of RRS1, inhibition of its DNA binding activity provoked by particular
mutations in its WRKY domain leads to autoimmunity in a RPS4-dependent manner
(Noutoshi et al., 2005; Sohn et al., 2014). Therefore, RRS1-R was initially considered
a negative regulator of immune-related genes. The autoimmune phenotype of the
RRS1-R *slh1* variant is conditioned by low humidity, suggesting that RRS1-R,
besides its function in pathogen recognition, could also sense particular
environmental disturbances such as drought stress. In response to specific biotic and
abiotic stresses, RRS1-R likely acts directly at genomic DNA and, together with
RPS4, behave as a reactive switch for transcriptional and signaling reprogramming.
Whether NLR-IDs possess broader sensing functions remains to be determined, but
almost certainly the functional characterization of their IDs will help to elucidate
potential additional NLRs functions.

6. Towards the engineering of synthetic NLR-IDs with extended recognition
capabilities

The high diversity of NLRs provides plants with versatile options for effective plant
immunity. However, on one hand, the ability of effectors to evolve rapidly and, on the
other hand, the necessary fine-tuning of NLR functions to avoid autoimmunity (see
Box 1) restrict the possibilities to transfer immune receptors into other plant genomes
for improved crop protection. Hence, strategies for NLR engineering are of particular
interest in plant immunity research. If successful, it could be possible to modify NLR
recognition specificity, or alternatively widen the range of effector recognition while
making it less feasible for pathogens to bypass NLRs. One possibility is to alter the
structure of NLR proteins itself. For example, a study conducted by Segretin et al.
showed that few amino acid changes in the LRR domain of R3a in potato (*Solanum
tuberosum*) enabled this NLR to recognize another isoform of AVR3a from
*Phytophthora infestans* (Segretin et al., 2014). However, direct NLR effector
recognition (modified self) is likely to be less tolerant to variations compared to the
indirect and integrated decoy models. Hence, decoy engineering represents a
more suitable approach. A proof-of-concept study was recently performed by Kim and colleagues (2016), demonstrating successfully the modification of the Arabidopsis protein kinase PBS1 that represents a decoy for pathogen-derived proteases (Kim et al., 2016). PBS1 is involved in basal immune response and acts as target/decoy for the effector AvrPphB from *Pseudomonas syringae*, which can cleave PBS1 due to its protease activity (Shao et al., 2003; Ade et al., 2007). PBS1 cleavage products are recognized by the NLR protein RPS5 (DeYoung et al., 2012) that in turn initiates resistance signaling. For a modified PBS1 decoy variation, the proteolytic target site, which is normally cleaved by AvrPphB, is exchanged with another proteolytic site. This modification enables the cleavage of PBS1 by other effectors, such as AvrRpt2 of *P. syringae*, Nla protease of Tobacco Etch Virus (TEV) and Nla protease of Turnip Mosaic Virus (TuMV). AvrRpt2 also triggers activation of RPS2 NLR by cleaving a nitrate-induced (NOI) domain present in RIN4, a negative regulator of plant defense which is guarded by multiple NLR proteins (Mackey et al., 2002; Mackey et al., 2003; Kim et al., 2005). Such protease cleavage sites are found in IDs present in a subset of NLRs, paving the way to the design of single NLR-IDs combining multiple protease recognition sequences. The modifications of NLR IDs, as well as the replacement of IDs with other identified effector targets from the host genome, display other promising tools to create novel effector traps. Considering that modifications of NLRs as well as of their IDs can compromise homo/heterocomplex formation required for their function and trigger either inactivation or autoactivation of the receptor, it is important for this approach to identify which IDs and which NLRs are best suited for fusion manipulation (e.g. IDs swapping/shuffling). For this, we need to gain better knowledge of interaction sites within NLR and its IDs, but also of the interaction of sensor and signaling NLRs. Furthermore, it is important to determine the layout of effector-ID complexes. Recently, the structure of several complexes has been resolved (Maqbool et al., 2015; Ortiz et al., 2017; Zhang et al., 2017). Determination of the molecular and structural bases of these interactions is crucial for the successful design of synthetic “multi-sensor NLR-IDs” made from the juxtaposition of different IDs, giving them extended recognition capabilities.

**Conclusion**
The past 20 years have brought major advances in plant NLR biology. The discovery of NLR-IDs represents a significant step towards a better understanding of mechanisms involved in the evolution and function of NLRs. During plant evolution, NLR-IDs appeared independently several times and in different configurations. Besides the sensing and regulating functions of IDs, important questions remain unanswered (see “Outstanding questions” box). There is still a critical need for further in-depth studies to establish the biological functions of IDs fused to NLRs and to elucidate the molecular events which link NLR-IDs activation with immunity pathways. Overall, NLR-IDs provide promising tools for the design of new strategies to protect plants against pathogens. Nevertheless, it remains to be determined whether engineered synthetic NLR-IDs can provide sustainable disease resistance, especially in different crop species.

Acknowledgements
E.G, D.T and L.D are supported by the ANR project RADAR (ANR15-CE20-0016-01). The authors apologize to researchers whose contributions could not be cited more extensively because of space limitations.
Figure legends.

Figure 1:
NLRs can directly or indirectly detect the presence of pathogen effectors by monitoring the manipulation of their host targets (baits or decoys, a). According to the integrated decoy model, integrated domains (IDs) in NLRs behave as decoy of effector targets enabling recognition of effector activities. This recognition can be direct or indirect (b and c, respectively). Different studies reported the existence of diverse IDs (in sequence and predicted molecular functions), which can be present at various positions within the modular structure of NLRs (d). NLR-IDs can be engineered using different strategies aimed at providing (i) extended specificity (i.e., specific point mutations in IDs enabling recognition of various allelic forms of a pathogen effector), (ii) multi-recognition capabilities (by integrating IDs from different NLRs within a single NLR), or (iii) new recognition specificities (by integrating previously characterized effector targets that then act as sensors) (Nt : N-terminal domain ; NBS : nucleotide binding site ; LRR : leucine-rich repeat ; E : Effector ; T : effector target ; ID : integrated domain ; ET : previously characterized effector target)

Table 1. Overview of NLR-ID repertoires in plant species

<table>
<thead>
<tr>
<th>Reference</th>
<th>No. of species investigated</th>
<th>No. of species with NLR-ID</th>
<th>No. of NLR-IDs</th>
<th>No. of NLR-IDs / species</th>
<th>Average percent of NLR-IDs (of all NLRs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kroj et al., 2016</td>
<td>33</td>
<td>23</td>
<td>34</td>
<td>from 1 to 16</td>
<td>3.5 %</td>
</tr>
<tr>
<td>Sarris et al., 2016</td>
<td>36</td>
<td>35</td>
<td>717</td>
<td>from 1 to 93</td>
<td>6.8 %</td>
</tr>
<tr>
<td>Bailey et al., 2018</td>
<td>9</td>
<td>9</td>
<td>331</td>
<td>from 7 to 133</td>
<td>7.9 %</td>
</tr>
<tr>
<td>Funk et al., 2018</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>24</td>
<td>14.0 %</td>
</tr>
<tr>
<td>Stein et al., 2018</td>
<td>13</td>
<td>13</td>
<td>446</td>
<td>_</td>
<td>8.2 %</td>
</tr>
</tbody>
</table>
References


Kroj T, Chanclaud E, Michel-Romiti C, Grand X, Morel JB (2016) Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. New Phytopath


Single amino acid mutations in the potato immune receptor R3a expand response to Phytophthora effectors. Mol Plant Microbe Interact 27: 624-637


ADVANCES

• Unusual domains integrated in the architecture of some plant NLRs (NLR-IDs) behave as decoys that mimic virulence targets of effector proteins, thus enabling pathogen detection.

• IDs in NLRs are extremely diverse in sequence and predicted molecular functions.

• The widespread distribution of NLR-IDs suggests an evolutionary mechanism of NLR diversification used by plants to expand their pathogen recognition capabilities.

• Identification of NLR-IDs paves the way to the development of novel strategies of plant disease control.
OUTSTANDING QUESTIONS

• What are the various mechanisms and components involved in the activation of NLR-ID proteins? Can we identify novel immunity components by characterizing IDs?
• Are NLR-IDs further involved in the downstream signaling mechanisms after activation by pathogen effectors?
• What additional functions controlled by plant NLRs can be revealed by analyzing NLR-IDs?
• What are the dynamics of host-pathogen co-evolution? How quickly can NLR-IDs be bypassed by pathogen effectors? What are the requirements necessary to successfully engineer NLR-IDs? What are the precautions to be considered when selecting the ID and structure modifications to ensure that novel NLR-IDs confer durable resistance, with minimal risk of functionality-loss or autoactivation?
• Is it possible to engineer a universal NLR-ID that confers broad pathogen resistance in various plant species?
BOX 1. Immunity Under Strict Control

The plant immune response is tightly adjusted to biotic stress and not activated in a non-challenging environment. In case of inappropriate defense activation, plant fitness is affected, and drastic developmental phenotypes such as tissue necrosis or dwarfism can be observed (Chakraborty et al., 2018). These phenotypes are similar to what is observed in plant hybrids when genetic incompatibility, resulting from incompatible allelic forms of parental genes, occurs (Chen et al., 2016). Chae et al. (2014) revealed that NLR alleles are responsible for most incompatibilities in A. thaliana (Chae et al., 2014). Accordingly, mutations in NLRs efficiently suppress hybrid incompatibility (Atanasov et al., 2018). NLR-IDs that negatively regulate the function of signaling NLRs could play a major role for appropriate activation of immune-related mechanisms. Hence, mutations in IDs can be responsible for autoactivation of NLRs. This is exemplified by the slh1 (susceptible to low humidity) mutation in the integrated WRKY domain of RRS1 that, under drought conditions, leads to constitutive activation of plant defense and a severe autoimmune dwarf phenotype (Noutoshi et al., 2005). A similar phenotype is obtained upon expression of an RRS1-R variant mimicking a modification of the WRKY domain by its cognate effector and triggering constitutive activation of the RPS4/RRS1-R complex (Le Roux et al., 2015).

This very sensitive recognition device is accompanied by a functional separation with signaling, through the involvement of NLR pairs: one NLR detects the presence of the pathogen, the other initiate downstream signaling (see Introduction). Independent evolution of the two NLR proteins adds flexibility to ensure the effectiveness of the immune system. New effectors can thus be recognized and activate existing signal pathways. It is also possible to modify signaling without affecting the recognition of the effectors, this can especially be useful in changing environmental parameters. A strict control of NLR structure and expression is therefore essential to ensure plant fitness and development. Hence, to avoid a dramatic reduction of plant vigor in future transgenic plants expressing engineered NLR-IDs, and to promote durable resistance engineering, it is necessary to perform more extensive structural analyses of NLR-IDs, and to gain a more comprehensive understanding of the mechanisms controlling their activation.
NLRs can directly or indirectly detect the presence of pathogen effectors by monitoring the manipulation of their host targets (baits or decoys, a). According to the integrated decoy model, integrated domains (IDs) in NLRs behave as decoy of effector targets and enable recognition of effector activities. This recognition can be direct or indirect (b and c, respectively). Different studies reported the existence of diverse IDs (in sequence and predicted molecular functions) which can be present at various positions within the modular structure of NLRs (d). NLR-IDs can be engineered using different strategies aimed at providing (I) extended specificity (i.e., specific point mutations in IDs enabling recognition of various allelic forms of a pathogen effector), (II) multi-recognition capabilities (by integrating IDs from different NLRs within a single NLR), or (III) new recognition specificities (by integrating previously characterized effector targets that then act as sensors) (Nt: N-terminal domain; NBS: nucleotide binding site; LRR: leucine-rich repeat; E: Effector; T: effector target; ID: integrated domain; ET: previously characterized effector target).

**Figure 1**: Different mechanisms of effector recognition mediated by NLR-IDs.


Google Scholar: Author Only


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title


Copyright (c) 2020 American Society of Plant Biologists. All rights reserved.


Kroj T, Chanclud E, Michel-Romiti C, Grand X, Morel JB (2016) Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. New Phytol


Sarris PF, Cevik V, Dagdas G, Jones JD, Krasileva KV (2016) Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. BMC Biol 14: 8


