Natural and CRISPR-induced genetic variation for plant immunity

Baptiste CASTEL

Thesis submitted to the University of East Anglia for the degree of

Doctor in Philosophy

The Sainsbury Laboratory

February 2019

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.
Abstract

Our understanding of the genetic basis of a trait primarily relies on analysing heritable phenotypic diversity.

For instance, different accessions of *Arabidopsis thaliana* (Arabidopsis) can either be resistant or susceptible to a given strain of *Albugo candida*, an oomycete that causes the white rust disease. The virulent *Albugo candida* race Exeter1 (AcEx1) can grow on most Arabidopsis accessions. Using the resistant Arabidopsis Oy-0, I mapped and cloned the gene responsible for AcEx1 resistance: *White Rust Resistance 4A* (WRR4A). Arabidopsis Col-0 also contains WRR4A but does not resist AcEx1. I found that WRR4A<sup>Col-0</sup> has an early stop codon, which is responsible for the recognition specificity of some effector candidates from *Albugo candida*. This example illustrates how natural diversity can be used to identify Resistance-genes and reveal components of the plant immune system.

However, natural diversity is not always available. Clustered and regularly interspaces short palindromic repeats (CRISPR) from bacterial genomes defines an immune system, re-invented for genome editing. I optimized a CRISPR-Cas9 method to generate null alleles in Arabidopsis. Using this method, I produced a double mutant of two immunity-related gene candidates that are in tandem in the genome: AtNRG1A and AtNRG1B. I confirmed the 7-year-old hypothesis that NRG1A and NRG1B are redundantly required for signalling downstream of multiple Resistance-genes, mainly from the TIR-NLR immune receptor family. So far very few genes required for immunity upon Resistance-protein activation were defined. This second example illustrates that CRISPR can be used to generate variation to unravel redundant genetic pathways.

The widespread adoption of CRISPR tools is likely to lead to a better understanding of the plant immune system. Ultimately, it will result in solutions to deploy genetics-based resistance to protect our crops from disease, reducing the need for chemicals.
# Table of contents

Abstract ............................................................................................................................. 3
Table of contents .............................................................................................................. 4
Acknowledgments ............................................................................................................ 8
Major abbreviations ........................................................................................................ 9
List of publications .......................................................................................................... 10

Chapter 1: General Introduction ..................................................................................... 11

1.1 Molecular Plant-Microbe Interactions ..................................................................... 11
   1.1.1 The plant immune system .................................................................................. 11
   1.1.2 NLRs: Structure ................................................................................................ 13
   1.1.3 NLRs: Activation .............................................................................................. 17
   1.1.4 NLRs: Signalling ............................................................................................... 20

1.2 *Albugo candida*, a plant pathogen oomycete ......................................................... 25
   1.2.1 Taxonomy ........................................................................................................ 25
   1.2.2 Life cycle .......................................................................................................... 25
   1.2.3 Impact and management .................................................................................. 26
   1.2.4 Virulence .......................................................................................................... 26
   1.2.5 *White Rust* Resistance-genes discovery ......................................................... 28

1.3 CRISPR for genome editing ...................................................................................... 29
   1.3.1 CRISPR: a bacterial immune system ................................................................. 29
   1.3.2 CRISPR for genome editing .............................................................................. 30
   1.3.3 CRISPR for plant immunity .............................................................................. 37

Chapter 2: Materials and methods .................................................................................. 39

2.1 Material ..................................................................................................................... 39
   2.1.1 Plant material ................................................................................................... 39
   2.1.2 Microbial material ........................................................................................... 39
4.2.2 Allyl alcohol enables to select CRISPR-induced Arabidopsis mutations ..........68
4.2.3 An overdrive sequence at the T-DNA right border does not affect the CRISPR activity ......................................................................................................................................................... 70
4.2.4 UBI10, YAO or RPS5a promoter-controlled Cas9 expression enhances mutation rates ........................................................................................................................................................................ 72
4.2.5 Codon optimization of Cas9 and presence of an intron elevate mutation rates ........................................................................................................................................................................................................... 75
4.2.6 A modified sgRNA triggers CRISPR-induced mutations more efficiently ..........77
4.2.7 The 3’ regulatory sequences of Cas9 and the sgRNA influence mutation rates 79
4.2.8 Divergent transcription of Cas9 and sgRNA can elevate mutation rates .......... 82
4.2.9 Most of the stable double events are homozygous rather than biallelic .......... 84
4.3 Discussion ........................................................................................................................................................................................................................................................................... 85
4.3.1 Identification of a CRISPR T-DNA leading to elevated mutation rates .......... 85
4.3.2 DSB-induced allelic recombination results in more homozygous than biallelic lines...................................................................................................................................................................................................................... 86
4.3.3 Non-lethal selectable markers to obtain T-DNA-free mutants ..................... 86

Chapter 5 : NLR immune receptors signal via the RPW8-NLR NRG1 .......................88
5.1 Introduction ........................................................................................................................................................................................................................................................................... 88
5.2 Results ........................................................................................................................................................................................................................................................................... 89
5.2.1 Identification of ADR1 in N. benthamiana genome ........................................ 89
5.2.2 CRISPR enables recovery of nrg1 mutants in Arabidopsis and N. benthamiana and a wrr4b mutant in Arabidopsis ........................................................................................................................................................................................................................................ 92
5.2.3 NRG1 is required for RRS1/RPS4-mediated HR but not bacterial resistance in Ws-2 ........................................................................................................................................................................................................................................ 94
5.2.4 Arabidopsis nrg1 mutants show impaired TIR-NLR-dependent resistance to oomycete pathogens ........................................................................................................................................................................................................................................ 100
5.3 Discussion ........................................................................................................................................................................................................................................................................... 106
5.3.1 NRG1 is a conserved clade of RPW8-NLR that contributes to TIR-NLR signalling ........................................................................................................................................................................................................................................ 106
5.3.2 Disease resistance is possible without HR ......................................................... 108
5.3.3 TIR-NLRs activate in parallel an SA-dependent pathway and an NRG1-dependent pathway that results in HR ................................................................................................. 109
5.3.4 Some TIR-NLRs signal via NRG1 only, some signal via both ADR1 and NRG1 .. 109

Chapter 6 : General discussion ................................................................................. 112
6.1 History of plant-microbe interaction genetics ..................................................... 112
6.2 Limit to R-gene discovery .................................................................................... 113
6.3 Meeting the challenge of lack of natural diversity .............................................. 113
  6.3.1 Successful approaches .................................................................................. 113
  6.3.2 Unsuccessful approach ................................................................................ 113
6.4 Beyond R-gene discovery: characterization of NLR signalling components ...... 114
  6.4.1 A black box between plant NLR and actual resistance ............................... 114
  6.4.2 RPW8-NLRs are candidates to be major ETI components ....................... 114
  6.4.3 RPW8-NLRs are indeed major ETI regulators, as proven using CRISPR ...... 115
6.5 How do RPW8-NLRs regulate ETI? .................................................................. 115
  6.5.1 What can we learn from RPW8-NLR homologues in fungi and animals? .... 115
  6.5.2 What do we currently know about RPW8-induced immunity? .................... 116
  6.5.3 Can we formulate a unifying hypothesis for RPW8-NLR function? ............ 117
6.6 What to expect from CRISPR? .......................................................................... 117
  6.6.1 A powerful genetic tool for biology ............................................................. 117
  6.6.2 Deployment of knock-in breeding ............................................................... 118
  6.6.3 Deployment of base editing ........................................................................ 118
  6.6.4 Perspective for agriculture ........................................................................ 119
References .............................................................................................................. 120
Acknowledgments

I feel very fortunate to have completed my PhD at The Sainsbury Laboratory with Professor Jonathan Jones. Jonathan has been a constant source of support, encouragement and novel ideas for the projects. He will remain an example and a source of inspiration for the rest of my scientific career.

I am also grateful for the team members he managed during my PhD. Particularly I would like to thank Dr Volkan Cevik, Dr Pingtao Ding and Dr Laurence Tomlinson for their valuable mentoring. Their professional and talented guidance helped me achieve rapid progress during my first years in the laboratory. Their enthusiasm for science largely contributed to develop my own.

I want to acknowledge all the Albugans for their help and for running such an enjoyable working team: Volkan Cevik, Oliver Furzer, Agathe Jouet, Amey Redkar, Sebastian Fairhead and Dae-Sung Kim. I also want to thank the CRISPRers: Laurence Tomlinson, Lila Grandgeorge, Federica Locci, Ying Yang, Oleg Raitskin and Nicola Patron; and the RRS1/RPS4 team including but not limited to Yan Ma, Pingtao Ding, Hannah Brown, Joanna Feehan, Panos Sarris, Panos Moschou, Sung Un Huh, Bruno Ngou, Zane Duxbury, Billy Taker-Brown, Hailong Guo, Hee-Kyung Ahn and Shanshan Wang.

I would like to thank my secondary supervisor Professor Cyril Zipfel, as well as the committee members Dr Volkan Cevik, Dr Nicola Patron and Professor Wendy Harwood. They have been a precious support for my PhD management.

The supportive team at TSL and JIC greatly facilitated my projects: the SL-JJ and JIC horticultural services, the administration team, the lab support team, the bioinformatics team, the IT support at TSL, the synthetic biology team and the tissue culture team.

Finally, I would like to thank all the people I met in Norwich including my numerous housemates (particularly Sylvain Capdevielle), my TSL, JIC, TGAC/EI, IFR/QIB, UEA mates, my climbing mates, my salsa mates, my couchsurfing mates, my cycling mates, my jiu-jitsu mates, my gym mates, my JIC Recreational Centre and pub mates.

To my family, my master cohort BFP, “Les Jozés” and all these people who made my PhD journey such a wonderful experience, thank you all again.
**Major abbreviations**

| Ac       | *Albugo candida*          |
| Avr      | Avirulence                |
| c-terminal | Carboxyl-terminal         |
| Cas      | CRISPR-associated         |
| CC       | Coiled-coil               |
| CDS      | Coding sequence           |
| CRISPR   | Clustered and regularly inter-spaced palindromic repeat |
| DSB      | Double-strand break       |
| DNA      | Deoxyribonucleic acid     |
| Dpi/hpi  | Day(s) / hour(s) post-inoculation |
| EDS1     | Enhanced disease susceptibility 1 |
| ETI      | Effector-triggered immunity |
| ETS      | Effector-triggered susceptibility |
| GFP      | Green fluorescent protein |
| HDR      | Homology-directed recombination |
| Hpa      | *Hyaloperonospora arabidopsis* |
| HR       | Hypersensitive Response   |
| KIB      | Knock-in breeding         |
| LRR      | Leucine-rich repeat       |
| N-terminal | Amino-terminal            |
| NB       | Nucleotide binding        |
| NHEJ     | Non-homologous end joining |
| NLR      | NB and LRR                |
| NRG1     | N-requirement gene 1      |
| OD       | Optical density           |
| PAMP     | Pathogen-associated molecular pattern |
| PCR      | Polymerase chain reaction |
| PRR      | Pattern recognition receptor |
| PTI      | PAMP-triggered immunity   |
| R        | Resistance                |
| RNA      | Ribonucleic acid          |
| ROS      | Reactive oxygen species   |
| RPP      | Resistance to *Hyaloperonospora arabidopsis* |
| RPW8     | Resistance to powdery mildew 8 |
| SA       | Salicylic acid            |
| TIR      | Toll, interleukin-1 receptor and many R-proteins |
| WRR      | White Rust Resistance     |
| WT       | Wild type                 |
List of publications


Chapter 1: General Introduction

1.1 Molecular Plant-Microbe Interactions

1.1.1 The plant immune system

1.1.1.1 Two layers of plant defence

Plants can interact with micro-organisms including fungi, oomycetes, bacteria and viruses. Many are found at the plant surface, forming the phyllosphere in the shoots and the rhizosphere in the roots. Some can penetrate within the plant tissue. They form biological interactions from symbiotic to parasitic. Many cases of beneficial interactions have been reported, such as mycorrhiza and nitrogen fixing symbioses (Oldroyd, 2013). Microorganisms can also negatively affect plant physiology to promote their own, in so-called pathogenic interactions.

The outcome of such interactions relies on a molecular dialogue between the pathogen and its host (Jones & Dangl, 2006). On the one hand, plants can recognize conserved microbial molecules resulting in pattern-triggered immunity (PTI). On the other hand, pathogens deliver virulence factors, known as effectors, to colonize the plant, feed, reproduce, counteract plant defences and enhance their fitness resulting in effector-triggered susceptibility (ETS). Plants have also evolved intracellular immune surveillance devices, called Resistance (R)-proteins, which recognise effectors, resulting in effector-triggered immunity (ETI). This three-component conceptual framework has been called the zig-zag-zig model (Jones & Dangl, 2006).

1.1.1.2 Pattern-triggered immunity (PTI)

Plants detect apoplastic elicitors via cell surface Pattern Recognition Receptors (PRRs) (Zipfel, 2008). PRRs can be Receptor-Kinases (RKs) or Receptor-like proteins (RLPs). Both harbour an extracellular domain involved in elicitor perception; RKs also have an intracellular kinase domain. Their cognate elicitors are in general conserved pathogen/microbe-associated molecular patterns (PAMPs/MAMPs), such as flagellin and elongation factor thermo unstable (EF-Tu) for bacteria, chitin for fungi or coat protein for viruses (Boutrot & Zipfel, 2017). Some PAMPs are even shared across microbial kingdoms, such as the “necrosis- and ethylene-inducing protein 1”-like proteins (NLPs) that are found
in oomycetes, bacteria and fungi and recognised by the PRR RLP23 (Albert et al., 2015; Lenarčič et al., 2017).

PRRs can also recognise endogenous compounds degraded during pathogen attack, called danger-associated molecular patterns (DAMPs). For instance, some pathogens use polygalacturonases to degrade the host cell wall and facilitate penetration. Polygalacturonases cleave homogalacturonan into oligogalacturonide, which are DAMPs recognised by the PRR WAK1 (Lorenzo et al., 2011).

PRR activation induces ion fluxes, oxidative burst, mitogen-activated protein kinases (MAPKs) activation, changes in gene expression and in protein phosphorylation, receptor endocytosis and callose deposition, resulting in local and systemic resistance (Couto & Zipfel, 2016). In general, PTI is sufficient to prevent infection, but in some cases the pathogen produces effectors to overcome PTI resulting in Effector-Triggered Susceptibility (ETS).

1.1.1.3 Effector-triggered susceptibility (ETS)

Effectors are pathogen-produced molecules contributing to virulence (Cornelis & Van Gijsegem, 2000). Many are delivered inside the plant cell via the Type III Secretion System (T3SS) of bacteria or the haustorium of filamentous pathogens (Cornelis, 2006; Yoshida et al., 2016). For instance, the effector AvrPphB is a bacterial T3SS effector from the bacteria *Pseudomonas syringae* that can cleave host kinases involved in PTI, such as BIK1, PBL1 and PBL2 (Zhang et al., 2010). Some effectors are also delivered into the apoplast (Kamoun, 2006). The oomycete *Phytophthora sojae* can secrete glucanase inhibitors in the apoplast of soybean to block the action of the host glucanases EGaseA and EGaseB, which otherwise degrade the pathogen cell wall (Rose et al., 2002). Beside inhibition of host defence, effectors can target nutritional pathways. One example is the transcriptional hijacking of the rice sugar transporter-encoding gene SWEET11 by the transcriptional activator-like (TAL) effector PthXo1 from the bacteria *Xanthomonas oryzae* (Chen et al., 2010). These examples illustrate the virulence function of effectors.

1.1.1.4 Effector-triggered immunity (ETI)

Effectors can be recognised by R-proteins, resulting in effector-triggered immunity (ETI). Most of the plant R-genes cloned to date belong to a class of immune receptors called NLR (nucleotide binding domain, leucine-rich repeat) (Jones et al., 2016). NLR-encoding genes are also found in genomes of animals and fungi. In plants, their activation results in
induction of PR (Pathogenesis-Related) proteins, salicylic acid (SA) production, transcriptional reprogramming and often leads to cell death at infection sites. This cell death, called the Hypersensitive Response (HR) helps to stop the propagation of the pathogens.

Some effectors can be recognised by non-NLR type of R-proteins. For instance, the effector AvrBs3 from Xanthomonas campestris triggers HR in Nicotiana benthamiana in presence of the pepper flavin monoxygenases Bs3 (Römer et al., 2013). Thus, Bs3 is an R-protein despite not being an NLR. The small non-NLR R-protein RPW8 is necessary and sufficient for resistance against powdery mildew fungi (Xiao et al., 2001).

1.1.1.5 Beyond the model

Some pathogen elicitors/host protein pairs do not fall in one or the other category. For instance, the PRR Cf-2 from tomato recognises the effector Avr2 from the fungus Cladosporium fulvum. Thus, the categorization of this pair within PTI or ETI can be discussed. As Avr2 is an effector, it activates ETI; as Cf-2 is a PRR, it activates PTI. Some models have proposed to replace the dichotomy PTI/ETI by a continuum of immunogenic proteins (Thomma et al., 2011; Cook et al., 2015) or by a nine-category classification (Kourelis & Van Der Hoorn, 2018). The zig-zag-zig (i.e. PTI/ETS/ETI) model will be used as a conceptual framework for this thesis, focusing on the ETI mediated by NLRs.

1.1.2 NLRs: Structure

1.1.2.1 Evolution

Since the cloning of the first R-gene, Hm1 from Maize, 25 years ago (Johal & Briggs, 1992), 323 other immunity-activating protein have been described so far in plants (Kourelis & Van Der Hoorn, 2018). 191 of them (60.8 %) are NLRs (Kourelis & Van Der Hoorn, 2018). These immune receptors are also found in animals and fungi, although they emerged convergently (Jones et al., 2016). In plants, NLRs emerged before the charophyte-embryophyte separation, i.e. before the colonization of the land (Gao et al., 2018; Han, 2018).

Canonical plants NLRs are composed of three domains: (i) an N-terminal Toll, Interleukin 1 receptor, Resistance protein (TIR), Coiled-Coil (CC) or Resistance to Powdery Mildew 8 (RPW8) domain, (ii) a central Nucleotide-Binding adaptor shared by Apaf-1, certain R-gene products and CED-4 (NB-ARC) domain followed by a (iii) Leucine-rich repeats (LRR)
domain. Several NLRs show variations such as domain duplication, domain truncation and/or fusion of extra domain. Animal NLRs usually have a NAIP, CIITA, HET-E and TP1 (NACHT) central domain and a pyrin (PYD) or a caspase recruitment (CARD) N-terminal domain (Jones et al., 2016). Fungal NLRs harbour a central NACHT or NB-ARC domain bordered by variable N-terminal (HeLo-like, Goodbye-like, sesB-like, PNP_UDP or HET) and C-terminal (WD40, Ankyrin or tetratricopeptide repeats) domains. Unlike in animals and plants, none of the 5616 fungal NLRs reported carry LRRs (Dyrka et al., 2014).

Based on their conserved central NB-ARC domain, there are three monophyletic classes of plant NLRs. Interestingly, each class is associated with TIR, CC or RPW8 N-terminal domain. The three classes are thus called TIR-NLRs, CC-NLR and RPW8-NLRs (Shao et al., 2014). Each plant NLR can be assigned to a class based on its NB-ARC domain and independently of its N-terminal domain. For instance the NB-ARC domain of ADR1-L3 from Arabidopsis (Arabidopsis thaliana) enables to be classified as an RPW8-NLR although it lacks an RPW8 N-terminal domain (Bonardi et al., 2011; Andolfo et al., 2014). In general, RPW8-NLRs are conserved and low copy number in all angiosperms, TIR-NLRs are diverse and expanded in rosids and CC-NLRs are diverse and expanded in asterids and monocots (Shao et al., 2016). TIR-NLRs are found in charophytes, bryophytes and in the basal angiosperm Amborella trichopoda but have been lost in monocots (Shao et al., 2014; Gao et al., 2018).

1.1.2.2 NB-ARC central domains

NLR activity can result in HR. Interestingly, there is homology between a specific region of plant NLRs and the cell death regulators CED-4 (from Caenorhabditis elegans) and Apaf-1 (from Homo sapiens). This related region is the NB-ARC domain (Van Der Biezen & Jones, 1998a). NB-ARC consists in three subdomains: NB, ARC1 and ARC2 (Takken et al., 2006). Several sequences are highly conserved among NB-ARCs and are crucial in NLR function. These motifs are P-loop (or Walker A), RNBS-A, kinase-2 (or Walker B), RNBS-B, RNBS-C, GLPL, RNBS-D and MHD (Van Der Biezen & Jones, 1998a; Takken et al., 2006; van Ooijen et al., 2008; Shao et al., 2016). For instance, mutations in the RNBS-A, Walker B or MHD motifs of the tomato NLR I-2 (S233F, D283E and D495V respectively) result in autoactive gain-of-function alleles (Tameling et al., 2006). A K207R mutation in the P-loop of the tomato NLR I-2 results in a loss-of-function allele, which is epistatic to the RNBS-A, Walker B and MHD autoactive mutations (Tameling et al., 2006).
1.1.2.3 LRR domain

LRR domains are found towards the C-terminus of NLRs, and additional domains can be found C-terminal to the LRRs. Each repeat consists of 20-29 amino acids including a consensus motif LxxLxLxxN/CxL. LRRs are involved in protein-protein interactions (Kobe & Kajava, 2001). In NLRs, LRRs are structurally irregular with varying repeat lengths and non-canonical LRR motifs (Takken & Goverse, 2012). Two LRR functions are proposed: (i) effector recognition and (ii) maintenance of NLR inactivation in the absence of pathogen (Takken & Goverse, 2012).

1.1.2.4 N-terminal domains

TIR domains

TIR domains are found in plant NLRs and in animal Toll, Toll-like and Interleukin-1 receptors. In both cases, they play a role in immunity; although this function is likely a convergence (Ausubel, 2005). In plants, TIR domains are thought be the signalling domain of TIR-NLRs because their overexpression often causes cell death (Nimma et al., 2017). Dimerization is required for signalling. Structural analysis revealed that the TIR-TIR interaction of the Arabidopsis TIR-NLRs SNC1, RPP1 and RPS4 involves the α-helices A and E (AE interface) and the α-helix D and E (DE interface) (Zhang et al., 2016a). At least for RRS1/RPS4, the integrity of both interfaces is required for cell death induction (Williams et al., 2014; Zhang et al., 2016a). Both AE and DE interfaces are required for signalling of RBA1, a plant TIR-only protein (Nishimura et al., 2016). In animals, TIR-containing protein activation involves TIR-TIR interactions via different interfaces for different proteins (Akira & Takeda, 2004; Nimma et al., 2017).

Since RBA1 is a TIR only protein, its TIR domain is necessarily involved in the direct or indirect recognition of its cognate effector HopBA1, from P. syringae. Moreover, the TIR domain of the tobacco NLR N is necessary and sufficient to interact with the p50 viral elicitor (Burch-Smith et al., 2007). This indicates that TIR domains are not only involved in signalling but can also play a role in effector recognition.

The TIR domain of the mammalian protein SARM1 has intrinsic enzymatic activity (Essuman et al., 2017). It directly degrades NAD (Nicotinamide adenine dinucleotide) resulting in axonal degeneration. The glutamic acid E642 is conserved between SARM1, other TIRs and other NADases and is required for SARM1 function. A glutamate at this position is required for function of RPS4 (Swiderski et al., 2009; Sohn et al., 2014). Bacterial
and archaeal TIR domains also have an NADase activity (Essuman et al., 2018). Whether plant TIR domains have intrinsic enzymatic activity is not known.

**CC domains**

Coiled-coil domains are found in various proteins such as leucine zipper transcription factors, tropomyosin and plant R-proteins. They comprise two or three α-helixes folded to form a superhelix. One of their features is to be able to form homo- and hetero-dimers or oligomers (Lupas, 1996). For example, the CC domains of MLA10 from barley and ZAR1 from Arabidopsis can form homomultimers, an association required to trigger cell death (Maekawa et al., 2011; Baudin et al., 2017). The CC domain of the maize NLR Rp1 also induces cell death in N. benthamiana (Wang et al., 2015a). Like the TIR domains, NLR CC domains can be involved in cell death signalling.

An EDVID motif has been identified in CC domains of NLRs (Rairdan et al., 2008). At least in the CNL Rx from potato, the EDVID motif is involved in interaction with the NB-ARC domain. Correspondingly, the CC domain of Rp1 can interact with the NB-ARC domain, suppressing the CC-induced cell death (Wang et al., 2015a).

In addition, some CC domains have been shown to interact with other proteins. For example, the CC domains of RPS5 from Arabidopsis can interact with its guardee PBS1 (Ade et al., 2007) and the CC domain of MLA10 can interact with some WRKY transcription factors (Shen et al., 2007). The CC domain of the wheat NLR Lr10 is under diversifying selection and therefore thought to be involved in effector recognition (Loutre et al., 2009).

NLR CC domains can interact with themselves, other CC domains, other subdomains from the NLR, different endogenous proteins and perhaps with effectors. This suggests that there may be not a single model but several modes of signalling.

**RPW8 domains**

RPW8 was first characterised as a small R-gene involved in powdery mildew resistance (Xiao et al., 2001). Arabidopsis accession Ms-0 contains two functional paralogs: RPW8.1 and RPW8.2, absents in the susceptible accession Col-0. Instead, Col-0 contains four paralogs: HR1, HR2, HR3 and HR4. They quantitatively contribute for a weak immune response to powdery mildew, although Col-0 remains largely susceptible (Berkey et al., 2017). During infection, RPW8 is targeted to the extra-haustorial membrane, likely directed by a predicted signal anchor at its N-terminus (Wang et al., 2009).
Some NLRs have an N-terminal domain that resembles RPW8. Although RPW8 may have a coiled-coil structure, the RPW8-NLR clade is not contained in, but sister to the CC-NLR clade (Xiao et al., 2001; Shao et al., 2014). RPW8 domains are also characterised by the absence of an EDVID motif in their N-terminal domain (Collier et al., 2011). The most characteristic motifs of RPW8-NLRs are found in the NB-ARC domain. Their RNBS-B is often LV(T/V)SR, while it is often xxTTR in TIR- and CC-NLRs (Shao et al., 2016).

Compared to the TIR-NLRs and CC-NLRs, the RPW8-NLR genes have not expanded intensively (Zhang et al., 2016b). For instance, in Arabidopsis, only five genes encode a RPW8-NLR: ADR1, ADR1-L1, ADR1-L2, NRG1A and NRG1B.

### 1.1.3 NLRs: Activation

#### 1.1.3.1 Effector recognition

NLR activation is initiated after direct or indirect effector recognition. For example, in Arabidopsis the recognition of ATR1 effector (from the oomycete *Hyaloperonospora arabidopsidis*, *Hpa*, the cause of downy mildew) was shown to be the consequence of a direct interaction with the TIR-NLR RPP1 (Krasileva et al., 2010).

On the other hand, the ‘guard model’ explains how NLRs can recognise indirectly an effector (Van Der Biezen & Jones, 1998b; Dangl & Jones, 2001). For example, the CC-NLR ZAR1 from Arabidopsis can trigger ETI after recognition of the effectors AvrAC, HopF2a (both from *P. syringae*) or HopZ1a (from the bacteria *Xanthomonas campestris*) (Lewis et al., 2013; Wang et al., 2015b; Seto et al., 2017). However, neither AvrAC nor HopZ1a can directly interact with ZAR1. In fact, ZAR1 ‘guards’ intermediate Receptor-Like Cytoplasmic Kinases (RLCKs). HopZ1a can acetylate the RLCK ZED1, which activates ZAR1 (Lewis et al., 2013). AvrAC can uridylylate PBL2 triggering its recruitment by a pre-complex formed by ZAR1 and the RLCK RKS1 (Wang et al., 2015b). The RLCK ZRK3 is required for HopF2a recognition by ZAR1 (Seto et al., 2017).

Avirulence factors such as AvrAC, HopZ1a and HopF2a are recognized and prevent infection. Unless they contribute to virulence in the absence of their cognate R-protein, they should be under negative selection pressure in pathogen genomes. Their conservation is consistent with the ‘decoy hypothesis’ (van der Hoorn & Kamoun, 2008). Decoys are host proteins that evolved, from an effector target or any host gene, to mimic an effector target, so that they can recognize an effector action and enable ETI. For
instance, PBL2 is a paralog of BIK1, which is uridylylated by AvrAC to suppress immunity. PBL2 mimics BIK1, acting as a decoy target of AvrAC to triggers immunity via ZAR1 (Wang et al., 2015b). The decoy terminology implies (i) that recognized effector has a virulence function in the absence of cognate R-protein and (ii) that the decoy protein, unlike the authentic effector target, does not contribute to virulence (van der Hoorn & Kamoun, 2008). Since AvrAC promotes infection in the absence of ZAR1, PBL2 or RKS1 and PBL2 uridylylation does not contribute to virulence, PBL2 is a good example of a decoy in Arabidopsis. Another is Pto, which is guarded by the tomato NLR Prf and mimics the AvrPto target FLS2 (van der Hoorn & Kamoun, 2008).

1.1.3.2 Effector-independent activation

Gain-of-function NLR alleles have been characterized in Arabidopsis. They constitutively activate ETI responses such as cell death and PR gene activation. They enable to study NLR signalling without the use of an exogenous effector.

For instance, suppressor screens of npr1 (which lacks expression of PR genes) revealed gain-of-function alleles of the TIR-NLR encoding genes SSI4 and SNC1 (Li et al., 2001; Shirano et al., 2002). Many NLR encoding genes were identified in an EMS mutagenesis screen for chilling sensitive (chs) mutants (Schneider et al., 1995). These mutants grow normally at 22 °C but show ETI-like responses at 16 °C. CHS1 encodes a TIR-NB protein, CHS2 encodes the TIR-NLR RPP4 and CHS3 encodes a non-canonical TIR-NLR having an extra C-terminal LIM domain (Huang et al., 2010; Yang et al., 2010; Wang et al., 2013a). All three are gain-of-function mutations. The reason why some R-proteins are activated at cold temperature is unclear. chs3-1 renders the plant more resistant to a freezing treatment (-6 °C for two to four hours) compared to wild type (Yang et al., 2010). This suggest that NLRs can play a role in freezing tolerance, independently of their role in plant-pathogen interaction. A suppressor screen of chs3-2D, another gain-of-function allele of CHS3, identified the TIR-NLR CSA1 as required for CHS3 function (Xu et al., 2015). Similarly, the activity of CHS1 requires the TIR-NLR SOC3 (Zhang et al., 2016c). SOC3 and CHS1 are also known to guard the homeostasis of the E3 ligase SAUL1, which might be a pathogen target (Tong et al., 2017). SOC3 pairs with CHS1 to guard the absence of SAUL1 function and pairs with TN2 to guard the over-activity of SAUL1 (Liang et al., 2018b).

Autoimmune phenotypes have also been observed in crosses between wild type plants resulting in hybrid necrosis. For example, DM1 (a wild type allele of the TIR-NLR SSI4) was identified as incompatible with DM2d (a wild type allele of the TIR-NLR RPP1) (Chae et al., 2001).
A cross between the accession Uk1 and Uk3 carrying one or the other of these alleles results in hybrid necrosis. The study of incompatible R-gene loci also revealed incompatibility between some alleles of RPW8 and some other alleles of the CC-NLR RPP7 (Chae et al., 2014).

Characterization of autoimmune phenotype can be misleading. CAMTA3 is a transcription factor that can bind the promoter of some defence-related genes such as EDS1 (Du et al., 2009). Since the camta3 mutant displays an autoimmune phenotype and many ETI marker genes are upregulated in the mutant, it was proposed that CAMTA3 is a transcriptional repressor of these ETI-related genes (Galon et al., 2008; Prasad et al., 2016). In fact, CAMTA3 activity is guarded by the TIR-NLRs DSC1 and DSC2. The phenotype and the gene activation observed are the result of the DSC1 and DSC2 activation rather than the loss of transcription activity of CAMTA3 (Lolle et al., 2017).

1.1.3.3 NLR pairs

Several plant NLRs have been found to work in pairs, with both proteins required for signalling. For instance, SOC3 and CSA1 are required for the signalling of CHS1 and CHS3 (Xu et al., 2015; Zhang et al., 2016c). In Arabidopsis accession Ws-2, RRS1/RPS4 form another TIR-NLR pair required for the recognition of the effectors AvrRPS4 and PopP2, from the bacteria Pseudomonas and Ralstonia respectively (Narusaka et al., 2009). Its paralog pair RRS1B/RPS4B also recognizes AvrRPS4, but not PopP2 (Saucet et al., 2015). Interestingly, the functional pairs CHS3/CSA1, CHS1/SOC3, RRS1/RPS4 and RRS1B/RPS4B also form physical pairs in the genome, linked in a head-to-head orientation (Narusaka et al., 2009). There are other TIR-NLR pairs found in a head-to-head orientation in Arabidopsis, but they have not been functionally characterized. The TIR-NLR-paired encoding genes constitute two monophyletic groups (Andolfo et al., 2014). This indicates that the current TIR-NLR pair repertoire is the result of duplication events from an original pair. Some CC-NLR genetic and functional pairs have also been described such as Pik-1/Pik-2 and RGA4/RGA5 in rice, involved in the recognition of AVR-Pik and AVR1-CO39 from the blast fungus Magnaporthe oryzae (Kanzaki et al., 2012; Cesari et al., 2013).

Some genes have been shown to function in pairs without being clustered in a head-to-head orientation. RPP2A and RRP2B encode two TIR-NLRs both required for resistance against Hpa race Cala2 and are clustered in a head-to-tail orientation (Sinapidou et al., 2004). RPW8 and the CC-NLR encoding cluster RPP7, as well as some SSL4 and RPP1 alleles, are involved in hybrid necrosis, indicating that they can induce immunity as a pair (Chae
et al., 2014). However, they are physically unlinked. The TIR-NLR N from tobacco confers resistance to TMV, and requires the RPW8-NLR NRG1 but their encoding genes are not linked in the tobacco genome (Peart et al., 2005). Roq1 and RPP1 are unrelated TIR-NLRs that also require NRG1 to signal (Qi et al., 2018). Similarly, the unrelated NRLs RPS2, RPP4, SNC1, UNI-1D, CHS3 and RPP1 require function of ADR1 to signal. ADR1s are RPW8-NLRs from the sister clade of NRG1 (Bonardi et al., 2011; Dong et al., 2016).

1.1.3.4 Sensor and helper NLRs

The requirement of an NLR by several NLRs is explained by the “sensor-helper” model (Bonardi et al., 2012). Some diverse NLRs “sense” the infection and are “helped” by an NLR which executes the immune response. ADR1 and NRG1 are two examples of helper NLRs.

Similarly, NRC proteins are CC-NLR helpers in Solanaceae (Wu et al., 2016). They redundantly contribute to the resistance mediated by the sensor NLRs Rpi-blb2, Mi-1.2, Sw5b, R8, R1, Prf, Rx and Bs2 (Wu et al., 2017). Interestingly, all these NRC-dependent sensors are part of a monophyletic group that emerged over 100 Mya and is broadly distributed among Asterids. Since the NRC helper clade and the NRC-dependent sensor clade are monophyletic, they form a network expanded from a single sensor-executor pair. The sensor clade is more diversified than the NRC clade. Sensors have expanded to recognize effectors from many pathogens and helpers are more constrained as immune signalling nodes (Wu et al., 2017).

In Arabidopsis, DM1 (an allele of SSI4 from Arabidopsis accession Uk-3) and DM2d (an allele of RPP1 from Arabidopsis accession Uk-1) can interact to trigger HR (Tran et al., 2017). DM1- and DM2-related NLRs might constitute a local sensor-executor network within Arabidopsis.

1.1.4 NLRs: Signalling

1.1.4.1 Conformational change model

The NB-ARC domain belongs to the superfamily of nucleotide-binding domains signal transduction ATPases with numerous domains (STAND), which also includes NACHT from animal NLRs (Leipe et al., 2004). The STAND protein MalT from Escherichia coli was used to describe the ADP/ATP switch model. MalT switches between an ADP-bound resting form and an ATP-bound active form that oligomerises. The ATPase activity of MalT is not
required for transcription activation. In fact, the Walker B hydrolyses ATP to ADP to reset MalT into its ADP-bound off state (Marquenet & Richet, 2007). It was proposed that this model could apply for all STAND proteins. Structural data of the NB-ARC protein Apaf-1 supports this model and indicate that the P-loop is involved in nucleotide binding (Cheng et al., 2016). In plant NLRs, P-loop mutants are in general loss-of-function and Walker B mutants are in general autoactive (Zhang et al., 2017). It supports the idea that P-loop is involved in ATP binding and required to switch to the active form, while Walker B is involved in ATP hydrolysis required to switch back from active to ADP-bound inactive form. So far there are no structural data available to confirm this model for plant NB-ARCs.

The effectors PopP2 and AvrRps4 interact with the C-terminal of the Arabidopsis NLR RRS1, resulting in activation of immunity (Ma et al., 2018). Similarly, the C-terminal domain of the NLRs RPP1 (from Arabidopsis) and Roq1 (from N. benthamiana) physically interacts with the cognate effectors ATR1 (from Hpa) and XopQ (from Xanthomonas) to trigger immunity (Krasileva et al., 2010; Qi et al., 2018). On the other hand, TIR, CC and RPW8 domains are in general sufficient to activate an immune response on their own (Bernoux et al., 2011; Collier et al., 2011; Bai et al., 2012). Thus it was proposed that, for plant NLRs, the C-terminal domain is involved in sensing, NB-ARC is involved in switching via conformational change and the N-terminal domain is involved in signalling (Collier & Moffett, 2009). For the Solanaceae NLRs N and Rx, elicitor sensing is regulated by the N-terminal domain (Burch-Smith et al., 2007; Rairdan et al., 2008), indicating that pathogen detection by NLRs is more complex than in this proposed model. The signalling role of the N-terminal domain is clearly established for all the mammalian NLRs described so far (e.g. NLRP1, NLRP3 and NLRC4) (Lechtenberg et al., 2014; Broz & Dixit, 2016), the fungal NLR NWD2 and PLP-1 (Riek & Saupe, 2016; Heller et al., 2018) and for many plant NLRs including RPS4, RPP1, At4g19530, NRG1, ADR1, MLA10, Rx and RPS5 (Ade et al., 2007; Rairdan et al., 2008; Swiderski et al., 2009; Maekawa et al., 2011; Collier et al., 2011).

1.1.4.2 In animals and fungi

Animal NLRs

Animal NLRs can initiate an immune response often culminating in a form of regulated programmed cell death (PCD) called pyroptosis, which resembles the plant HR (Duxbury et al., 2016).

Pyroptosis is driven by specific caspases (e.g. caspases 1, 4 and 5 in humans) and initiated by inflammasomes (Vande Walle & Lamkanfi, 2016). At least three NLRs can form
inflammasomes upon activation: NLRP1, NLRP3 and NLRC4. They exist as monomeric surveillance devices in the cytosol. After recognition of an effector from the bacteria Bacillus anthracis for NLRP1, a Type 3 Secretion System (T3SS) subunit or flagellin for NLRC4 (the stimulus of NLRP3 is not known), they oligomerize to form a ring-shape complex: the inflammasome (Broz & Dixit, 2016). The inflammasome activates inflammatory caspases, resulting in the cleavage of gasdermin D. The N-terminal fragments of gasdermin D oligomerize at the plasma membrane to form pores (Liu et al., 2016b). This plasma membrane permeabilization characterises the pyroptosis but is also observed during necroptosis but not during apoptosis (Vande Walle & Lamkanfi, 2016). Similarly, Apaf-1 can oligomerize after cytochrome c activation to form a ring-shaped apoptosisome (Zhou et al., 2015).

**Fungal NLRs**

Fungi can recognise non-self during cell fusion between two genetically distinct individuals and respond with cell death. This rejection is called heterokaryon incompatibility and is genetically controlled by het (heterokaryon incompatibility) loci. Interestingly, some HET proteins (e.g. HET-D, HET-E, HET-R, NWD2) are NLRs with an N-terminal HET domain (Paoletti & Saupe, 2009).

The NWD2/HET-S system is one of the best-described incompatibility system in fungi (Daskalov & Saupe, 2015; Riek & Saupe, 2016). NWD2 is an NLR with an N-terminal HET domain, a central NACHT domain and a C-terminal WD-repeats domain. HET-S is a non-NLR two-domain protein with an N-terminal executor HeLo domain and a C-terminal prion forming domain. The prion-forming domain can adopt an amyloid β-solenoid fold (prion form) leading to a refolding of the HeLo domain which then acquires a toxic pore-forming activity at the plasma membrane. The amyloid templating of HET-S can be induced independently by [HET-s], an incompatible allele of HET-S that lacks the pore-forming activity, or the NLR NWD2. When activated by a specific ligand at the WD40 repeats region, NWD2 induces the prionisation of HET-S that leads to the pore-forming activity of its HeLo domain. This system illustrates the role of an NLR in a PCD-mediated defence strategy in fungi. Similarly, the fungal NLR PLP-1 triggers cell death upon recognition of SEC-9 alleles from incompatible strains (Heller et al., 2018).

PNT1/HELLP is a Chaetomium globosum pair, homolog of NWD2/HET-S (Daskalov et al., 2016). Like HET-S, HELLP has a HeLo N-terminal domain, can form amyloid fibrils, is membrane-targeted and has a cell death-inducing activity. The HeLo executor domain of
HELLP and HET-S shares sequence similarities with RPW8 and NRG1 from plants (Daskalov et al., 2016). This suggests that RPW8 and/or RPW8-NLRs might trigger HR via a related pathway.

1.1.4.3 In plants: Downstream signalling

Plant NLR-mediated resistance requires many proteins and hormones. NDR1 and EDS1 are two important proteins acting downstream of NLRs (Glazebrook et al., 1997). CC-NLR signalling is often dependent on NDR1 while TIR-NLR signalling is often dependent on EDS1 (Century et al., 1995; Parker et al., 1996). The apparent specific requirement of EDS1 by TIR-NLRs and NDR1 by CC-NLRs is not strict. For instance, EDS1 regulates RPP8- and RPS2-mediated immunity, redundantly with SID2 (Venugopal et al., 2009).

EDS1 is a nucleocytoplasmic lipase-like protein, which interacts with its sequence-related partners PAD4 or SAG101 (Wagner et al., 2013). The mechanism of EDS1 has not been fully solved yet. It can be required for ROS burst, HR, gene induction and SA accumulation (Wiermer et al., 2005; Cui et al., 2017, 2018). PAD4 and SAG101 interact with EDS1 at the same interface, suggesting co-existence of both EDS1-PAD4 and EDS1-SAG101 complexes (Wagner et al., 2013).

NDR1 plays a role in the plasma membrane/cell wall adhesion during infection (Knepper et al., 2011). It regulates expression of genes from LEA and PIP gene families. LEA and PIP proteins control fluid loss and entry, suggesting a role of NDR1 in electrolytes leakage during infection (Knepper et al., 2011).

1.1.4.4 In plants: Phytohormone signalling

Some plant hormones, such as jasmonic acid (JA), Salicylic Acid (SA) and Ethylene (ET) are reported to regulate plant immunity (Bari & Jones, 2009). Typically, SA is associated with resistance against biotrophic / hemibiotrophic pathogens, whereas JA and ethylene are associated with resistance against necrotrophic pathogens (Buscaill & Rivas, 2014).

SA is a positive regulator of plant immunity. SA treatment in tobacco enhances resistance to TMV while overexpression of NahG, encoding a salicylate hydroxylase that catabolizes SA, compromises immunity (Durner et al., 1997). SA accumulates at infection points but also contributes to systemic resistance. Accumulation in non-infected tissues mediates and enhances a long-lasting resistance to secondary pathogen challenge, called Systemic Acquired Resistance (SAR) (Durrant & Dong, 2004). NPR1 is the central protein required for SAR and its activity is regulated by SA (Yan & Dong, 2014). The related proteins NPR3
and NPR4 are also SA receptors. In high SA, NPR1 interacts with TGA transcription factors to activate SA-responsive genes, such as PR1. In low SA, NPR3/NPR4 play the opposite role by interacting with other TGAs to repress gene expression (Ding et al., 2018). This SA regulation is described in Arabidopsis immunity but may not be generalised to all plants. In rice, the basal level of SA is higher than in Arabidopsis and defence gene activation by SA is not reported. However, NPR1 is present and regulates immunity in rice (Chern et al., 2005). The roles of SA in rice immunity remain to be explored.

Jasmonate related metabolites, including the active form jasmonyl-isoleucine (JA-Ile), are lipid-derived compounds rapidly synthesized upon pathogen attack and leaf damage (Pieterse et al., 2012). JAZ proteins, MYC2 and COI1 are central for JA-triggered immunity (Bari & Jones, 2009). JAZ proteins can bind to transcription factors, such as MYC2, 3 and 4 and EIN3/EIL1, to block their activities (Campos et al., 2014). JA-Ile produced upon infection can bind COI1, resulting in JAZ protein degradation and de-repression of the transcription factors. Release of EIN3/EIL1 activates PDF1.2 expression (indirectly, through activation of intermediate transcription factors EFR1 and ORA59). Release of MYC2 directly activates transcription of VSP292. Typically, ERF1-activated genes are associated with necrotrophic resistance and MYC-activated genes are associated with resistance against herbivores (Pieterse et al., 2012).

JA and SA pathways seem antagonistic: activation of one pathway leads to repression of the other (Bari & Jones, 2009). The hemibiotrophic bacterium P. syringae produces a JA analogue, called coronatine, to repress the SA/biotrophic resistance pathway (Zhao et al., 2003). However, both SA and JA accumulate after activation of the Arabidopsis NLR RPS2, activated by AvrRpt2 from P. syringae. This suggests that JA and SA could interplay to maintain resistance against necrotrophic pathogens while building an effective defence against a biotrophic pathogen (Liu et al., 2016a).

Hormones are also involved in priming. Priming occurs during infection to prepare plants for secondary pathogen challenge. Primed plants respond to infection faster and stronger than non-primed ones. Priming correlates with specific chromatin marks, e.g. on some promoters, which have been reported to be inherited to the next generation. SA, JA, piperolic acid and azelaic acid are indispensable for priming (Conrath et al., 2015; Mauch-Mani et al., 2017).
1.2 *Albugo candida*, a plant pathogen oomycete

1.2.1 Taxonomy

*A. candida* causes white rust of Brassicaceae. It is a filamentous biotrophic pathogen. It can affect both wild and crop species such as *Arabidopsis*, cabbage (*Brassica oleracea*), radish (*Raphanus sativus*) and brown mustard (*Brassica juncea*) (Pidskalny & Rimmer, 1985; Holub et al., 1995).

*A. candida* belongs to the oomycetes, a group of filamentous non-photosynthetic eukaryotes. Oomycetes used to be classified as fungi, but cladistic methods placed them as a monophyletic group, independent of fungi and more related to brown algae (Beakes et al., 2012). Unlike fungi, oomycete hyphal walls contain cellulose and glucan but not chitin. Most of the oomycetes are marine organisms, obligate parasites of seaweeds, crustacea and nematodes. Others can be freshwater or soil saprophytes or insect, animal or plant parasites (Beakes & Sekimoto, 2008). Plant pathogen oomycetes includes the genera *Phytophthora*, *Hyaloperonospora* and *Albugo* (Kamoun et al., 2015). Among oomycetes, *A. candida* belongs to an independent group called Albuginales, related to Peronosporales, the group of *Phytophthora* and *Hyaloperonospora* (Thines, 2014). Albuginales contain only one family, the Albuginaceae, that encompasses three genera: *Albugo*, *Pustula* and *Wilsoniana* (Choi et al., 2009). At least 17 lineages have been defined within *A. candida*, each colonizing a specific range of host and having a distinct effector set (Jouet et al., 2018). Despite their unique host range, there is evidence of recombination between lineages (McMullan et al., 2015).

1.2.2 Life cycle

*A. candida* reproduces sexually by oospores and asexually by zoospores (Saharan et al., 2014). During the sexual cycle, an antheridium (male gamete, multinucleate) delivers its nuclei to an ooplasm (female gamete, mononucleated). The ooplasm nucleus fuses to one of the nuclei released by the antheridium, forming an oospore. In favourable conditions, oospores can germinate in short germ tubes that produce 40 to 60 zoospores each. Similarly, zoospores can form germ tubes that penetrate the leaf through wound or stomata. Then hyphae develop in the host mesophyll and produce asexual zoosporangia, each containing ~5 zoospores. Pressure applied by numerous zoosporangia on the lower
epidermis results in symptomatic white pustules. Increases in pressure result in cracks that release zoosporangia out of the leaf (Holub et al., 1995; Saharan et al., 2014).

In crop fields, A. candida reproduces asexually on its host and sexually at the end of the host growing season. Indeed, oospores are thick-walled and can survive in the absence of host during winter.

1.2.3 Impact and management

A. candida sporangia production results in symptomatic white pustules on the lower surface of the leaf. Due to this symptom, the resulting disease is called white rust disease. All members of Albuginales (i.e. Albugo, Pustula and Wilsoniana) can cause white rust symptoms. A. candida is characterized by its specificity to infect Brassicaceae (Thines et al., 2009), although it has been reported on some Cleomaceae and Capparaceae (Choi et al., 2009). Some other Albugo species, such as A. laibachii and A. koreana, can also grow on Brassicaceae (Thines et al., 2009). A. candida is distributed worldwide and can infect both wild species and crops including radish, cabbage, wasabi, indian mustard and turnip. 30-60 % of yield losses have been reported on rapeseed in Manitoba, Canada, 1971; 5-10 % in Australia, 1981; up to 89.8 % in an indian mustard, India, 1988 (Saharan et al., 2014). Moreover, plants infected by white rust are often co-infected by other pathogens such as the oomycetes Hpa and Phytophthora infestans, increasing yield losses (Cooper et al., 2008; Saharan et al., 2014; Belhaj et al., 2017; Prince et al., 2017).

There is no specific fungicide for white rust disease control, but some broad-spectrum fungicides such as Metalaxil, Mancozeb and Blitox can reduce white rust development (Saharan et al., 2014). Since A. candida grows ideally in cold and humid environments, keeping leaves away from moisture can also reduce white rust propagation (Saharan et al., 2014).

1.2.4 Virulence

1.2.4.1 CCG effectors

As a biotrophic pathogen, A. candida requires living plant tissues to complete its life cycle. Its effectors must be particularly specialized to maintain host cells alive and, at the same time, enable its propagation within plant tissues. Natural selection that maximises parasite fitness on one host species might lead to specialisation and reduced host range
In fact, *A. candida* evolved races, each specialised to infect a distinct host range (McMullan *et al.*, 2015; Jouet *et al.*, 2018).

Genome analysis of several *A. candida* races and the closely related species *A. laibachii* revealed a set of secreted proteins (Kemen *et al.*, 2011; Links *et al.*, 2011). Based on comparison with previously identified secreted effector families, it has been shown that the *A. candida* secretome contains 13 glycosyl hydrolases, 3 chitinases, 9 elicitors or elicitin-like, one CBEL-like, 6 Crinklers and 26 proteins that carry an RxLR motif (Links *et al.*, 2011). Surprisingly, the extensively described RxLR effector family (Morgan & Kamoun, 2007) is not expanded in the *A. candida* secretome. They are 26 RxLRs in *A. candida* compared to 563 in *P. infestans* (Links *et al.*, 2011). On the other hand, a new family of secreted proteins was identified in *A. laibachii* (Kemen *et al.*, 2011) and observed as well in *A. candida* (Links *et al.*, 2011). They carry a conserved CxxCxxxxxG motif and are consequently called "CCG" effector candidate proteins. There are ~100 CCGs in each *A. candida* race (Jouet *et al.*, 2018). Since CCGs are specific and abundant in *Albugo* species, they may be key effectors in white rust disease development. Currently, their function in virulence and avirulence is under investigation (Redkar *et al.*, in prep).

1.2.4.2 Immunosuppression

*A. laibachii* race Nc14 (an isolate of Acem1)-induced ETS not only facilitates its own colonisation, but also enables the growth of otherwise incompatible pathogens (Cooper *et al.*, 2008). Nc14 disrupts resistances against powdery mildew (via RPW8), *Hpa* races Cal2 (via RPP2) and Hiks1 (via RPP7) and the oomycete *Phytophthora infestans* (Cooper *et al.*, 2008; Belhaj *et al.*, 2017). In contrast, it does not affect the growth of already compatible pathogens such as the fungus *Colletotrichum higginsianum* (Cooper *et al.*, 2008).

*A. laibachii* imposes a reduction of tryptophan-derived antimicrobial metabolites in the host, which explains partially the immunosuppression (Prince *et al.*, 2017).

Despite the host specificity of each race, recombination has occurred between races (McMullan *et al.*, 2015). Immunosuppression by a compatible *A. candida* race could enable co-infection with an incompatible race, thus allowing sexual reproduction and emergence of a new race. Recombination of effectors in this new race could potentially facilitate the colonization of a new hosts (McMullan *et al.*, 2015).
1.2.5 **White Rust Resistance-genes discovery**

1.2.5.1 Using natural variation

White rust is a significant disease of *B. juncea*, a major crop in India. Although the cultivated lines Varuna, Pusa bold, Pusa jai kisan and Rohini are susceptible, some European lines such as Heera and Donskaja-IV are fully resistant. This variation was used to map and clone the CC-NLR encoding gene *BjuWRR1* from Donskaja-IV. *BjuWRR1* confers full resistance to white rust in Varuna, Pusa bold, Pusa jai kisan and Rohini (Arora *et al.*, 2019). However, the lack of genetic data for *B. juncea* impede the rapid cloning of WRR-genes from this species.

Most *A. candida* races will not colonize Arabidopsis, but if they do, they fail to grow on every Arabidopsis accession. This variation enables genetic analysis to identify White Rust Resistance (WRR)-genes in Arabidopsis, that could be transformed into crops.

The *Capsella bursa-pastoris*-infecting race AcEm2, the *Brassica juncea*–infecting race Ac2V and the *Brassica rapa*-infecting race Ac7V cannot grow on the Arabidopsis accession Col-0 but can trigger a weak chlorosis on cotyledons in Nd-1. Map-based cloning was used to clone the underlying resistance gene: the TIR-NLR WRR4A (Borhan *et al.*, 2008). The Col-0 allele of WRR4A confers resistance to a broad range of *A. candida* races including AcEm2, Ac2V, Ac7V and the *Brassica oleracea*-infecting race AcBoT. Arabidopsis WRR4A can be transformed into the brown mustard crop (*B. juncea*) providing full resistance to Ac2V (Borhan *et al.*, 2010).

Similarly, the race AcNc2 can grow on the Arabidopsis accession Ws-2 but not Col-0. A cross between Col-0 and Ws-2 highlighted the role of two WRR loci. One is the TIR-NLR pair WRR5A/WRR5B and the second is the atypical RPW8-NLR WRR7 (Cevik *et al.*, in prep).

Variation observed in white rust disease resistance among Arabidopsis accessions was successfully used to identify the broad-spectrum gene WRR4A. However, WRR4A resistance could theoretically be broken if deployed alone. Thus, it remains crucial to identify other WRR genes against the crop-infecting *A. candida* races.

1.2.5.2 Revealing novel WRR-genes using transgressive segregation

Genetic mapping of WRR genes is possible when both susceptible and resistant parents can be crossed. However, all the Arabidopsis accessions tested are resistant (at leaf stage) to Ac2V, Ac7V and AcBoT races (Cevik *et al.*, 2019). One reason could be the
presence of non-overlapping sets of resistance genes within each Arabidopsis accession. This hypothesis was tested using transgressive segregation. Multiparent Advanced Generation Inter-Cross (MAGIC) lines, deriving from 19 Arabidopsis parents (Kover et al., 2009), were screened for susceptibility to Ac2V. Although all 19 parents are resistant to Ac2V, two MAGIC lines (MAGIC23 and MAGIC329) are susceptible to this race due to transgressive segregation of WRR-genes (Cevik et al., 2019). These lines are also susceptible to Ac7V. However, they show weak or no susceptibility to AcBoT. Backcross of MAGIC329 and MAGIC23 with some of the MAGIC parents enabled to identify new WRR-genes. WRR4B from Ws-2 and Col-0, WRR8 from Sf-2 and WRR9 from Hi-0 are all TIR-NLR encoding genes that confer resistance to Ac2V.

An F2 population between MAGIC23 and MAGIC329 resulted in “Double-Magic” lines that are susceptible to AcBoT. Backcross of one AcBoT-susceptible “Double-Magic” line with the resistant parent MAGIC329 enabled the identification of a novel TIR-NLR involved in AcBoT resistance: WRR12 (Cevik et al., 2019).

### 1.3 CRISPR for genome editing

#### 1.3.1 CRISPR: a bacterial immune system

CRISPR genome editing methods derive from a bacterial immune system. Bacteria can be infected by viruses and carry an adaptive immune system to be “vaccinated” in case of secondary infection (Horvath & Barrangou, 2010). It is based on CRISPR (Clustered Regularly Inter-Spaced Palindrome Repeat) sequences and Cas (CRISPR-associated) genes. Cas proteins share similarities with various known proteins such as nucleases and polynucleotide-binding proteins. CRISPR loci comprise 23 to 47 bp repeats separated by highly polymorphic 21 to 72 bp spacer sequences. Spacers consist mostly of viral DNA fragments integrated after infection.

Several CRISPR immune systems (including Type I to Type VI) are reported in bacteria. They all function in three steps: (i) spacer acquisition, (ii) CRISPR component expression and (iii) cleavage of foreign nucleic acid (Karvelis et al., 2013). Spacer acquisition consists in integration of foreign DNA (typically from a bacteriophage) in CRISPR loci by Cas1 and Cas2. Then, CRISPR sequences are transcribed in long pre-crRNAs (precursor CRISPR RNAs). pre-crRNAs are cleaved into short crRNAs containing a single spacer flanked by CRISPR repeats. In CRISPR Type I, III, IV, V and VI, the single crRNA directly binds it cognate
endonuclease. In CRISPR Type II, each mature crRNA hybridizes with a tracrRNA (trans-acting CRISPR RNA) at a ~25-nucleotide interface that has complementarity between both crRNA and tracrRNA (Deltcheva et al., 2011). Each crRNA-tracrRNA forms a complex with a single protein from the Cas9 family. Complexed (Class I including Type I, III and IV) or individual (Class II including Type II, V and VI) ribonucleoproteins can cleave foreign DNA that shares complementarity with the crRNA spacer sequence (Makarova et al., 2017a,b).

To avoid cleavage of endogenous CRISPR sequences, ribonucleoproteins cleave spacer-similar DNA only when it is next to a specific PAM (protopspacer adjacent motif), e.g. "NGG" for Streptococcus pyogenes Cas9 (SpyCas9, all occurrences of 'Cas9', unless otherwise stated, will refer to 'SpyCas9' in the rest of the text) (Sander & Joung, 2014). Since spacers integrated in endogenous CRISPR regions are not next to a PAM, they are not cleaved by ribonucleoproteins. The Type VI CRISPR protein Cas13 (previously C2c2) has the particularity to cleave RNA instead of DNA. As in Type II and Type V, Cas13 mediated cleavage is processed by individual ribonucleoproteins (Gootenberg et al., 2017; Smargon et al., 2017).

The CRISPR immune system illustrates how integration of viral DNA during a primary infection leads to homologous DNA or RNA cleavage and resistance during secondary infection. In 2012, two groups proposed the use of CRISPR for genome editing application (Gasiunas et al., 2012; Jinek et al., 2012). Indeed, cleaved DNA is processed in vivo by repair mechanisms, sometimes introducing short insertions of deletions (indels). This method, so-called CRISPR system, was successfully applied to induce targeted mutations in vivo, in animal cells (Cong et al., 2013; Mali et al., 2013) and in plant cells (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013).

1.3.2 CRISPR for genome editing

1.3.2.1 Targeted mutagenesis

Type II, Type V and Type VI single ribonucleoproteins can cleave nucleic acid in vitro (Gasiunas et al., 2012; Jinek et al., 2012; Zetsche et al., 2015; Abudayyeh et al., 2016). Type II-based systems, using the endonuclease Cas9, are the most commonly used. Cas9-mediated cleavage requires presence of DNA template and each of three components of the ribonucleoprotein: Cas9, tracrRNA and crRNA (Gasiunas et al., 2012). They cause a DSB (double strand break) on a template DNA complementary to the crRNA if it is located at the 3′ of a PAM. The DSB occurs three nucleotides before the PAM.
To simplify CRISPR system, Jinek et al. engineered a tracrRNA/crRNA fusion by adding a linker loop, forming a single guide RNA (sgRNA, hereafter gRNA) (Jinek et al., 2012). They showed that the Cas9/gRNA complex is sufficient to induce DSB on targeted DNA template in vitro. In addition to the gRNA, the second central component of CRISPR system is Cas9, an endonuclease with two nuclease domains: RuvC and HNH. The crystal structure of Cas9 reveals a conformation in two lobes (Nishimasu et al., 2014). The recognition lobe is involved in interaction with the gRNA. The nuclease lobe contains both RuvC and HNH domains and executes the DSB. It also contains a PAM-interacting domain. Mutations in this domain modify the PAM recognition from NGG to NGAA, NGAC, NGAT or NGAG for Cas9-VQR (Kleinstiver et al., 2015) and to NG for xCas9 (Hu et al., 2018) and Cas9-NG (Nishimasu et al., 2018).

In vivo and particularly in planta, DSB are processed by two primary repair mechanisms: NHEJ (non-homologous end joining) and HDR (homology-directed recombination) (Puchta, 2005; Lieber, 2008). NHEJ can join two DNA ends, without homologous template and at any stage of the cell cycle. It is quick but error-prone, sometimes inducing small insertions or deletions (indels). The cNHEJ (classic NHEJ) is based on the KU70/KU80 complex, which binds and links two ends resulting from a DSB. More NHEJ pathways can take over in plants, including a-NHEJ (alternative -NHEJ), b-NHEJ (backup-NHEJ) and MMEJ (microhomology-mediated end joining) (Schmidt et al., 2019). In mouse, a polq-ku80 double mutant, i.e. impaired in the MMEJ and cNHEJ pathways, displays an increased frequency of HDR (Zelensky et al., 2017). It indicates that MMEJ and cNHEJ are the two major NHEJ pathways in mouse.

HDR occurs during late S and G2 phases of cell cycle, because it requires an undamaged chromatid as template to repair DSB on sister chromatid. It is precise but occurs much less frequently than NHEJ (Steinert et al., 2016).

NHEJ-induced indels can disrupt the coding DNA sequence (CDS) or cis regulatory elements. HDR can be used to insert a sequence through recombination between endogenous target and exogenous DNA donor template. CRISPR cleavage has been demonstrated in vivo in heterologous systems including human, mouse, frog, zebrafish, fruit fly, silkworm, bacteria, rice, tobacco and Arabidopsis (Sander & Joung, 2014).
1.3.2.2 Optimization in Arabidopsis

First reports
In 2013, CRISPR for genome editing was reported for the first time in Arabidopsis and *N. benthamiana* protoplasts (Li et al., 2013), in wheat protoplasts and rice stable lines (Shan et al., 2013) and *N. benthamiana* leaves (Nekrasov et al., 2013). In protoplasts and rice stable lines, authors reported indels at a gRNA target and a 48-nucleotide deletion using two gRNA targets. Furthermore, they observed HDR-mediated gene replacement after CRISPR-induced DSB, in presence of a donor template (Li et al., 2013; Shan et al., 2013). Indels were also observed in *N. benthamiana* leaves after transient expression of Cas9 and gRNA (Nekrasov et al., 2013).

Many biotechnological applications, including generation of knockout lines, require stable CRISPR-mediated genome editing, i.e. CRISPR events in germ cell line inherited to the progeny. In 2014, it was reported that, among the second generation of CRISPR-transformed Arabidopsis lines, some plants bear indels inherited from their parent (Feng et al., 2014).

Variation in efficiency
In plant, CRISPR typically involves the transgenic expression of Cas9 and one or several gRNA(s). The expression of Cas9 is regulated by an RNA-Polymerase II-dependent promoter. The gRNA expression is regulated by a RNA-Polymerase III-dependent promoter to avoid post-transcriptional modifications, such as poly-adenylation and capping on the gRNA. Whereas Feng et al. reported 27.4% of inherited events in T2 using the 35S (from *Cauliflower Mosaic Virus*) promoter to drive Cas9 expression (Feng et al., 2014), Fauser et al. obtained 8 to 70% of CRISPR-induced homozygous mutants in T2, from different T1 parents using *Ubi4-2* (from *Petroselinum crispum*) promoter (Fauser et al., 2014). By using one gRNA targeting *FT* and *ICU2* promoter to drive Cas9 expression, Hyun et al. recovered 1/34 heterozygous mutant among a T2 family (Hyun et al., 2014). Xing et al. tried to knockout three genes in a single multiplex assay. They obtained 52.9% of triple mutants from five independent T2 populations. In their experiment, Cas9 expression was driven by the 35S promoter (Xing et al., 2014). These methods for Arabidopsis reported contrasting efficiencies. One reason of such variability could be the use of different CRISPR components (Volpi e Silva & Patron, 2017).
Germline-specific promoters

Based on the hypothesis than Cas9 cis-regulation can result in variable CRISPR efficiency, Wang et al. and Mao et al. developed a germ line specific CRISPR system to increase inheritable events (Mao et al., 2015; Wang et al., 2015c). They expressed Cas9 under the control of EC1.1, EC1.2 or SPL promoters, specifically active in the female (EC1.1 and EC1.2) or male (SPL) germ line. These systems trigger less somatic activity in T1 but stable CRISPR-induced mutations in T2. Cas9 regulation by EC1.2 promoter resulted in nine triple homozygous mutants out of 108 tested T1 lines tested (Wang et al., 2015c). This system is the first to report stable CRISPR-induced mutations in the first generation after transformation with such efficiency. Interestingly in their experiment, Cas9 terminator seems to influence the mutation rate. Higher mutation rate was observed when Cas9 expression was regulated by E9 terminator (rbcS E9 from Pisum sativum) compared to the Nos terminator (Nopaline synthase from Agrobacterium tumefaciens). The idea of using a germline specific promoter to drive Cas9 expression has been further investigated. A high CRISPR activity has been reported using SPL (Mao et al., 2015), MGE1 (Eid et al., 2016) and YAO (Yan et al., 2015) promoters. Particularly, the YAO promoter can lead to 90.5% of somatic activity whereas, in the same conditions, the 35S promoter led to 4.3% of somatic activity. The RPS5a promoter is also shown to be extremely efficient to induce inheritable mutations (Tsutsui & Higashiyama, 2017).

gRNA transcription

The gRNA expression can also be crucial for CRISPR activity. Generally, the gRNA expressing cassette contains the U6-26 promoter, one gRNA and a seven-thymine repeat that constitutes an RNA-polymerase III terminator. U6-26 promoter is indeed the most active U6 promoter, although U6-1, U6-4, U6-5, U6-6 and U6-29 are also effective in Arabidopsis (Li et al., 2007). Nevertheless, RNA polymerase III-dependent promoters are not tissue specific and less diverse than RNA polymerase II-dependent ones. A strategy has been proposed to drive gRNA expression using RNA polymerase II, based on RNA self-cleavage. Hammerhead type and hepatitis delta virus ribozymes can be self-cleaved from RNA 5′- and 3′-ends respectively. Gao and Zhao proposed the use of RGR (ribozyme-gRNA-ribozyme) in CRISPR systems (Gao & Zhao, 2014). RGR can be transcribed by RNA polymerase II and the self-cleavage of ribozymes can release a desired gRNA. Similarly, the endogenous tRNA-processing system can cleave and release tRNA precursors. A tRNA-gRNA architecture enables the production of multiple gRNAs using a single promoter (Xie et al., 2015). Csy4 is an RNase of Pseudomonas aeruginosa that has a specific
recognition site. A single transcript of gRNAs flanked by Csy4 recognition sites can produce multiple gRNAs using a single promoter (Tsai et al., 2014). tRNA, RGR and Cys4 gRNA processing system can successfully trigger CRISPR-mediated disruption of targeted gene in vivo (Čermák et al., 2017).

**gRNA backbone**

In addition to Cas9 and gRNA expression regulations, the gRNA structure may enhance CRISPR activity (Jinek et al., 2013). Most of the CRISPR systems reported so far are based on a crRNA-tracrRNA fusion (Jinek et al., 2012). This gRNA coding sequence contains a succession of four thymines that can act as a terminator for RNA-polymerase III. Thus, Chen et al. proposed an improved version of this gRNA (Chen et al., 2013). First, they broke the succession of four thymines by doing an A>T transversion and secondly, they added a five base-pair extension in the stem, since Cas9 binds gRNA through its duplex stem (Nishimasu et al., 2014). As this gRNA has a nucleotide “flip” and an “extension”, it is referred as gRNA-EF. It was used by Chen et al. to direct a dead version of Cas9 (dCas9, bearing a D10A mutation in RuvC domain and H840A mutation in HMH domain, rendering inactive both nuclease domains) fused to KRAB, a human transcription repressor, in human cells. dCas-KRAB was able to repress transcription two-fold more when led by gRNA-EF compared to the classic gRNA (Chen et al., 2013). Three subsequent CRISPR assays conducted in mammals (Dang et al., 2015; Peng et al., 2015) and Caenorhabditis elegans (Ward, 2014) reported elevated mutation rates when gRNA-EF was used instead of classic gRNA.

**gRNA spacer**

Moreover, the spacer sequences of the gRNA may play a critical role in CRISPR efficiency. Three different gRNAs targeting AtBR11 triggered variable CRISPR activity, from 30% to 84.2%, when expressed in the same conditions and in same CRISPR system (Feng et al., 2014). Based on the crystal structure of Cas9, recognition of gRNA spacer should be sequence independent although Arg71 may induce a preference to guanine in the four nucleotides at 3’ end of the spacer sequence (Nishimasu et al., 2014). Prediction of CRISPR efficiency based on gRNA spacer sequence remains largely empirical. A comparative study of 1841 gRNAs in mammal cells revealed that some nucleotides are statistically favoured or disfavoured at certain positions, especially at 3’ end of the spacer sequence (Doench et al., 2014). It is not known if these observations are valid in plants.
Influence of temperature on Cas9 activity

LeBlanc et al. used a different approach to optimise CRISPR in Arabidopsis. Cas9 is originally a protein from *S. pyogenes*; its enzymatic activity optimum is 37°C *in vitro*. In Arabidopsis, a series of 37°C treatments of plants expressing Cas9 results in elevated activity, as compared to plant grown at 22°C (Le Blanc et al., 2017). The heat treatments not only increased the Cas9 activity but also gRNA expression.

1.3.2.4 Diversity of CRISPR endonucleases

Cas9 from *S. pyogenes* is widely used in CRISPR systems. However, other CRISPR endonucleases can be used.

Cas9 from *Staphylococcus aureus* (SaCas9) is smaller and equally or more efficient than SpCas9 (Steinert et al., 2015; Raitskin et al., 2018; Wolter et al., 2018). However, it has a more specific PAM: NNGGGT, which reduces the targeting range.

The Type V CRISPR endonuclease Cas12 (previously Cpf1) from *Lachnospiraceae bacterium* and *Acidaminococcus sp.* functions in human cells (Zetsche et al., 2015). Instead of the dual crRNA-tracrRNA recognized by Cas9, Cas12 uses a single and shorter crRNA. Also, Cas12 recognizes a TTTV PAM located at the 5' of the target and releases fragments with a five-nucleotide overhang after cleavage, whereas Cas9 produces blunt ends. The Cas12 system has been successfully applied for targeted mutagenesis in rice and tobacco (Endo et al., 2016; Hu et al., 2017; Tang et al., 2017).

Cas13a (previously C2c2), the endonuclease of CRISPR Type VI, has been discovered in 2015 and characterised in 2016 (Shmakov et al., 2015; Abudayyeh et al., 2016). Another CRISPR Type VI endonuclease, Cas13b, was characterized in 2017 (Smargon et al., 2017).

Unlike Cas9 and Cas12, Cas13 cleaves specifically RNA. Importantly, Cas13 can cleave collateral RNA molecules even if they do not share identity with the template crRNA-targeted RNA (Gootenberg et al., 2017).

The addition of Cas12 and Cas13 in the genome editing toolbox has facilitated the development of novel CRISPR applications, beyond the generation of targeted random mutations (Murugan et al., 2017; Schindele et al., 2018).

1.3.2.5 Genome editing and beyond

NHEJ-mediated gene editing

The imprecise repair of Cas9-targeted DSB by the NHEJ have been extensively use to generate null alleles (Hsu et al., 2014). It enabled recapitulation of key events of tomato
domestication in only few months (Li et al., 2018; Zsögön et al., 2018), generation of transgene-free powdery mildew resistant tomato plants (Nekrasov et al., 2017) and correction of a heart genetic disease in mice during early postnatal stage (Xie et al., 2016) and in human zygote (Ma et al., 2017). The precise RNA-guided targeting of Cas proteins to nucleic acid can also lead to much diverse applications (Langner et al., 2018; Schindele et al., 2018).

Transcriptional regulation
RuvC and HNH, the two nucleases of Cas9, can be mutated to obtain a nuclease dead Cas9 (dCas9). A fusion of the demethylase TET1CD or the acetyltransferase p300 to dCas9 successfully up-regulated the transcription of targeted gene (Hilton et al., 2015; Choudhury et al., 2016). In addition, transcriptional repressors such as KRAB and the methyltransferase DNMT3A were fused to dCas9 to hypermethylate targeted gene promoters, resulting in transcriptional repression (Thakore et al., 2015; Vojta et al., 2016).

Base editing
In order to generate precise mutations at Cas9-binding site, two groups engineered a fusion of Cas9 with the human cytidine deaminases APOBEC or AID (Komor et al., 2016; Nishida et al., 2016). Such cytidine base editor (CBE) modules catalyse the deamination of cytosine, resulting in uracil which can be processed as thymine during replication. Following the same strategy, David Liu’s laboratory generated an adenine base editor (ABE) using the adenine deaminase TadA (Gaudelli et al., 2017). Deamination of adenosine results in inosine, which can be processed as guanosine during replication. The activity of ABE and CBE is very precise, occurring in a 7 bp frame (N3-N9 in N20-PAM) (Shan & Voytas, 2018). Base editors have proven to be efficient in planta (Zong et al., 2017; Kang et al., 2018).

Dynamic loci visualisation
dCas9 is also used as a fluorophore-carrier to highlight specific loci. Combined with high resolution microscopy, this system enabled to follow multiple loci in human cell (Chen et al., 2013) or telomeres in plant cells (Dreissig et al., 2017) during different stages of the cell cycle.

Cas13 applications
The specificity of Cas13 for RNA binding has open the doors to novel CRISPR applications (Wolter & Puchta, 2018). Targeting of mRNA enabled to knock down gene expression at similar level than RNAi (RNA interfering, a typical method to knockdown gene expression)
but with a higher specificity (Abudayyeh et al., 2017). The transcriptional status of tested cell was otherwise normal, indicating that the collateral activity of Cas13 in vitro is not observed in vivo.

The in vitro collateral activity of Cas13 was used as an advantage to design the SHERLOCK (specific high-sensitivity Enzymatic Reporter unlocking) detection system (Gootenberg et al., 2017). By adding a fluorescent reporter (quenched fluorescent RNA), collateral Cas13 activity results in fluorescence. Such collateral activity occurs only in presence of a template RNA identical to the crRNA in complex with Cas13. The system was able to detect Zika virus fragments with high sensitivity and specificity.

Cas13 was also used to defend against RNA virus, such as TuMV (turnip mosaic virus) in plants (Aman et al., 2018). Transient expression of Cas13 and crRNA targeting TuMV reduced the proliferation of the virus by 50%. This is one example among many other CRISPR applications for plant disease resistance (Langner et al., 2018).

1.3.3 CRISPR for plant immunity

1.3.3.1 Knocking-out Susceptibility-genes

R-genes are sufficient to confer resistance to a given pathogen. In contrast, some genes are required for susceptibility and are consequently called Susceptibility(S)-genes (Eckardt, 2002). Several examples of CRISPR-mediated S-gene disruption have resulted in resistant plants (Zaidi et al., 2018). MLO is a dominant S-gene that confers broad spectrum susceptibility to powdery mildew in diverse plants from monocots to dicots. An Arabidopsis mlo2-mlo6-mlo12 triple mutant is fully resistant to an Arabidopsis-compatible powdery mildew strain (Acevedo-Garcia et al., 2014). CRISPR was used to knock-out MLO genes in wheat and tomato (Wang et al., 2014; Nekrasov et al., 2017). The edited plants were resistant to powdery mildew. This approach enables to generate resistant crops in few months, with mutation undistinguishable from naturally occurring mutation (Nekrasov et al., 2013).

1.3.3.2 Knock-in breeding Resistance-genes

HDR-mediated repair of CRISPR-induced DSB can result in gene replacement or insertion in plant and animal (Li et al., 2013; Mali et al., 2013). Gene targeting via HDR occurs via homologous recombination between a template and genomic DNA at very low frequency. It is drastically enhanced when DSB occurs (Puchta & Fauser, 2013). Although
CRISPR-mediated gene targeting is a routine assay in animal, it remains currently a challenge in plant. Stable gene targeting was achieved in Arabidopsis and tomato in 6% to 25% of transformed plants (Dahan-Meir et al., 2018; Wolter et al., 2018; Vu et al., 2019). Theoretically, CRISPR could be used to introgress an R-gene in a susceptible line quickly and with no other DNA introgression as compared to classic breeding. Currently it has not been demonstrated.

In only five years, CRISPR has been deployed to build significant resistance in plant via RNA virus targeting and S-gene knock-out. If gene targeting is one of the current challenges, we can expect to see much more CRISPR-derived applications in the next few years, for plant disease resistance and beyond.
Chapter 2 : Materials and methods

2.1 Material

2.1.1 Plant material

_Arabidopsis thaliana_ (Arabidopsis) wild type lines used in this study are Col-0, Oy-0, Ler-0 and Ws-2. The Recombinant Inbred Line population between Oy-0 and Col-0 is the collection “27RV” from the “Versailles Arabidopsis Stock Centre”. The MAGIC329 and MAGIC23 are from the MAGIC collection (Kover et al., 2009). The mutant lines used in this study are Be-o_ADHR002 (NASC: N8102), Col-0_wrr4-6 (NASC: N667351), Col-0_adrl-adrl-1l-adrl-1l (Bonardi et al., 2011), Ws-2__rrs1-rps4-rps4b (Saucet et al., 2015), Ws-2_eds1-1 (Falk et al., 1999), Col-0_eds1-2 (Bartsch et al., 2006), Col-0_eds1-12 (Ordon et al., 2017), and Col-0_RPW8-S5 (Xiao et al., 2003). _Camelina sativa_ wild type used in this study is cultivar Celine. Seeds were sown directly on compost and plants were grown at 21 °C, with 10 hours of light and 14 hours of dark, 75 % humidity. For seed collection, 4- to 5-weeks old plants were transferred under long-day condition: 21 °C, with 16 hours of light and 8 hours of dark, 75 % humidity.

The Solanaceae plants used in this study were _Nicotiana benthamiana_ and _Nicotiana tabacum_ cultivar “Petit Gerard”. Seeds were sown directly on compost and plants were grown at 21 °C, with 16 hours of light and 8 hours of dark, 55 % humidity.

2.1.2 Microbial material

2.1.2.1 Escherichia coli

_Escherichia coli_ strain DH10B genotype F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara leu) 7697 galU galK rpsL nupG λ– was used for cloning purposes. They were grown at 37 °C overnight on LB media supplemented with antibiotic according to the plasmid-carried selectable marker.

2.1.2.2 Agrobacterium tumafaciens

_Agrobacterium tumafaciens_ strains used in this study were GV3101, Agl-1 and GALLS (Hodges et al., 2004). They were grown at 28 °C overnight for 48 hours on LB media
supplemented with gentamycin (GV3101 and GALLS) or carbenicillin (Agl-1) along with rifampicin and a third antibiotic according to the plasmid-carried selectable marker.

2.1.2.3 Pseudomonas spp.

Pseudomonas syringae pv tomato strain DC3000 (DC3000) and Pseudomonas fluorescens strain Pfo-1 (Pfo-1) engineered with a Type III Secretion System (Thomas et al., 2009) were used in this study. They were grown at 28 °C for 48 hours on KB media supplemented with rifampicin and kanamycin (DC3000) or chloramphenicol, tetracycline and gentamycin (Pfo-1).

DC3000 strains were expressing the following constructs: pVSP61 (empty vector), pVSP61::AvrRps4, pVSP61::AvrRps4-KRYYAAAA, pVSP61::PopP2, pVSP61::PopP2-C321A, pVSP61::AvrRpt2, pVSP61::AvrRpm1 or pVSP61::AvrPphB (Axtell & Staskawicz, 2003; Sohn et al., 2009; Williams et al., 2014).

Pfo-1 strains were expressing the following constructs: pBBR1MCS-5::AvrRps4, pBBR1MCS-5::AvrRps4-KRYYAAAA, pEDV6::PopP2, pEDV6::PopP2_C321A or pEDV6::AvrRpt2 (Sohn et al., 2007, 2014).

2.1.2.4 Oomycetes

The following Albugo candida races were used in this study: the Brassica oleracea-infecting race AcBoT, the Brassica juncea-infecting race Ac2V, the Brassica rapa-infecting race Ac7V, the Capsella spp.-infecting race AcEm2 and the Arabidopsis spp.-infecting races AcNc2 and AcEx1 (Liu & Rimmer, 1993; McMullan et al., 2015; Prince et al., 2017).

The following Hyaloperonospora arabidopsidis races were used in this study: Emoy2 and Cala2 (Parker et al., 1996).

2.1.3 Media

All recipes are for scale of 1 litre.

2.1.3.1 Lysogeny Broth (LB)

10 g tryptone, 5 g yeast extract, 5 g NaCl, 1 g glucose, pH 7.0. For solid medium, 10 g agar was included.
2.1.3.2 Lennox (L)

10 g tryptone, 5 g yeast extract, 10 g NaCl, pH 7.0. For solid medium, 10 g agar was included.

2.1.3.3 King’s B (KB)

20 g peptone, 10 ml glycerol, 1.6 g Potassium Hydrogen Phosphate, 10 ml glycerol. pH was adjusted to 5.8 with NaOH. For solid medium, 15 g agar was included.

2.1.3.4 Murashige and Skoog ½ (MS ½)

2.2 g of Murashige and Skoog medium (micro and macro elements including vitamins), 30 g sucrose. pH was adjusted to 5.8 with NaOH. For solid medium, 10 g agar was included.

2.1.4 Antibiotics

Stock solutions were stored at -20 °C, except for rifampicin, which was stored at 4 °C. Working concentrations indicate the final concentrations used in selective media.

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock concentration</th>
<th>Working concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbenicillin</td>
<td>100 mg/ml in H₂O</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>10 mg/ml in H₂O</td>
<td>10 µg/ml</td>
</tr>
<tr>
<td>Gentamycin</td>
<td>10 mg/ml in H₂O</td>
<td>10 µg/ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>150 mg/ml in H₂O</td>
<td>150 µg/ml</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>10 mg/ml in methanol</td>
<td>10 µg/ml</td>
</tr>
<tr>
<td>Spectinomycin</td>
<td>100 mg/ml in H₂O</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>5 mg/ml in ethanol</td>
<td>5 µg/ml</td>
</tr>
</tbody>
</table>

Table 2-1: Antibiotics used in this study

2.2 Methods

2.2.1 Molecular Biology

2.2.1.1 Genomic DNA isolation

For short-term storage of plant DNA (e.g. for genotyping by PCR), ~0.5 cm² of leaf tissue was printed by mechanical pressing on a Whatman® FTA® card. Card was air dried and stored at room temperature. For PCR application, a 1.2 mm diameter disc was extracted from the card-printed leaf with a Harris Uni-Core Punch and incubated for 2 hours at room
temperature with 50 µl FTA extraction buffer (10mM TrisHCl, 0.01 %v/v Tween 20, 2mM EDTA). The buffer was discarded and the disc was washed with 180 µl water before being used as PCR template.

For long term storage, tissue was ground and mixed with 700 µl CTAB extraction buffer (2 g CTAB, 2 g PVP MW 40,000, 28 ml NaCl 5 M, 4 ml EDTA 0.5M pH 8, 10 ml TrisHCl 1M pH 8 in water for a total volume of 100 ml) and placed at 65 °C for 2 hours. DNA was extracted with 500 µl Chloroform:Isoamylalcol, by mixing the sample by inversion and centrifuging for 10 minutes at 20,000 RCF. The upper phase was transferred in a new tube and DNA was precipitated with 2/3 volume of isopropanol for 5 minutes at room temperature. DNA was span down by centrifuging at 20,000 RCF. Supernatant was discarded. Pellet was washed with ethanol 76 %, air-dried, re-suspended in 300 µl water and stored at -20 °C.

2.2.1.2 Polymerase chain reaction

Standard PCR were performed using Taq DNA Polymerase (NEB). For cloning of PCR products (except for USER® cloning), Q5® High-Fidelity DNA Polymerase (NEB) was used. For cloning of PCR products via USER® cloning, KAPA HiFi HotStart Uracil+ ReadyMix (KAPABIOSYSTEM) was used. Reaction mix, annealing and extension temperature were defined following the manufacturer’s instructions. PCR were carried out in a thermocycler.

To estimate PCR product size and quantity, gel electrophoresis was used. Gels were made with 1.5 % agarose, TAE buffer 1X and 0.01 µl/ml ethidium bromide. Gels were run at ~100-150 V and were visualized using UV light. Gel extractions were performed using the QIAquick Gel Extraction Kit (QIAGEN), following the manufacturer’s instructions.

2.2.1.3 Golden Gate Cloning

The Golden Gate assembly enables a single-step assembly of multiple DNA fragments into a destination vector, utilising the activities of type II endonucleases BsaI and BpiI and DNA ligase T4 (Engler et al., 2009; Weber et al., 2011). For each assembly, a unique set of compatible 4 bp overhangs at the junction of each modules are designed to ensure the assembly in desired order. The overhang were chosen following the plant Golden Gate standards (Engler et al., 2014). The modules were assembled in a single digestion and ligation reaction called diglig.

Digligs were done in the following conditions: 0.02 pmol of each module and the destination vector were mixed with 0.5 µl BsaI-HF® enzyme (20,000 units/ml, NEB) or
BpiI® enzyme (10,000 units/ml, ThermoFisher), 0.15 µl BSA 10X (2 mg/ml, NEB), 1 µl T4 DNA ligase (400,000 units/ml, NEB), 1.5 µl CutSmart® buffer (NEB) and water to a final volume of 15 µl. Digligs were carried out in a thermocycler with the following steps: initial digestion at 37 °C for 1 minute, 25 cycles of digestion at 37 °C for 3 minutes and ligation at 16 °C for 4 minutes, denaturation of enzymes at 50 °C for 5 minutes and 80 °C for 5 minutes. 2 to 4 µl were used for transformation in electro-competent E. coli.

2.2.1.4 USER cloning

The USER cloning assembly enables to integrate one or several PCR amplicons in a USER-compatible vector (Geu-Flores et al., 2007). The procedure is based on the use of PCR primers that contain a single deoxyuridine near their 5’ end. Treatment of the PCR products with deoxyuridine-excision enzyme generates long overhangs designed to specifically complement each other.

USER-compatible vector was pre-digested with PacI and Nt.BBvCI. USER ligation was conducted by mixing 0.02pmole of gel purified PCR product(s) with the pre-digested destination vector with 1 µL USER® enzyme (NEB), 1 µL Taq DNA Polymerase Buffer (NEB) and water in a total reaction volume of 10 µl. The mix was incubated at 37 °C for 15 minutes then 25 °C for 15 minutes.

2 to 4 µl were used for transformation in chemical-competent E. coli.

2.2.1.5 Gibson assembly

Gibson assembly enables to join multiple linear DNA fragments sharing identity at their extremity (Gibson et al., 2009). It is based on the concerted action of a 5’ exonuclease, a DNA polymerase and a DNA ligase. Two DNA fragment can be assembled if the ~20 nucleotides 5’ end of one is identical to the 3’ end of the other.

The reagents are the 5X ISO buffer (for 6 ml: 3 ml TrisHCl [1 M, pH 7.5], 150 µl MgCl₂ [2 M], 60 µl of dGTP [100 mM], 60 µl of dATP [100 mM], 60 µl of dTTP [100 mM], 60 µl of dCTP [100 mM], 300 µl DTT [1 M], 1.5 g PED-8000, 200 µl NAD [100 mM] and water) and the Gibson Assembly Mix (for 1.2 ml: 320 µl 5X ISO buffer, 0.64 µl T5 exonuclease [10,000 units/ml, NEB], 20 µl Phusion® High-Fidelity DNA Polymerase [2,000 units/ml, NEB], 160 µl Taq DNA ligase [10,000 units/ml, NEB] and water).
Two or more Gibson compatible linear DNA molecules (typically PCR products) were mixed in an equimolar amount with 15 µl Gibson Assembly Mix and water in a total volume of 20 µl. The mix was incubated at 50 °C for 1 hour.

2 to 4 µl were used for transformation in chemical-competent E. coli.

2.2.1.6 Transformation in E. coli and A. tumefaciens

Plasmid were transformed using ~5x10⁸ cells in a 0.1cm cuvette and a micropulser (Bio-Rad) on settings recommended by the manufacturer. Bacteria were suspended in 300 µl of LB and incubated 1 hour at 37 °C (E. coli) or 2 hours at 28 °C (A. tumefaciens). 50 µl were spread on L plate with the appropriate antibiotic. For transformation of a Golden Gate cloning product, 10 µl of X-Gal (40mg/ml) was added on top of the plate for white/blue selection. The plate was incubated at 37 °C overnight (E. coli) or at 28 °C for 48 hours (A. tumefaciens).

2.2.1.7 Plasmid purification and confirmation

For plasmid purification, two white colonies were re-cultivated on liquid LB with the appropriate antibiotic and incubated overnight at 37 °C (E. coli) or at 28 °C (A. tumefaciens). The plasmid DNA was extracted using a QIAprep Spin Miniprep Kit (QIAGEB) following the manufacturer’s instruction. The plasmid DNA was eluted in 50 µl water and stored at -20 °C.

Successful DNA insertions into the plasmid were checked by restriction enzyme analysis. Plasmids sequences were confirmed by Illumina sequencing using service provided by GATC Biotech. The sequencing data were analysed using CLC Main Workbench 7.7.1. Correctly transformed bacteria were stored as glycerol stock (1 ml of liquid culture from a single colony with 1 ml of 60 % glycerol) at -80 °C.

2.2.1.8 Gene expression measurement by RT-qPCR

For gene expression analysis, RNA was isolated from three biological replicates and used for subsequent reverse transcription quantitative PCR (RT-qPCR) analysis. Briefly, RNA was extracted using the RNeasy Plant Mini Kit (QIagen) and treated with RNase-Free DNase Set (QIagen). Reverse transcription was carried out using the SuperScript IV Reverse Transcriptase (ThermoFisher). qPCR was performed using CFX96 Touch™ Real-Time PCR Detection System. Data were analysed using the double delta Ct method (Livak & Schmittgen, 2001).
2.2.1.9 Protein extraction and Western Blot

Proteins were extracted from leaf tissue using TruPAGE™ LDS Sample Buffer (Sigma-Aldrich) following the manufacturer recommendations. They were separated by SDS-PAGE and analysed by immunoblotting. Tris-Glycine polyacrylamide (PAA) gels were prepared with 5 % polyacrylamide for the stacking gel, and 10 or 12 % polyacrylamide for resolving gels in this study. The pre-stained protein ladder (PageRuler, ThermoFisher) was used as molecular weight marker. Proteins were transferred to Immobilon-P PVDF membranes (Merck Millipore), using a semi-dry transfer apparatus supplied by Trans-Blot Turbo (Bio-Rad). Membranes were blocked for 1 h at room temperature or overnight at 4 °C in TBST (Tris-Buffered Saline with 0.1 % Tween) containing 5 % (w/v) non-fat dry milk. Membrane incubation with Horseradish Peroxidase (HRP) conjugated antibodies (AntiFLAG M2, 1:10000 dilution, Sigma; Anti-GFP, 1:10000 dilution, Santa Cruz Biotechnology) was carried out in TBST supplemented with 5 % milk by gentle agitation at room temperature for 1 h. The membrane was then rinsed 3 times in TBST (10 min), and once in TBS (Tris-Buffered Saline). Chemiluminescence detection for proteins of interest was carried out firstly by incubating the membrane with developing reagents (SuperSignal West Pico & West Femto), using ImageQuant LAS 4000 (Life Sciences).

2.2.2 Plant biology

2.2.2.1 Transient expression

A. tumefaciens strains were streaked on selective media and incubated at 28 °C for 48 hours. A single colony was transferred to liquid LB medium with appropriate antibiotic and incubated at 28 °C for 24 hours in a shaking incubator (200 rotations per minute). The resulting culture was centrifuged at 3000 rotations per minute for 5 minutes and resuspended in 2 ml of infiltration buffer (10mM MgCl2, 10mM MES, pH 5.6) at OD600=0.5 (2.5x10⁸ colony forming unit [cfu/ml]). For co-expression, each bacterial suspension was adjusted to OD600=0.5 in the final mix. The abaxial surface of 5-weeks old N. tobacco or N. benthamiana were infiltrated with 1 ml needle-less syringe. Leaves were phenotyped for cell death at 3 to 7 dpi.

2.2.2.2 Stable transformation

A. tumefaciens mediated transformation of Arabidopsis was carried out using the floral dip method (Clough & Bent, 1998). Briefly, 6- to 8-weeks old plants were dipped in a solution of A. tumefaciens at OD600=0.5 (2.5x10⁸ cfu/ml).
Recovered seeds were selected with either herbicide (glufosinate sprayed three times on 1- to 3-weeks old plants at a concentration of 0.375g/l), antibiotic or fluorophore expression in seeds. The fluorescence in seeds is the result of the expression of RFP or GFP under the control of the seed-specific AtOLEOSIN1 promoter, described as the FAST-Red/FAST-Green method (Shimada et al., 2010). The seeds were manually selected using a fluorescence stereomicroscope.

2.2.2.4 CRISPR event characterization

Six to twelve independent transformants were tested for CRISPR activity. The loc(i)us containing the gRNA-target(s) was(were) amplified by PCR. The amplicon was either sent for sequencing or digested with a restriction enzyme having its recognition site at the gRNA target. The loss of restriction site was assessed on an electrophoresis gel. Transformants showing a high undigested vs digested ratio and/or a sequencing trace with elevated polymorphism at the gRNA target was considered as somatically CRISPR active lines. The progeny of the most CRISPR active line was screened for homozygous events, using the same method. Lines showing full resistance to digestion and/or a sequencing trace with a clear single indel were considered as putative homozygous mutant. Such lines were generally identified in T2 but in some cases, homozygous mutants were identified directly in T1.

Once a line was genotyped as putative homozygous mutant, its progeny was selected for the absence of T-DNA and screened by PCR and sequencing to confirm that the mutation was indeed inherited from the parent. The seeds of least one non-transgenic progeny carrying the CRISPR-induced homozygous mutation were collected and stored for further experiments.

For the characterization of CRISPR event at the ADH1 locus specifically, plants were screened with allyl-alcohol. ~100 seeds were sterilized, immersed in water and incubated at 4 °C in the dark overnight. Then they were treated with 30mM allyl-alcohol (ALDRICH) at room temperature for 2 hours, shaken at 750 rpm. Then they were rinsed three times with water and sown on MS½ medium. After two weeks, the number of germinated and non-germinated seeds was monitored. The adh1 mutation was confirmed by PCR and sequencing for up to six allyl-alcohol resistant lines per genotype. Sequencing results were compared to the Col-0 sequence of ADH1 using CLC Main Workbench 7.7.1. ADH1 genotypes were reported as WT (identical to Col-0), heterozygous (both Col-0 and single mutation detected), biallelic (two different mutations detected), homozygous (single
mutation detected) or somatic (more than two signals detected). For each T2 family, the CRISPR efficiency was defined as the ratio of homozygous and biallelic mutants compared to the total number of seeds sown.

2.2.3 Pathology assay

2.2.3.1 Albugo candida infection

For propagation of *A. candida*, zoospores were suspended in water (~$10^5$ spores/ml) and incubated on ice for 30 min. The spore suspension was then sprayed on plants using a Humbrol® spray gun (~700 μl/plant) and plants were incubated at 4 °C in the dark overnight. Infected plants were kept under 10 hours light (20 °C) and 14 hours dark (16 °C) cycles. Plants were scored as susceptible if a pathogen was capable of accomplishing its life cycle and sporulation was macroscopically visible within 3 weeks after plant inoculations.

2.2.3.2 *Hyaloperonospora arabidopsidis* infection

For propagation of *H. arabidopsidis* (*Hpa*), 1-week old Arabidopsis seedlings were sprayed with fresh *Hpa* spores at a concentration of $10^4$ spores/ml using a Humbrol® spray gun (~700 μl/plant). Sprayed seedlings were covered with a plastic lid and were kept under 10 hours light (20 °C) and 14 hours dark (16 °C) cycles. Susceptibility was measured as the number of spores per plant. Approximately 80 plants were bulked in 2 ml of water and spores were counted using a haemocytometer. Results are expressed as the number of spores per plant.

2.2.3.3 Hypersensitive Response assay

*Pseudomonas fluorescens* (*Pfo-1 EtHAn*) or pVSP61:AvrPphB or *Pseudomonas syringae* (DC3000) carrying denoted constructs were grown on selective KB-medium agar plate for 48 hours at 28 °C. Bacteria were harvested from plate, re-suspended in 10 mM MgCl₂ and concentration was adjusted to OD₆₀₀ = 0.2 ($10^8$ cfu/ml). The abaxial surface of 4- to 5-weeks old Arabidopsis leaves were hand-infiltrated with 1 ml needle-less syringe. Cell death was monitored qualitatively 12 to 24 hours after infiltration. For quantitative measurement of HR, electrolyte leakage assay was conducted. Leaf discs were taken with a cork borer from infiltrated leaves. Discs were dried, washed in deionized water for half an hour before being floated on deionized water (16 discs per sample, 3 samples per
biological replicate for 3 biological replicates). Electrolyte leakage was measured on a Horiba LAQUQtwin-EC-33 conductivity meter at indicated time points.

2.2.3.4 Bacterial growth assay

*Pseudomonas syringae* (DC3000) carrying denoted constructs were grown on selective KB-medium agar plate for 48 hours at 28 °C. Bacteria were harvested from plate, re-suspended in 10mM MgCl₂ and concentration was adjusted to OD₆₀₀ = 0.001 (5×10⁵ cfu/ml). The abaxial surface of 4- to 5-weeks old Arabidopsis leaves were hand-infiltrated with 1 ml needle-less syringe. For quantification, leaf sample were harvested with a 6 mm diameter cork borer, resulting in a ~0.283 cm²-sized leaf disc. Two leaf discs per leaf were harvested and used as single sample. For each condition, four samples were collected just after infiltration and eight samples were collected 72 hours after infiltration. Samples were ground in 200 µl of 10mM MgCl₂, serially diluted (5, 50, 500, 5000 and 50000 times) and spotted (6 to 10 µl per spot) on selective KB-medium agar plate to grow 48 hours at 28 °C. The number of colonies (cfu per drop) was monitored and the bacterial growth was expressed in cfu/cm² of leaf tissue.

2.2.3.7 Salicylic acid measurement

*Pseudomonas fluorescens* (Pf0-1) (Thomas et al., 2009) carrying denoted construct were grown on selective KB-medium agar plate for 48 hours at 28 °C. Bacteria were harvested from the plate, re-suspended in infiltration buffer (10mM MgCl₂, pH 5.6) and concentration was adjusted to OD₆₀₀ = 0.05 (2.5×10⁷ cfu/ml). The abaxial surface of 5-week old Arabidopsis leaves was hand-infiltrated with 1 ml needle-less syringe. Leaves were harvested 24 hours post inoculation and freeze dried. Salicylic acid was extracted from 10 mg of ground dry tissue with 400 µl of 10 % methanol and 1 % acetic acid in water on ice for 30 minutes. The solution was centrifuged at 13000 rpm for 10 minutes. A second extraction was carried out on the pellet in the same conditions and both supernatants were mixed. Samples were analysed on an Acquity UPLC attached to a TQS tandem mass spectrometer (both from Waters). Detection was by negative electrospray MS. Spray chamber conditions were 600 °C desolvation temperature, 900 L.hr⁻¹ desolvation gas, 150 L.hr⁻¹ cone gas, and 7.0 bar nebulizer pressure. The spray voltage was 1.5 kV in negative mode.
2.2.3.6 Trypan Blue staining

Cell death and filamentous pathogen development were studied in leaf mounts stained with lactophenol-trypan blue. Whole leaves were boiled for 1 minute in stain solution (10 ml lactic acid, 10 ml glycerol, 10g phenol, 10mg trypan blue and water in a final volume of 10 ml) and then decolorized in chloral hydrate (2.5 g chloral hydrate and water in final volume of 1 ml) for several hours to several days. They were mounted in 60 % glycerol and examined on a microscope.
Chapter 3 : Allelic variation in the TIR-NLR WRR4 defines white rust resistance specificity

3.1 Introduction

Albugo candida is an oomycete causing white rust in Brassicaceae. There are many genetic races of A. candida, which are characterized by a strong host specificity (Jouet et al., 2018). Sexual recombination has been reported between races, indicating that gene exchanges and emergence of new races are possible (McMullan et al., 2015). Despite the host specificity, some crop-infecting races can grow in Arabidopsis. This compatibility has been exploited to map White Rust Resistance (WRR)-genes in Arabidopsis, to be cloned into crops. For instance, WRR4A and its paralog WRR4B from Arabidopsis Col-0 confer resistance to the Brassica juncea-infecting race Ac2V and to the Brassica oleracea-infecting race AcBoT (Borhan et al., 2008; Cevik et al., 2019). These resistances are effective in Arabidopsis and in crops (Borhan et al., 2010; Cevik et al., 2019).

Although WRR4A and WRR4B from Col-0 confer broad spectrum white rust resistance, A. candida race Exeter1 (AcEx1) can grow on Col-0 (Prince et al., 2017). If WRR4A and WRR4B are deployed in Brassica crops, these resistances can conceivably be overcome. Moreover, AcEx1 can grow on the oil-producing crop Camelina sativa.

C. sativa has been engineered to produce a PUFA (polyunsaturated fatty acid)-rich oil. PUFA, particularly the long chain (LC-PUFA) such as eicosapentaenoic (EPA) and docosahexaenoic (DHA), are recognized for their health benefit. For example, they reduce the risk of coronary heart diseases (Sidhu, 2003). Land plants do not synthesise EPA and DHA. The main source for human consumption remains fish and seafood, which accumulate LC-PUFA from marine microalgae in the wild. In aquaculture, the main source of LC-PUFA is fish oil, which is an unsustainable practice (Tocher, 2015). C. sativa modified to express algal genes can produce fish-like oil containing 12% of DHA (Petrie et al., 2014). This oil can effectively replace fish-oil from salmon feed, without affecting their nutritional quality for human consumption (Betancor et al., 2015). The culture of transgenic C. sativa for fish oil-like production may increase in the next few years and cultures would be exposed to white rust, particularly the race AcEx1. WRR genes against AcEx1 from Arabidopsis could be used to protect C. sativa fields.
The Arabidopsis accessions Oy-0 and HR-5 resist AcEx1. Recombinant Inbred Lines (RILs) between Oy-0 and the AcEx1-compatible accession Col-0 were used to map and clone an AcEx1 WRR gene from Oy-0.

3.2 Results

3.2.1 Col-0 displays a weak WRR4A-dependent response to AcEx1

Although AcEx1 can achieve its life cycle on Col-0, a weak chlorotic response can be observed (Figure 3-1). This chlorotic response is not seen in the fully susceptible accession Ws-2 or in a Col-0_wrr4a knock-out mutant. It indicates that WRR4A is activated upon AcEx1 infection but cannot trigger a sufficient immune response for pathogen growth arrest. Six weeks after inoculation, Col-0 WT (carrying WRR4A) shows less disease symptoms than Col-0_wrr4a. Ws-2 lines expressing WRR4A from Col-0 under the control of native 5’ and 3’ regulatory sequences are partially resistant. The resistance appears higher than in Col-0 WT likely due to elevated expression of the transgene compared to natural expression level. Since WRR4A response is weak and is not sufficient to arrest the pathogen propagation, Col-0 is considered as susceptible in this study.
3.2.2 Oy-0 resists AcEx1 via two loci WRR13 and WRR11

Volkan Cevik and Sebastian Fairhead previously conducted a Quantitative Trait Locus (QTL) analysis among RILs between a resistant parent Oy-0 and a susceptible parent Col-0. It indicates that two major loci are involved in AcEx1 resistance in Oy-0: the WRR13 locus on chromosome 1 (20,384,267 - 22,181,333 bp) and the WRR11 locus on chromosome 3 (17,282,622 - 19,628,061 bp). There is also a small QTL on chromosome 5, called WRR15. This QTL is below the LOD score threshold and was not further investigated.

RenSeq (R-gene enrichment sequencing) helps define the NLR repertoire of a sample (Jupe et al., 2013). Oliver Furzer conducted RenSeq in Oy-0 and identified all the NLRs associated with WRR13 and WRR11. The WRR13 locus contains a cluster of TIR-NLRs, orthologous to the WRR4 cluster in Col-0, and a CC-NLR cluster, orthologous to the RPP7 cluster in Col-0. The organisation of the WRR4 locus is different between Oy-0 and Col-0.
In Col-0 there are three genes referred as WRR4A, WRR4B and At1g56520 (WRR4C) (Cevik et al., 2019). In Oy-0, WRR4A and WRR4B are present, WRR4C is pseudogenised and a fourth paralog, called WRR4D, is present distal to WRR4A. The WRR11 locus contains two TIR-NLR encoding genes, paired in a head-to-head orientation (At3g51560-At3g51570), the CC-NLR encoding gene ZAR1, the RPW8 cluster and a CC-NLR encoding gene absent in Col-0.

3.2.3 Fine mapping of WRR13

Twelve RILs recombine within the WRR13 locus. Based on their phenotype and on their genotype at WRR11 and WRR15 loci, three have the Col-0 allele of WRR13, three have the Oy-0 allele and six were not considered as they could be either (Table 3-1).

<table>
<thead>
<tr>
<th>Line</th>
<th>WRR13</th>
<th>WRR11</th>
<th>WRR15</th>
<th>Phenotype</th>
<th>If Col-0</th>
<th>If Oy-0</th>
<th>WRR13 allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>506</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>5</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>?</td>
</tr>
<tr>
<td>252</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>4</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>?</td>
</tr>
<tr>
<td>315</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>5</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>?</td>
</tr>
<tr>
<td>189</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Oy-0</td>
<td>3</td>
<td>3 to 5</td>
<td>1 to 4</td>
<td>?</td>
</tr>
<tr>
<td>64</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>5</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>?</td>
</tr>
<tr>
<td>103</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>5</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>?</td>
</tr>
<tr>
<td>91</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>Oy-0</td>
<td>3</td>
<td>3 to 4</td>
<td>0 to 2</td>
<td>Col-0</td>
</tr>
<tr>
<td>298</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>Oy-0</td>
<td>3</td>
<td>3 to 4</td>
<td>0 to 2</td>
<td>Col-0</td>
</tr>
<tr>
<td>118</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>Oy-0</td>
<td>3</td>
<td>3 to 4</td>
<td>0 to 2</td>
<td>Col-0</td>
</tr>
<tr>
<td>273</td>
<td>recombinant</td>
<td>Col-0</td>
<td>Col-0</td>
<td>3</td>
<td>4 to 5</td>
<td>3 to 5</td>
<td>Oy-0</td>
</tr>
<tr>
<td>19</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>Oy-0</td>
<td>0</td>
<td>3 to 4</td>
<td>0 to 2</td>
<td>Oy-0</td>
</tr>
<tr>
<td>336</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>Oy-0</td>
<td>2</td>
<td>3 to 4</td>
<td>0 to 2</td>
<td>Oy-0</td>
</tr>
</tbody>
</table>

Table 3-1: Phenotypes and genotypes of WRR13 recombinant lines

The phenotype is scored on a scale from 0 (fully resistant) to 5 (as susceptible as Col-0). The “if Col-0” column indicates the phenotypic score range in RILs having the indicated genotype at WRR11 and WRR15 loci and being Col-0 at the WRR13 locus. The “if Oy-0” column indicates the phenotypic score range in RILs having the indicated genotype at WRR11 and WRR15 loci and being Oy-0 at the WRR13 locus. Oy-0 alleles are highlighted in bold. The score range matching the score of each line is underlined.

I identified a SNP (Single Nucleotide Polymorphism; C1_21,195) and a RFLP (Restriction Fragment Length Polymorphism; C1_21,691) marker within WRR13, between the WRR4.
and RPP7 clusters. These markers enabled fine mapping (Figure 3-2). Four lines eliminate the RPP7 cluster from the WRR13 QTL. None eliminates WRR4, indicating that WRR13 resistance is mediated by one or several of the WRR4 paralogs or by a non NLR-encoding gene in the vicinity.

Figure 3-2: The WRR13 QTL includes the WRR4 cluster but not the RPP7 cluster
Six lines were genotyped using two markers within the WRR13 QTL. Dark grey indicates the presence of the gene underlying WRR13 resistance, light grey indicates its absence. Dash grey indicates the recombination locus, which can contain the gene underlying WRR13 resistance. Arrows indicates genes from the Oy-0 resistant genotype. The names of the markers refer to the position in kb on chromosome 1. Figure is not on scale.
3.2.4 Fine mapping of WRR11

Eight RILs recombine within the WRR11 locus. Based on their phenotype and on their genotype at WRR13 and WRR15 loci, three have the Col-0 allele of WRR11, three have the Oy-0 allele and two were not considered as they could be either (Table 3-2).

<table>
<thead>
<tr>
<th>Line</th>
<th>WRR13</th>
<th>WRR11</th>
<th>WRR15</th>
<th>Phenotypes</th>
<th>If Col-0</th>
<th>If Oy-0</th>
<th>WRR11 allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>162</td>
<td>Col-0</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>4 - 4 - 4</td>
<td>4 to 5</td>
<td>3 to 4</td>
<td>?</td>
</tr>
<tr>
<td>304</td>
<td>Col-0</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>3 - 3 - 2</td>
<td>4 to 5</td>
<td>3 to 4</td>
<td>Oy-0</td>
</tr>
<tr>
<td>471</td>
<td>Oy-0</td>
<td>recombinant</td>
<td>Col-0</td>
<td>4 - 4 - 4</td>
<td>4 to 5</td>
<td>1 to 3</td>
<td>Col-0</td>
</tr>
<tr>
<td>472</td>
<td>Oy-0</td>
<td>recombinant</td>
<td>Col-0</td>
<td>4 - 4 - 3</td>
<td>4 to 5</td>
<td>1 to 3</td>
<td>Col-0</td>
</tr>
<tr>
<td>497</td>
<td>Col-0</td>
<td>recombinant</td>
<td>Col-0</td>
<td>5 - 4 - 4</td>
<td>4 to 5</td>
<td></td>
<td>Col-0</td>
</tr>
<tr>
<td>155</td>
<td>Oy-0</td>
<td>recombinant</td>
<td>Oy-0</td>
<td>0 - 0 - 0</td>
<td>1 to 3</td>
<td>0 to 2</td>
<td>Oy-0</td>
</tr>
<tr>
<td>367</td>
<td>Col-0</td>
<td>recombinant</td>
<td>Col-0</td>
<td>3 - 3 - 4</td>
<td>4 to 5</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>403</td>
<td>Col-0</td>
<td>recombinant</td>
<td>Col-0</td>
<td>3 - 2 - 3</td>
<td>4 to 5</td>
<td>3</td>
<td>Oy-0</td>
</tr>
</tbody>
</table>

Table 3-2: Phenotypes and genotypes of WRR11 recombinant lines

The phenotype is scored on a scale from 0 (fully resistant) to 5 (as susceptible as Col-0). Lines have been phenotyped three times independently. The “If Col-0” column indicates the phenotypic score range in RILs having the indicated genotype at WRR13 and WRR15 loci and being Col-0 at the WRR11 locus. The “If Oy-0” column indicates the phenotypic score range in RILs having the indicated genotype at WRR13 and WRR15 loci and being Oy-0 at the WRR11 locus. Oy-0 alleles are highlighted in bold. The score range matching the score of each line is underlined.

I identified three RFLP (C3_18,016; C3_18,535; C3_18,850), one CAPS (Cleaved Amplified Polymorphic Sequence; C3_19,122) and one SNP (C3_18,937) markers within WRR11. These markers enabled fine mapping (Figure 3-3). Three lines eliminate the TIR-NLR pair, two lines eliminate ZAR1 and one line eliminates the RPW8 cluster. It suggests that WRR11 resistance is mediated by the CC-NLR absent from Col-0 or by a non NLR-encoding gene in its vicinity.
Figure 3.3: The WRR11 QTL includes a CC-NLR absent from Col-0

Six lines were genotyped using five markers within the WRR11 QTL. Dark grey indicates the presence of the gene underlying WRR11 resistance, light grey indicates its absence. Dash grey indicates the recombination locus, which can contain the gene underlying WRR11 resistance. Arrows indicates genes from the Oy-0 resistant genotype. The names of the markers refer to the position in kb on chromosome 3. Figure is not on scale.

3.2.5 WRR4A^{Oy-0} confers full resistance to AcEx1

In chromosome 1, WRR4A, WRR4B and WRR4C are three TIR-NLRs candidates to be WRR13. On chromosome 3, a CC-NLR absent from Col-0 is candidate to be WRR11 (hereafter CWR11 for Candidate to be WRR11). I cloned WRR4A, WRR4B, WRR4C and CWR11 from Oy-0 with their natural 5' and 3' regulatory sequences by the USER method (Geu-Flores et al., 2007)
and expressed them in the fully susceptible accession Ws-2 (Figure 3-4). Only WRR4A$_{Oy-o}$ confers full resistance to AcEx1; all the other genes do not contribute to resistance at all. As CWR11 is the only NLR candidate at the WRR11 locus, I used CRISPR to test also its role in AcEx1 resistance by loss-of-function. Briefly, I assembled Cas9 with the UBI10 promoter along with two sgRNAs targeting specifically CWR11 (TGAAGTACTTGCAAGCTAAGnGG and AGGAAGATCAAGGTTTGACnGG) and a FAST-Red selectable marker. The construct was expressed in Oy-0. Three T1 lines out of seven showed somatic activity. I selected eight Cas9-free (i.e. selected against FAST-Red) T2 progenies for each of the three lines. Out of these 24 lines, 17 were WT, five were heterozygous and one was homozygous cwr11 mutant. The homozygous mutation is a c.197del (of 2571 bp total CDS), resulting in an early stop codon in the CC domain encoding
region. This line is fully resistant to AcEx1 (Figure 3-4). Gain-of-function and loss-of-function approaches show that CWR11 is not sufficient nor required for AcEx1 resistance.
Four genes were cloned from Oy-0 into the AcEx1-susceptible accession Ws-2. CWR11 was also removed (Oy-0_cwr11) from the resistant accession Oy-0. Plants were sprayed inoculated with AcEx1 and pictures were taken two weeks after inoculation. The phenotypes are indicated under each picture. The numbers indicate the number of lines displaying the indicated phenotype out of the number of independent T1 lines transformed with the indicated gene.

Figure 3-4: WRR4A<sup>Oy-0</sup> confers full resistance to AcEx1

Susceptible 10/10 12/12 6/6 5/5

Resistant Resistant Resistant

AcEx1 Two weeks post inoculation
3.2.6 An early stop codon in WRR4A confers specificity in *Albugo candida* secreted protein recognition

The Col-0 allele of WRR4A confers a response to AcEx1 which is not full resistance (}

![Image](image-url)

**AcEx1**

Two weeks post inoculation
Figure 3-1), while the Oy-o allele of WRR4A confers full resistance to AcEx1 (Figure 3-4). I investigated the polymorphism between Oy-o and Col-o alleles to explain this dissimilar function. There are 26 non-synonymous SNPs, a 1- and an 8-amino acids
deletion in WRR4A°y° compared to the Col-0 allele (Figure 3-5). The most striking difference is an early stop codon in Col-0 that causes an 89-amino acid c-terminal deletion.
To test the relevance of this polymorphism, I cloned four alleles of WRR4A with the USER method under the control of the 35S promoter: Col-o WT, Oy-o WT, Col-o without its stop codon and including the following 267 nt (Col-o_ΔSTOP) and Oy-o with a restored stop codon and lacking the following 267 nt (Oy-o_+STOP). I tested these constructs in transient expression in tobacco leaves for recognition of *Albugo candida* secreted protein.
from the CCG family (Figure 3–6). HR (hypersensitive response causing macroscopic cell death) was used as a marker of CCG recognition. Amey Redkar and Volkan Cevik previously identified nine CCG proteins recognised by WRR4A<sup>Col-0</sup> (Redkar et al., in prep.). These nine CCGs are still recognised by WRR4A<sup>Col-0ΔSTOP</sup>. Only one is also recognised by WRR4A<sup>Oy-0</sup>, three are weakly
recognised and five are not recognised at all (Figure 3-6). It indicates that the recognition of these nine avirulence factors is mediated by the core section of the TIR-NLR. In contrast, three CCGs from AcEx1 are specifically recognised by the Oy-o allele of WRR4A. They are recognised by WRR4A<sup>Col-0ΔSTOP</sup> even
more strongly, but they are not recognised by WRR4A<sup>0+STOP</sup> (Figure 3-6). It indicates that the c-terminal extension is required for their recognition.
Tobacco leaves (cv Petit Gerard) were infiltrated with *Agrobacterium tumefaciens* delivering denoted constructs via T-DNA somatic integration, at $\text{OD}_{600} = 0.3$. WRR4A alleles were infiltrated in the positions indicated in the cartoon on the left. Pictures were taken 3 days after infiltration. CCG alleles are indicated in superscript, according to the *Albugo candida* race they are cloned from. a. Test of nine CCGs recognised by WRR4A<sup>Col-0</sup>. b. Test of three CCGs from AcEx1. Data from Amey Redkar.
3.2.7 WRR4A<sup>Oy-0</sup> resistance can be transferred in the crop *Camelina sativa*

The oil-producing crop *C. sativa* is susceptible to AcEx1 (Figure 3-7). I challenged ten *C. sativa* cv Celine WT plants with AcEx1 and observed disease symptoms two weeks after inoculation. Eight plants displayed white rust on one single leaf; two plants were not infected at all. AcEx1 is not as virulent on *C. sativa* than on Arabidopsis. I transformed WRR4A<sup>Oy-0</sup> with its natural 5’ and 3’ regulatory sequences in *C. sativa* cv Celine. I identified independent transformants with WT-like symptoms, reduced symptoms or no symptoms upon AcEx1 inoculation. All the lines with reduced or no symptoms expressed WRR4A<sup>Oy-0</sup> (Figure 3-7). I tested the transgenic (i.e. selected for FAST-Red) T2 progeny of four independent transformants. Lines #1 and #7 segregate 15:1 and lines #5 and #10 segregate 3:1 for FAST-Red, indicative of two and one locus T-DNA insertion respectively. Transgenic (homozygous or heterozygous) progenies of lines #5 and #10 contains 5/9 and 4/10 resistant plants respectively. I harvested seeds from these nine resistant plants for further investigation.
Figure 3.7: Some C. sativa expressing WRR4A<sup>0y-0</sup> display enhanced resistance to AcEx1

a. Expression of WRR4A<sup>0y-0</sup> in eight Camelina sativa lines showing resistance or intermediate phenotype in response to AcEx1. Expression is relative to EF1α. RNA was extracted from AcEx1 inoculated plants, 3 weeks post inoculation. – ctrl = negative control, i.e. C. sativa line transformed with CWR11. WRR4A<sup>0y-0</sup> is expressed at different levels and is not expressed in a line not transformed with WRR4A<sup>0y-0</sup>. b. Examples of AcEx1 symptoms on WT or WRR4A<sup>0y-0</sup> transgenic Camelina sativa cv Celine plants. Pictures were taken at 12 dpi. c. Phenotypes of ten Camelina sativa cv Celine plants WT, two weeks after AcEx1 inoculation. Some plants do not show symptoms. d. Phenotypes of nine or ten Camelina sativa cv Celine plants, two weeks after AcEx1 inoculation. Lines are WRR4A<sup>0y-0</sup> transgenic T2 lines, either homozygous or heterozygous.
3.3 Discussion

3.3.1 WRR11 QTL on chromosome 3 remains unrevealed

WRR11 is a major QTL involved in AcEx1 resistance to AcEx1 in Oy-0. The only NLR encoding gene associated with the locus is not involved in AcEx1 resistance, as showed by gain-of-function in Ws-2 and loss-of-function in Oy-0 (Figure 3-4). It is possible that that the resistance is mediated by another gene from the locus. Based on the Col-0 genome, there are 645 genes linked to WRR11, from At3g46930 to At3G50010. These genes include SNIPER4, RLP44, PMEI11, EDS1 and many uncharacterised genes. SNIPER4 is involved in signalling of the TIR-NLR SNC1 (Huang et al., 2018). RLP44 interacts with BAK1 and is involved in development (Wolf et al., 2014). PMEI11 is a pectin methylesterase inhibitor involved in resistance against the fungus Botrytis cinerea (Lionetti et al., 2017). EDS1 is a lipase-like protein involved in NLR signalling (Parker et al., 1996).

3.3.2 An early stop codon in a TIR-NLR causes loss of recognition of AcEx1

WRR4A^*Oy* provides full resistance to AcEx1, while the Col-0 allele does not (Figure 3-1 and Figure 3-4). The major polymorphism is an early stop codon in Col-0, which causes an 89-amino acid c-terminal deletion. This extension, in Oy-0 or in a mutated Col-0 allele of WRR4A, enables recognition of at least three CCGs of AcEx1 (Figure 3-6). They are not recognised by the WT allele of Col-0 or by a mutated allele of Oy-0. Conceivably, the ancestral state of WRR4A had the c-terminal extension. In the absence of AcEx1 selection pressure, an early stop codon mutation occurred in the Col-0 lineage, which also evolved to recognise more CCGs from the races AcEm2, Ac2V and AcNc2. The loss of the extension may confer an advantage in Col-0, such as the recognition of additional avirulence factors or an intramolecular regulation for a balanced immune signalling, but the current data do not demonstrate it.

3.3.3 Crops can be protected using R-genes from Arabidopsis

WRR4A^*Col* confers resistance to Ac2V in B. juncea and to AcBoT in B. oleracea (Borhan et al., 2010; Cevik et al., 2019). As AcEx1 can grow on the crop C. sativa, I tested if the WRR4A^*Oy* resistance was transferable from Arabidopsis to C. sativa. I used the European cultivar Celine because it is the one used to engineer the DHA/EPA-producing line (Petrie et al., 2014). AcEx1 is not particularly virulent in C. sativa cv Celine (Figure 3-7). AcEx1 is part of A.
*candida* group II while the *C. sativa* infecting races are part of the group III (Jouet *et al.*, 2018). Groups II and III are very close but distinct. Group III effectors may be more specialised to infect *C. sativa* than the effectors from Group II, explaining the partial compatibility between AcEx1 and *C. sativa*.

I identified two WRR4A\textsuperscript{Oy}\textsuperscript{-} transgenic *C. sativa* lines with single locus insertion. The progeny of these lines contains slightly more resistant plants than a WT population. The occurrence of susceptible lines in transgenic plants can be explained by lowered expression level. It could also be that WRR4A\textsuperscript{Oy}\textsuperscript{-} is not functional in *C. sativa*. WRR8 and WRR9 are two genes that confers full resistance to Ac2V in Arabidopsis but are not functional in *B. juncea* (Cevik *et al.*, 2019). The progeny of the transgenic resistant plants will be investigated to identify a line with consistent resistance to AcEx1.
Chapter 4: Optimization of CRISPR-Cas9 method in Arabidopsis

4.1 Introduction

4.1.1 Aspiration

In the previous chapter, I took advantage of natural genetic variation to identify an R-gene and investigate the plant immune system. However, genetic variation is not always available. For instance, all the natural Arabidopsis accessions resist the A. candida crop-infecting races Ac2V, Ac7V and AcBoT (Cevik et al., 2019). Random mutagenesis can be used to generate polymorphism and characterise genes in a forward genetic approach. In addition, transgressive segregation between 19 resistant Arabidopsis parents was employed to generate Ac2V, Ac7V and AcBoT susceptible lines, enabling the discovery of three novel White Rust Resistance (WRR)-genes (Cevik et al., 2019). In 2012/2013, CRISPR emerged as a cheap, fast and precise tool to edit genomes and create genetic variation (Gasiunas et al., 2012; Jinek et al., 2012). Conceivably, CRISPR could generate susceptible lines to a given pathogen in a previously non-host species. It appeared that CRISPR-mediated genome editing was not as efficient in Arabidopsis compared to other systems such as human, mouse, tomato and tobacco. Thus, I tried to optimize the method for generation of null alleles in Arabidopsis. The content of this chapter is largely identical to (Castel et al., 2019), a publication under the CC0 license.

4.1.2 Current status of CRISPR in Arabidopsis

Clustered regularly interspaced short palindromic repeat (CRISPR)- CRISPR associated (Cas) site-specific nucleases are components of prokaryotic immunity against viruses and are widely deployed as tools to impose operator-specified nucleotide sequence changes in genomes of interest (Gasiunas et al., 2012; Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013). For instance, the RNA-directed endonuclease Cas9 from the Type II CRISPR system functions in heterologous organisms, enabling applications such as targeted mutagenesis, dynamic imaging of genomic loci, transcriptional regulation and base editing (Barrangou & Horvath, 2017; Dreissig et al., 2017; Shan & Voytas, 2018). Expression of a single-guide RNA (sgRNA, an artificial fusion of the dual endogenous crisprRNA/trans-acting-crisprRNA) with Cas9, causes targeted DNA mutations in animal and plant cells
Cas9-sgRNA ribonucleoprotein causes a Double strand break (DSB) on DNA template homologous to the sgRNA spacer sequence. Cleaved DNA strands can be re-ligated by the endogenous Non-Homologous End Joining (NHEJ) system, which can result in insertions or deletions (indels) at the repaired site. Indels in the coding DNA sequence (CDS) can cause a codon reading frame shift resulting in loss-of-function alleles.

Arabidopsis is widely used for plant molecular genetics and lines mutated for a gene of interest are a valuable resource. Expression of CRISPR-Cas9 components can result in loss-of-function alleles of targeted genes in Arabidopsis, with variable efficiency (Jiang et al., 2013; Fauser et al., 2014; Feng et al., 2014). To improve induced mutation rates in Arabidopsis, several groups have evaluated various promoters to drive Cas9 expression (Yan et al., 2015; Wang et al., 2015c; Tsutsui & Higashiyama, 2017). I set out to optimize mutation rates in Arabidopsis. I conducted an extensive comparison of construct architecture, varying promoters, Cas9 alleles, terminator, sgRNA backbones and transcriptional direction.

There are several methods to deliver Cas9-sgRNA ribonucleoprotein into plant cells. The most common one is the expression of Cas9 and sgRNA by the plant itself after transformation. On the other hand, ribonucleoproteins can be directly delivered by protoplast transformation or particle bombardment (Woo et al., 2015; Wolter & Puchta, 2017; Liang et al., 2018a). In my experiments, Cas9 and the sgRNA were delivered by Agrobacterium tumefaciens-mediated transgenesis to avoid the process of regeneration via tissue culture. The method requires three steps: (i) DNA assembly of a binary vector with a selectable marker, a Cas9 and an sgRNA expression cassettes, (ii) transformation of the plasmid via the floral dip method (Clough & Bent, 1998) and (iii) identification of mutants among the transformed lines. I tested multiple T-DNA architectures for their ability to trigger homozygous mutations in the ADH1 gene. ADH1 converts allyl alcohol into lethal allyl aldehyde. Thus adh1 mutant lines resist an allyl alcohol treatment, enabling facile measurement of CRISPR-induced mutation rates (Fauser et al., 2014; Tsutsui & Higashiyama, 2017). I defined combinations of CRISPR components that enable high efficiency recovery of stable homozygous mutants in the first generation after transformation.
4.2 Results

4.2.1 T-DNA assembly is facilitated by the Golden Gate cloning method

In Golden Gate modular cloning, the promoter, reading frame and 3’ end modules at ‘Level 0’, are assembled using Type IIIS restriction enzymes into ‘Level 1’ complete transcript unit. Several Level 1 transcript units can then be combined into T-DNAs at ‘Level 2’. This enables facile assembly of diverse T-DNA conformations (Weber et al., 2011; Engler et al., 2014). In this project, I used three Level 0 acceptor vectors designed to clone promoter, CDS or terminator fragments. I also used three Level 1 vectors: a glufosinate plant selectable marker in position 1 (pICSL11017, cloned into pICH47732), a Cas9 expression cassette in position 2 (cloned into pICH47742) and an sgRNA expression cassette in position 3 (cloned into pICH47751) (Figure 3-1). Some Cas9 expression cassettes were cloned into a Level 1 position 2 variant: pICH47811. This vector can be assembled in Level 2 in the same fashion as pICH47742, but it enables Cas9 transcription in the opposite direction as compared to the other Level 1 modules.

Figure 4-1: Golden Gate-based cloning method to assemble CRISPR constructs

Cas9 alleles, promoters and terminators were cloned into the indicated Level 0 acceptor vectors as described in Materials and Methods and were assembled in Level 1 acceptor vector pICH47742 or pICH47811. An sgRNA targeting AtADH1 was amplified by PCR and assembled with the U6-26 promoter vector pICSL90002 in the same manner. Both Cas9 and sgRNA expression units were assembled in Level 2 acceptors pAGM4723 (not containing an overdrive sequence) or pICSL4723 (containing an overdrive) along with a glufosinate resistance plant selectable marker. An end-linker pICH41766 (EL2;3) was used to link the sgRNA expression unit to the Level 2 acceptor vector. For a head-to-head orientation of the sgRNA and Cas9 expression cassettes, Cas9 allele, promoter and terminator were assembled into pICH47811 instead of pICH47742.
With the help of Laurence Tomlinson and Federica Locci, I assembled 25 different Level 1 Cas9 constructs and four sgRNA expression cassettes. The sequence targeted by the sgRNA was CGTATCTCGCCATGAAGCNGG (Protospacer Adjacent Motif indicated in bold and italics) which targets specifically ADH1 in Col-0, enabling selection of CRISPR-induced adh1 homozygous mutants by selecting with allyl alcohol (Fauser et al., 2014). Assembly of these Level 1 modules resulted in 39 Level 2 T-DNA vectors (Table 4-1).
<table>
<thead>
<tr>
<th>Cas9 Promoter</th>
<th>Cas9 CDS</th>
<th>Cas9 Term.</th>
<th>Orient.</th>
<th>sgRNA Backbone</th>
<th>sgRNA Term.</th>
<th>Vector</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>B</td>
</tr>
<tr>
<td>UBI10</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>D</td>
</tr>
<tr>
<td>UBI10</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>Original</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>E</td>
</tr>
<tr>
<td>ICU2</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>F</td>
</tr>
<tr>
<td>AG</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>H</td>
</tr>
<tr>
<td>35S</td>
<td>3</td>
<td>Ags</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>L</td>
</tr>
<tr>
<td>ICU2</td>
<td>1</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>M</td>
</tr>
<tr>
<td>ICU2</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>N</td>
</tr>
<tr>
<td>UBI10</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>O</td>
</tr>
<tr>
<td>UBI10</td>
<td>3</td>
<td>Ags</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>P</td>
</tr>
<tr>
<td>35S</td>
<td>3</td>
<td>Ags</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>Q</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>R</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>S</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>T</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>U</td>
</tr>
<tr>
<td>EC1.2</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>W</td>
</tr>
<tr>
<td>EC1.2</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>X</td>
</tr>
<tr>
<td>EC1.2</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>Y</td>
</tr>
<tr>
<td>EC1.2</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>Z</td>
</tr>
<tr>
<td>CsVMV</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>AA</td>
</tr>
<tr>
<td>CsVMV</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>Original</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>AB</td>
</tr>
<tr>
<td>CsVMV</td>
<td>1</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>AC</td>
</tr>
<tr>
<td>CsVMV</td>
<td>1</td>
<td>Ocs</td>
<td>H2T</td>
<td>Original</td>
<td>7*T</td>
<td>pAGM4723</td>
<td>AD</td>
</tr>
<tr>
<td>CsVMV</td>
<td>2</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>AE</td>
</tr>
<tr>
<td>CsVMV</td>
<td>1</td>
<td>Ocs</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_192</td>
<td>pAGM4723</td>
<td>AF</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pICSL4723</td>
<td>AJ</td>
</tr>
<tr>
<td>EC1.2</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pICSL4723</td>
<td>AK</td>
</tr>
<tr>
<td>MGE1</td>
<td>2</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pICSL4723</td>
<td>AL</td>
</tr>
<tr>
<td>RPS5α</td>
<td>2</td>
<td>Nos</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pICSL4723</td>
<td>AP</td>
</tr>
<tr>
<td>YAO</td>
<td>2</td>
<td>Nos</td>
<td>H2T</td>
<td>EF</td>
<td>7*T</td>
<td>pICSL4723</td>
<td>AQ</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>3</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AS</td>
</tr>
<tr>
<td>enh. 1.2 + EC1.1</td>
<td>4</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AT</td>
</tr>
<tr>
<td>RPS5α</td>
<td>3</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AV</td>
</tr>
<tr>
<td>UBI10</td>
<td>3</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AW</td>
</tr>
<tr>
<td>YAO</td>
<td>3</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AX</td>
</tr>
<tr>
<td>RPS5α</td>
<td>4</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AY</td>
</tr>
<tr>
<td>YAO</td>
<td>4</td>
<td>E9</td>
<td>H2T</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>AZ</td>
</tr>
<tr>
<td>RPS5α</td>
<td>4</td>
<td>E9</td>
<td>H2H</td>
<td>EF</td>
<td>U6-26_67</td>
<td>pICSL4723</td>
<td>BA</td>
</tr>
</tbody>
</table>

Table 4-1: 39 constructs were assembled using Golden Gate
For every construct, the sgRNA targets ADH1 and is transcribed under the control of the U6-26 promoter. Orientation H2T: head-to-head (with Cas9 expression in pICSL472311), H2T: head-to-tail (with Cas9 expression in pICSL47234). Cas9_1: (Mali et al., 2013); Cas9_2: (Fauser et al., 2014); Cas9_3: (Li et al., 2013); Cas9_4: (Cong et al., 2013). pICSL4723 has an overdrive, pAGM4723 does not.

4.2.2 Allyl alcohol enables to select CRISPR-induced Arabidopsis mutations

The 39 Level 2 plasmids were transformed in A. tumefaciens strain GV3101 and used to generate Arabidopsis Col-0 transgenic lines. ‘T1’ refers to independent primary transformants selected from the seeds of the diploid plant; ‘T2’ refers to the T1 progeny.
For each of the 39 constructs, about 100 T2 progenies from six independent T1 lines were tested for allyl alcohol resistance (Figure 4-2). T2 seeds were selected with 30 mM allyl alcohol for two hours. Six survivors (or all survivors if there were less than six) were screened by PCR amplification and capillary sequencing to confirm the mutation in ADH1 at the expected target site. This genotyping step enabled to estimate the percentage of non-mutated plants that escape the allyl alcohol selection. Indeed, I identified some lines surviving the allyl alcohol screen that are heterozygous (ADH1/adh1) or WT (ADH1/ADH1). CRISPR activity is expressed as \[ \frac{(\text{number of allyl alcohol surviving plants}) \times (\% \text{ of homozygous or biallelic mutants confirmed by sequencing among the surviving plants tested})}{(\text{number of seeds sown})} \]. It was measured for four to six independent T2 families, for each of 39 constructs.

Figure 4-2: Evaluation of mutation rates
Constructs were transformed into Arabidopsis accession Col-0 via Agrobacterium tumefaciens strain GV3101. Six independent transformants (T1) were selected using glufosinate. About 100 progeny (T2) of each transformant were selected for allyl alcohol resistance. For each independent T2 family, up to six allyl alcohol resistant plants were genotyped at the ADH1 locus. For each T2 family, the mutation rate was calculated as \[ (\% \text{ of homozygous or biallelic mutants confirmed by sequencing among the surviving plants tested}) / (\text{number of seeds sown}) \].
When more than 75% of the lines survived the allyl alcohol treatment and all the lines genotyped are knock-out alleles with the exact same mutation within one T2 family, the T1 parent was assumed to be a homozygous mutant. Such T2 families are indicated in red.

4.2.3 An overdrive sequence at the T-DNA right border does not affect the CRISPR activity

First generation of Golden Gate vectors contain a short version of the T-DNA right border, that lacks an overdrive sequence. The overdrive sequence can increase the integration efficiency (Peralta et al., 1986). Imperfect integration of the CRISPR T-DNA could explain the low CRISPR efficiency I observed. I assembled the same CRISPR components in pAGM4723, that lacks an overdrive, and in pICSL4723, that has an overdrive (Figure 4-3). In one comparison the presence of the overdrive resulted in slightly better activity but in another one it did not (Figure 4-3). I concluded that the presence of an overdrive does
not influence the CRISPR efficiency. Thus, I could compare constructs independently of the presence of an overdrive.

Figure 4-3: The presence of an overdrive sequence in the T-DNA right border does not affect the CRISPR efficiency

a. Sequence of the right border with (pICSL4723) or without (pAGM4723) an overdrive sequence.
b., c. Each panel represents a vector comparison in the same context. Vectors can be compared within each panel, not from one panel to another. The modules have been assembled by Golden Gate into pICSL4723 (RB+OD, with an overdrive) or pAGM4723 (RB, without an overdrive) and transformed into Col-0 via Agrobacterium tumefaciens strain GV3101. LB: Left Border. SM: Sel. Marker (Glufosinate resistance gene). EC1.2: 1014 bp of the At2g21740 promoter. EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. Cas9_2: (Fauser et al., 2014). E9T: 631 bp of the Pisum sativum rbcS E9 terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA_EF: “extension-flip” sgRNA. U6-26T: 7 bp of the At3g13855 terminator. RB: Right Border.
Border. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. **Bold and underlined:** Most active construct for each panel.

4.2.4 **UBI10, YAO or RPS5a** promoter-controlled Cas9 expression enhances mutation rates

CRISPR-mediated DNA sequence changes must occur in the germ-line to be inherited. Ubiquitous promoters are strongly expressed in most tissues, which theoretically include both somatic- and germ-lines. However, the most commonly used ubiquitous promoter, the 35S promoter from *Cauliflower Mosaic Virus*, is not always expressed at high level in the germ-line (Sunilkumar et al., 2002). I compared 35S with the Arabidopsis **UBI10**
promoters for CRISPR activity. More mutants were recovered using the UBI10 promoter, suggesting it is more active than 35S in the germ-line (Figure 4-4: a, d).
Figure 4-4: UB110, YAO and RPS5a promoter-regulated Cas9 expression enhances mutation rates

Each panel represents a promoter comparison in the same T-DNA context. Promoters can be compared within each panel, but not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an overdrive) or pAGM4723 (RB, without an overdrive) and transformed into Col-0 via Agrobacterium tumefaciens strain GV3101. LB: Left Border. SM: Selectable Marker (Glufosinate resistance gene). 35S: 426 bp of the 35S promoter from Cauliflower Mosaic Virus. UB110: 1327 bp of the At4g05320 promoter. EC1.2: 1014 bp of the At2g21740 promoter. EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. MGE1: 1554 bp of the At5g55200 promoter. AG: 3101 bp of the At4g18960 promoter. ICU2: 625 bp of the At5g67100 promoter. CsVMV: 517 bp of a promoter from Cassava Vein Mosaic virus. RPS5a: 1688 bp of the At3g11940 promoter. YAO: 596 bp of the At4g05410 promoter. Cas9_1: (Mali et al., 2013); Cas9_2: (Fauser et al., 2014); Cas9_3: (Li et al., 2013); Cas9_4: (Cong et al., 2013). E9T: 631 bp of the Pisum sativum rbcS E9 terminator. OcsT: 714 bp of the Agrobacterium tumefaciens octopine synthase terminator. AgsT: 410 bp of the Agrobacterium tumefaciens agropine synthase terminator. NosT: 267 bp of the Agrobacterium tumefaciens nopaline synthase terminator. pU6-26: 205 bp of the At3g13855 promoter. sgRNA: “extension-flip” sgRNA. U6-26T: 7, 67 or 192 bp of the At3g13855 terminator. RB: Right Border. d. Five lines were tested for UB110 and ICU2 and four lines for AG instead of six. f. Five lines were tested for YAO and RPS5a instead of six. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being de novo from the T2 line. Bold and underlined: Most active construct(s) for each panel.

Other groups have found that some germ-line specific promoters can enhance mutation rates in Arabidopsis (Mao et al., 2015; Yan et al., 2015; Wang et al., 2015c; Eid et al., 2016). Following these observations, I tested more germ-line-expressed promoters.

In the combinations tested, I detected low CRISPR activity using the meiosis I-specific promoter MGE1 (Eid et al., 2016) (Figure 4-4: c.), the homeotic gene promoter AG (Hong et al., 2003) (Figure 4-4: d.) and the DNA polymerase subunit-encoding gene promoter ICU2 (Hyun et al., 2014) (Figure 4-4: d.). They were tested with constructs inducing an overall low activity and I do not exclude that they can perform efficiently in other conditions. In one context specifically, ICU2 promoter resulted in moderate activity in four of the six T2 families tested, while only one T2 family showed activity with the UB110 promoter (Figure 4-4: e.).

EC1.2 and an EC1.2::EC1.1 fusion (referred as ‘EC enhanced’ or ‘ECenh’) are specifically expressed in the egg cell and were reported to trigger elevated mutation rates with CRISPR in Arabidopsis (Wang et al., 2015c). In the conditions tested, only ECenh induced homozygous mutants in T1 and at low frequency (Figure 4-4: b., g.). In one comparison, EC1.2 and ECenh performed slightly better than UB110 (Figure 4-4: d.), but in another, they induced lower activity (Figure 4-4: e.). EC1.2 and ECenh were modified during the Golden
Gate cloning process. Such modifications may have altered some cis regulatory elements, rendering them less active.

A promoter from Cassava Vein Mosaic Virus (CsVMV) was reported to mediate CRISPR activity in Brassica oleracea (Lawrenson et al., 2015). It induced more CRISPR activity than UBI10 in two combinations tested (Figure 4-4: d., e.).

I also tested the YAO and RPS5a promoters. Both were reported to boost CRISPR activity in Arabidopsis (Yan et al., 2015; Tsutsui & Higashiyama, 2017). They triggered elevated mutation rates compared with UBI10 (Figure 4-4: f.). In one comparison, RPS5a performed slightly better (Figure 4-4: g.), but in another, YAO performed better (Figure 4-4: h.).

As have others, I conclude that the promoter driving Cas9 expression influences CRISPR-mediated mutation rates (Yan et al., 2015; Wang et al., 2015c; Eid et al., 2016; Tsutsui & Higashiyama, 2017). I observed the best mutation rates using RPS5a, YAO and UBI10 promoters.

4.2.5 Codon optimization of Cas9 and presence of an intron elevate mutation rates

The activity of different constructs with the same promoter can be very different. For instance, RPS5a- and YAO-driven lines were recovered that displayed either high or low activity (Figure 4-4: f., h.). In general, the most active constructs carried Cas9_3 or Cas9_4 alleles. I thus compared four Cas9 alleles side-by-side (Figure 4-5). Cas9_1 is a human codon-optimized version with a single C-terminal NLS (Nuclear Localization Signal) (Mali et al., 2013). Cas9_2 is an Arabidopsis codon-optimized version with a single C-terminal NLS (Fauser et al., 2014). Cas9_3 is a plant codon-optimized version with both N- and C-terminal NLSs, an N-terminal FLAG tag and a potato intron IV (Li et al., 2013). Cas9_4 is a human
codon-optimized version with both N- and C-terminal NLSs and an N-terminal FLAG tag (Cong et al., 2013).
4.2.6 A modified sgRNA triggers CRISPR-induced mutations more efficiently

In the endogenous CRISPR immune system, Cas9 binds a crRNA (CRISPR RNA) and a tracrRNA (trans-acting CRISPR RNA) (Wiedenheft et al., 2012). An artificial fusion of both, called sgRNA (single guide RNA), is sufficient for CRISPR-mediated genome editing (Jinek et al., 2012). sgRNA stability was suggested to be a limiting factor in CRISPR system (Jinek et al., 2013). Chen et al. proposed an improved sgRNA to tackle this issue (Chen et al., 2013). It carries an A-T transversion to remove a TTTT potential termination signal, and an extended Cas9-binding hairpin structure (Figure 4-6 a.). I compared side-by-side the
‘Extended’ and ‘Flipped’ sgRNA (sgRNA\textsuperscript{EF}) with the classic sgRNA (Figure 4-6 b., c.). In two independent comparisons, the efficiency was higher with sgRNA\textsuperscript{EF}. The improvement was not dramatic but sufficient to recommend the use of ‘EF’-modified guide RNAs for genome editing in Arabidopsis. Consistently, a comparison of many sgRNA backbone revealed that a flip and an extension significantly increase the overall CRISPR efficiency (Dang\textit{et al.}, 2015).

**Figure 4-6:** A modified sgRNA is slightly more efficient to trigger mutations. 

\textbf{a.} Original sgRNA proposed by George Church (Mali\textit{et al.}, 2013). Extension-Flip (EF) sgRNA proposed by Lei Qi and Bo Huang (Chen\textit{et al.}, 2013). \textbf{b. and c.} Each panel represents an sgRNA backbone comparison in the same context. sgRNA backbones can be compared within each panel but not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an overdrive) or pAGM4723 (RB, without an overdrive) and transformed into Col-0 via \textit{Agrobacterium tumefaciens} strain GV3101. \textbf{LB:} Left Border. \textbf{SM:} Selectable Marker (Glufoisinate resistance gene). \textbf{CsVMV:} 517 bp of a promoter from \textit{Cassava Vein Mosaic virus}. \textbf{UBI10:} 1327 bp of the At4g05320 promoter. \textbf{Cas9\_2:} (Fauser\textit{et al.}, 2014). \textbf{OcsT:} 714 bp of the \textit{Agrobacterium tumefaciens} octopine synthase terminator. \textbf{U6-26p:} 205 bp of the At3g13855 promoter. \textbf{U6-26T:} 7 bp of the At3g13855 terminator. \textbf{RB:} Right Border. \textbf{c.} Five lines were tested for sgRNA\textsuperscript{EF} and four lines for sgRNA instead of six. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each
dot represents an independent T2 family. **Bold and underlined**: Most active construct(s) for each panel.

### 4.2.7 The 3' regulatory sequences of Cas9 and the sgRNA influence mutation rates

To avoid post-transcriptional modifications such as capping and polyadenylation, sgRNA must be transcribed by RNA polymerase III (Pol. III). Several approaches involving ribozymes, Csy4 ribonuclease or tRNA-processing systems have been proposed but were not tested here (Gao & Zhao, 2014; Tsai et al., 2014; Xie et al., 2015). U6-26 is a Pol. III-transcribed gene in Arabidopsis (Li et al., 2007). I used 205 bp of the 5’ upstream region of U6-26 as promoter and compared a synthetic polyT sequence (seven thymines) and 192 bp of the 3’ downstream region as terminator. A T-rich stretch has been reported to function as a termination signal for Pol. III (Waibel & Filipowicz, 1990).

In seven out of nine side-by-side comparisons, the authentic 192 bp of U6-26 terminator directed a higher efficiency of the construct, as compared to a synthetic polyT termination sequence (Figure 4-7). I speculate that a stronger terminator increases the stability of the sgRNA. For multiplex genome editing, the use of 192 bp per sgRNA will result in longer T-DNAs, increasing the risk of recombination and instability. In addition, I generated constructs with only 67 bp of the U6-26 3’ downstream sequence. Such constructs were not compared side-by-side with the ‘192 bp terminator’, although they enabled modest to high mutation rates (*e.g.* Figure 4-4 f., g.). With these results in mind, I recommend using 67 bp of the 3’ downstream sequence of U6-26 as terminator for the sgRNA.
Figure 4.7: The sgRNA expression regulated by an authentic 3’ regulatory sequence of U6-26 produces greater mutation rates.

a. to c. Each panel represents a terminator comparison in the same context. Terminators can be compared within each panel, not from one panel to another. The modules were assembled into pAGM4723 and transformed into Col-0 via Agrobacterium tumefaciens strain GV3101. LB: Left Border. SM: Selectable Marker (Glufosinate resistance gene). ICU2: 625 bp of the At5g67100 promoter. 35S: 426 bp of the 35S promoter from Cauliflower Mosaic Virus. CsVMV: 517 bp of a promoter from Cassava Vein Mosaic virus. Cas9_2: (Fauser et al., 2014). Cas9_3: (Li et al., 2013). OcsT: 714 bp of the Agrobacterium tumefaciens octopine synthase terminator. AgsT: 410 bp of the Agrobacterium tumefaciens agropine synthase terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA\textsuperscript{EF}: “extension-flip” sgRNA. U6-26T: 7 or 192 bp of the At3g13855 terminator. RB: Right Border. a. Five lines were tested for U6-26T\textsuperscript{67} instead of six. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Bold and underlined: Most active construct(s) for each panel.

Since 3’ regulatory sequences can influence sgRNA stability, I tested if the same was true for Cas9. I compared the Pisum sativum rbcS E9 with two A. tumefaciens terminators commonly used in Arabidopsis: Ocs and Ags (Figure 4.8). I did not observe consistent differences between E9 and Ocs (Figure 4.8). However, in one comparison, E9 outperformed Ags tremendously (Figure 4.8). This is consistent with previous observations that RNA Polymerase II (Pol. II) terminators quantitatively control gene expression and influence CRISPR efficiency in Arabidopsis (Nagaya et al., 2010; Wang et al., 2015c). I propose that a weak terminator after Cas9 enables Pol. II readthrough that could interfere with Pol. III transcription of sgRNAs in some T-DNA construct architectures. This limiting factor can be corrected by divergent transcription of Cas9 and sgRNAs.

Figure 4.8: A weak 3’ regulatory sequence reduces the CRISPR-induced mutation rate.

a. to c. Each panel represents a terminator comparison in the same context. Terminators can be compared within each panel, not from one panel to another. The modules were assembled into
pAGM4723 and transformed into Col-0 via Agrobacterium tumefaciens strain GV3101. **LB**: Left Border. **SM**: Selectable Marker (Glufosinate resistance gene). **EC_enh.**: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. **UBI10**: 1327 bp of the At4g05320 promoter. **Cas9_2**: (Fauser et al., 2014). **Cas9_3**: (Li et al., 2013). **E9T**: 631 bp of the Pisum sativum rbcS E9 terminator. **OcsT**: 714 bp of the Agrobacterium tumefaciens octopine synthase terminator. **AgsT**: 410 bp of the Agrobacterium tumefaciens agropine synthase terminator. **pU6-26**: 205 bp of the At3g13855 promoter. **sgRNA**: “extension-flip” sgRNA. **U6-26T**: 7, 67 or 192 bp of the At3g13855 terminator. **RB**: Right Border. For the comparison using the **UBI10** promoter, the **AgsT** is in combination with **U6-26T**; **OcsT** is in combination with **U6-26T**. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. **Red dot**: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being de novo from the T2 line. **Bold and underlined**: Most active construct(s) for each panel.

4.2.8 Divergent transcription of Cas9 and sgRNA can elevate mutation rates

The Golden Gate Level 1 acceptor vector collection contains seven ‘forward’ expression cassettes and seven ‘reverse’ expression cassettes, which are interchangeable (Engler et al., 2014). I assembled ‘RPS5a:Cas9_4:E9’ and ‘YAO:Cas9_3:E9’ in both the Level 1 vector position 2 forward (pICH47742) and reverse (pICH47811) (Figure 4-1). In one case, CRISPR activity was moderate when Cas9 and sgRNA are expressed in the same direction and high when they are expressed in opposite direction (Figure 4-9: a.). In another configuration, CRISPR activity was very high in both cases (Figure 4-9: b.).
Figure 4-9: CRISPR activity is similar or higher when the sgRNA and the Cas9 expression cassettes are in a head-to-head orientation

**a.** Each panel represents an orientation comparison in the same context. Orientations can be compared within each panel, not from one panel to another. The modules have been assembled by Golden Gate into pICSL4723 (RB+OD, with an overdrive) and transformed into Col-0 via Agrobacterium tumefaciens strain GV3101. **Left Border (LB):** Left Border. **SM:** Selectable Marker (Glufosinate resistance gene. **RPS5a:** 1688 bp of the At3g11940 promoter. **YAO:** 596 bp of the At4g05410 promoter. **Cas9_3:** (Li et al., 2013). **Cas9_4:** (Cong et al., 2013). **E9T:** 631 bp of the Pisum sativum rbcS E9 terminator. **U6-26p:** 205 bp of the At3g13855 promoter. **sgRNA:** “extension-flip” sgRNA. **U6-26T:** 67 bp of the At3g13855 terminator. **RB:** Right Border. **a.** Five lines were tested for H2H instead of six. **b.** Five lines were tested for H2T instead of six. The sgRNA targets ADH1. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. **Red dot:** All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being de novo from the T2 line. **Bold and underlined:** Most active construct(s) for each panel.
These data support the use of a strong terminator after Cas9 (e.g., E9 or Ocs) and express Cas9 and sgRNA in opposite directions.

**4.2.9 Most of the stable double events are homozygous rather than biallelic**

From the mutant screen, 315 allyl alcohol resistant lines were confirmed by capillary sequencing. 59% were homozygous (single sequencing signal, different than ADH1 WT), 11% were heterozygous (dual sequencing signal, one matching ADH1 WT) and 10% were biallelic (dual sequencing signal, none matching ADH1 WT), while 20% were difficult to assign (unclear sequencing signals, either biallelic or due to somatic mutations, but clearly different than WT, heterozygous or homozygous genotypes) (Figure 4-10). The recovery of heterozygous (ADH1/adh1) lines indicates that the loss of a single copy of ADH1 can enable plants to escape the allyl alcohol selection.

![Type of mutations in ADH1](image)

**Figure 4-10: Genotype at ADH1 locus confirmed by capillary sequencing**

For each T2 family tested, up to six allyl alcohol resistant plants were genotyped by capillary sequencing of an sgRNA target (ADH1) PCR amplicon. I retrieved a total of 315 sequences with a mutation. 59% (187) showed a single sequencing signal, different than ADH1 WT and were classified as “Homozygous”. 11% (33) showed an overlap of two sequencing signals, one matching ADH1 WT and one different; and were classified as “Heterozygous”. 10% (31) showed an overlap of two sequencing signals, none matching ADH1 WT; and were classified as “Biallelic”. 64 (20%) showed an overlap of signals different than WT but not clear enough to distinguish; and were classified as “Unknown”. The “Unknown” sequences can be biallelic or due to somatic mutations but are different than WT, heterozygous or homozygous genotypes.
4.3 Discussion

4.3.1 Identification of a CRISPR T-DNA leading to elevated mutation rates

CRISPR emerged in 2012/2013 as a useful tool for targeted mutagenesis in many organisms including plants (Gasiunas et al., 2012; Jinek et al., 2012; Nekrasov et al., 2013). In Arabidopsis, the transgenic expression of CRISPR components can be straightforward, avoiding tedious tissue culture steps. However, the mutation rates were variable. Often many lines have to be genotyped before identifying a mutant (Hyun et al., 2014). Many strategies to enhance the overall CRISPR-induced mutation rate have been proposed (Fauser et al., 2014; Yan et al., 2015; Wang et al., 2015c; Peterson et al., 2016; Tsutsui & Higashiyama, 2017). I conducted a systematic comparison of putative limiting factors including promoters, terminators, codon optimization, sgRNA improvement and T-DNA architecture.

I found that the best promoters to control Cas9 expression are UBI10, YAO and RPS5α. The best terminators were Ocs from A. tumefaciens and rbcS E9 from P. sativum. A plant codon-optimized, intron-containing Cas9 allele outperformed the other alleles tested. A modified sgRNA with a hairpin Extension and a nucleotide Flip, called sgRNAEF, triggered slightly elevated mutation rates. The sgRNA transcription regulation by the authentic 3' regulatory sequence of AtU6-26 results in better CRISPR activity. I get high mutation rates with either 67 bp or 192 bp of terminator and recommend using the shortest (67 bp). I hypothesise that a weak terminator after Cas9 enables RNA-polymerase II readthrough within the sgRNA expression cassette, preventing optimal expression of the sgRNA. Indeed, I recorded an elevated CRISPR-Cas9 efficiency by expressing Cas9 and sgRNA in opposite directions.

Considering the combinations of Cas9 and sgRNA genes tested in this comparison, I recommend using a ‘YAO:Cas9_3:E9’ and a ‘pU6-26:sgRNAEF:U6-26T67’ cassettes in head-to-head orientation. This combination is included in the constructs tested here (Figure 4-9: b.) and enabled us to recover one homozygous mutant in five T1 plants tested. I also noted useful rates with other constructs (e.g. Figure 4-4: f.), indicating that the CRISPR components do not entirely explain the final CRISPR activity. It was recently reported that heat stress increases the efficiency of CRISPR in Arabidopsis (Le Blanc et al., 2017). Environmental conditions may explain fluctuation of the CRISPR activity, independently of the T-DNA architecture.
4.3.2 DSB-induced allelic recombination results in more homozygous than biallelic lines

I was surprised to recover more homozygous than biallelic events. Stable double mutations are the result of two CRISPR events, on the male and female inherited chromosome respectively. In this scenario, lines can be recovered with two different mutations, resulting in a biallelic (e.g. \textit{adh1-2/adh1-3}) genotype, rather than having the same mutation on both chromosomes (e.g. \textit{adh1-1/adh1-1}). DSB-induced homologous recombination occurs between allelic sequences (Gisler et al., 2002). It has been reported that double strand breaks caused by CRISPR-Cas9 can increase this phenomenon (Hayut et al., 2017). Allelic recombination can explain our observation of the same mutation on both copies of \textit{ADH1}. The prevalence of homozygous over biallelic genotypes facilitates the genotyping and is an advantage for targeted mutagenesis using CRISPR-Cas9.

4.3.3 Non-lethal selectable markers to obtain T-DNA-free mutants

I used a glufosinate resistance selectable marker which enables easy selection of transgenic lines in T1. It can be important to segregate away the T-DNA in the CRISPR mutant line for multiple reasons. For instance, a loss-of-function phenotype must be confirmed by complementation of the CRISPR-induced mutation. A CRISPR construct still present in the mutant can target the complementation transgene and interfere with the resulting phenotypes. Selection of non-transgenic lines is possible but complicated with classic selectable markers such as kanamycin or glufosinate resistance, since a selective treatment kills the non-transgenic plants. FAST-Green and FAST-Red provide a rapid non-destructive selectable marker and involve expression of a GFP- or RFP-tagged protein in the seed (Shimada et al., 2010). Transgenic and non-transgenic seeds can be distinguished under fluorescence microscopy (Morineau et al., 2017; Tsutsui & Higashiyama, 2017; Wu et al., 2018c). This facilitates recovery of mutant seeds lacking the T-DNA (Figure 4-11). Homozygous mutants can be identified among the independent T1 lines. Non-fluorescent seeds can be selected from the T1 progeny seeds. The resulting T2 plants are homozygous mutant and non-transgenic.
Figure 4-11: FAST-Red combined with CRISPR to generate T-DNA free mutants

The five T1 lines are independent transformants. They are all hemizygous for the T-DNA. At the sgRNA target site, they can be WT, or display somatic, heterozygous, biallelic of homozygous mutations. All the possibilities are represented here. “Somatic” describes events happening in somatic cells, not inherited in the next generation. As somatic events can happen independently in each cell, they often result in mosaic pattern of mutations across the leaf. One line has homozygous mutation (mut1/mut1). It produces seeds segregating for the T-DNA, visible under microscope if using FAST-Red. The seeds will segregate 3:1 (Red : Non-red) if there is one locus insertion, 15:1 (Red : Non-red) if there are two loci insertion, etc. The T2 progeny of (mut1/mut1) is 100% homozygous for the mutation. The non-red seeds are also T-DNA free.

By comparing diverse combinations of CRISPR components, I identified a CRISPR- and Golden Gate-based method to generate stable Arabidopsis mutant lines in one generation. By trying to elevate mutation rates in Arabidopsis, I found several limiting factors mostly related to Cas9 and sgRNA transcription. Some of these findings can be tested for other plant species and for knock-in breeding. The generation of null alleles via CRISPR is today quick and simple, facilitating the investigation of gene function. For instance, many components of the plant immune system remain uncharacterised likely due to redundancy (Eulgem & Somssich, 2007; Zipfel, 2009). CRISPR enables to target several gene family members at the same time to test for their redundancy.
**Chapter 5 : NLR immune receptors signal via the RPW8-NLR NRG1**

This chapter is largely identical to (Castel et al., 2018), a publication under the CC-BY license.

5.1 Introduction

The plant immune system involves both cell-surface receptors that detect extracellular pathogen-associated molecular patterns (PAMPs) and intracellular receptors that detect pathogen ‘effector’ proteins that if not detected, usually contribute to pathogen virulence (Jones & Dangl, 2006). Most R (Resistance)-genes cloned encode nucleotide-binding, leucine-rich repeat (NLR) immune receptors (Kourelis & Van Der Hoorn, 2018). These receptors are widely deployed during breeding for crop disease resistance (Borhan et al., 2010; Jones et al., 2014; Witek et al., 2016). A better understanding of NLR-mediated immunity could facilitate their deployment to safeguard crops from pathogens, reducing the need for chemical applications.

Most NLRs comprise an N-terminal domain, a central NB-ARC (NB (nucleotide-binding) domain shared by APAF-1 (apoptotic protease activating factor 1), plant R-proteins and CED-4 (cell death protein 4) domain and C-terminal LRRs (Leucine-rich Repeats). At their N-termini, they usually carry a TIR (Toll/Interleukin-1 receptor/Resistance protein) domain or non-TIR N-terminus, often with CCs (Coiled-Coils). A phylogenetically distinct NLR subset carries an N-terminal RPW8 (Resistance to Powdery Mildew 8) domain. The corresponding TIR-NLR, CC-NLR and RPW8-NLR proteins, upon activation, trigger a complex network of responses including gene induction, production of reactive oxygen species and SA (salicylic acid), transcriptional reprogramming and a form of cell death called the HR (hypersensitive response), resulting in resistance (Jones & Dangl, 2006).

The RPW8-NLR NbNRG1 was first identified in *Nicotiana benthamiana* as required for resistance to *Tobacco Mosaic Virus* (TMV) mediated by the *N* gene (Peart et al., 2005). *N* encodes a TIR-NLR that activates resistance upon recognition of the TMV replicase component p50. *NRG1* is widespread in angiosperms suggesting an important role in immunity (Collier et al., 2011; Shao et al., 2014).

Two additional ‘helpers’ have been described in plants. ADR1 is part of a conserved clade within angiosperms and contributes to function of RPP2, RPP4, RPS2, SNC1, CHS3 and...
RRS1 in Arabidopsis (Bonardi et al., 2011; Dong et al., 2016). Interestingly, ADR1s are also RPW8-NLRs, phylogenetically close to the NRG1 clade (Collier et al., 2011; Shao et al., 2014). The NRC helper NLRs in Solanaceae are required for the sensor NLRs Rpi-blb2, Mi2-5, Sw5b, R8, R1, Prf, Rx, Bs2 and CNL-119900, amongst others (Wu et al., 2017). In both cases, the helper clade is relatively conserved and partially redundant, while their associated sensors are expanded and diversified. Phylogenetic analyses revealed that the conserved NRG1 clade and TIR-NLRs co-occur in angiosperm genomes: all clades that lack NRG1 also lack TIR-NLRs (Collier et al., 2011; Shao et al., 2016). Conceivably, all RPW8-NLRs are helper NLRs, with ADR1 signalling downstream of CC-NLRs and TIR-NLRs and NRG1s signalling downstream of TIR-NLRs only.

I set out to test if NRG1 is a helper for other TIR-NLRs. I generated nrg1a-nrg1b loss-of-function double mutants in Arabidopsis and a mutant of the single NRG1 copy in N. benthamiana using CRISPR/Cas9 and tested for loss of TIR-NLR activities. I found that NRG1 is at least partially required for the signalling of the TIR-NLRs WRR4A, WRR4B, RPP1, RPP2, RPP4 and the NLR pairs CHS1/SOC3, CHS3/CSA1 and RRS1/RPS4. In contrast, nrg1 loss-of-function does not compromise the CC-NLR R proteins RPS5 and MLA. RPM1 and RPS2 (CC-NLRs) function is slightly compromised in an nrg1 mutant. Thus, NRG1 is required for full TIR-NLR function and contributes to the signalling of some CC-NLRs. Recent studies also suggested a requirement of NRG1 for RPP1-mediated signalling and identified Roq1, an additional NRG1-dependent TIR-NLR (Brendolise et al., 2018; Qi et al., 2018). Surprisingly, I found that NRG1 is required the HR but not for the bacterial disease resistance conferred by RRS1/RPS4. I discuss the potentially complementary roles of NRG1 and its sister clade ADR1. I propose that some NLRs signal via NRG1 only (e.g. CSA1/CHS3), some via ADR1 only (e.g. RPS2) and some via both (e.g. RRS1/RPS4).

5.2 Results

5.2.1 Identification of ADR1 in N. benthamiana genome

I used the NB-ARC amino acid sequence of NbNRG1 (DLPQLQELKVKLLEEKEKVVVELSAPAGCGKTLAAAMLCLQEEDDIKDKYRDIFVTVSKKANIRIVGEIFE
MKGYKGPDFASEHAAVCQLNNLRSTSQPVVLVDVWSESDVIESFIFQIPFGKILVTSRVFPPFKFDTYKLNLLSEKDAKALFYSSAFKDSIPYVQLDLVHKAVRSCCGFPLALKVVGRSCLGQPELIWFNRVMLQSKRQILFPTENDLLRTLRSIDALDEIDLYSEATTTLRDCYLDGLSFPEDHRIHAAAIL

97
I identified three transcripts: *Niben101Scf02118g00018.1* (**NbNRG1**), *Niben101Scf03844g01015.1* (**NbNRG2**, pseudogene) and *Niben101Scf02422g02015.1* (uncharacterised). The other results were not considered due to low coverage or low identity. *Niben101Scf02422g02015.1* predicted transcript does not start with a START codon and does not finish with a STOP codon so could be incomplete. *Niben101Scf02422g02015.1* was used as blastn query on ‘N. benthamiana Genome v1.0.1 Contigs’ to identify the gene in its genomic context. I retrieved *Niben101Scf02422g02015.1* genomic sequence on contig *Niben101Scf02422Ctg032* and identified a START codon 23 bp upstream and a STOP codon 564 bp downstream of *Niben101Scf02422g02015.1*, framing a 2487 bp gene including exons and introns (**Figure 5-1**). The resulting protein was analysed using SMART protein domain annotation resource (Letunic et al., 2012). It contains an N-terminal RPW8 domain, a central NB-ARC domain followed by LRRs. A phylogenetic reconstruction of *Niben101Scf02422g02015.1* NB-ARC domain along with the NB-ARC domains of AtADR1, AtADR1-L1, AtADR1-L3, AtADR1-L3, AtNRG1A, AtNRG1B, AtDAR5, NbNRG1 and AtZAR1 as outer group places *Niben101Scf02422g02015.1* in the ADR1 clade (**Figure 5-1**).
Figure 5: Phylogenetic tree of RPW8-NLRs (NB-ARC domain) from Arabidopsis and Nicotiana benthamiana

The sequences used here are from TAIR10 for Arabidopsis and SolGenomicsNetwork for N. benthamiana. The CC-NLR AtZAR1 was used an outer group. NB-ARC domain amino acid sequences were aligned using ‘MUSCLE’. The phylogenetic reconstruction was made using the Maximum Likelihood method. The robustness of the tree was tested by Bootstraps (100 replicates). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 10 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 57 positions in the final dataset. Evolutionary analyses were conducted with the software MEGA7. b. Sequence of NbAD1R1 in genomic context from contig Niben101Sco2422Ctg032. Blue: cDNA of the annotated Niben101Sco2422Ctg032.1 gene from Sol Genomic Network. Green: Putative ADR1 CDS framed with ha START and STOP codon. The figure was made using CLC Main Workbench Version 8.0.1. MUSCLE was used to generate phylogenetic trees using the Maximum Likelihood methods with 100 bootstraps. This analysis was carried out using the MEGA7 software.
5.2.2 CRISPR enables recovery of nrg1 mutants in Arabidopsis and N. benthamiana and a wr4b mutant in Arabidopsis

Four CRISPR constructs (CRISPR-1, CRISPR-2, CRISPR-3 and CRISPR-4) were assembled using Golden Gate cloning method and expressed via Agrobacterium tumefaciens strain GV3101 in Arabidopsis Ws-2, Col-0 or N. benthamiana (Figure 5-2).

Figure 5-2: CRISPR enables to generate nrg1 mutants in plants
a. Schematic representation of the T-DNA containing a plant selectable marker (blue), a Cas9 (yellow) and a gRNA (orange) expression cassettes. The modules were assembled using the Golden Gate cloning method. b. Constructs were expressed in plants via Agrobacterium tumefaciens strain GV3101. In Ws-2, one progeny carries a T deletion in NRG1A and a T insertion in NRG1B. In Col-0, one progeny carries a T deletion in NRG1A and a G deletion in NRG1B. In Nicotiana benthamiana, one progeny carries an AAGAGAA deletion in NRG1. These mutations cause a codon reading frame shift resulting in an early stop codon in the NB-ARC domains. The sequences were analysed using CLC Main Workbench Version 7.7.1. Red Square: PAM (Protospacer Adjacent Motif). LB: Left Border, pUBI10: 1327 bp of the At4g05320 promoter, pRPS5a: 1688 bp of the At3g11940 promoter, 35S: 426 bp of the 35S promoter from Cauliflower Mosaic Virus, E9T: 631 bp of the Pisum sativum rbcS E9 terminator, U6-26p: 205 bp of the At3g13855 promoter, U6-26t: 67 bp of the At3g13855 terminator, RB: Right border.

For CRISPR-1, three sgRNAs targeting both AtNRG1A and AtNRG1B (GCTCATTACCAAACCTGAAA[nGG], GTGGAAGCTGGTCTGAAG[nGG]) and
GATGATTTGTTCTCATCGAAA[nGG]) were designed and assembled by PCR to a sgRNA backbone in a Golden Gate compatible fashion (Methods S1) (Castel et al., 2019). sgRNAs were assembled with AtU6-26 promoter in the Golden Gate Level 1 vectors pICH47751, pICH47761, pICH47772 respectively. A human codon optimized allele of Cas9 was assembled with the AtUBI10 promoter and Nos terminator in Golden Gate compatible Level 1 vector pICH47742. These four level 1 vectors, along with a FAST-Red selectable marker in Level 1 vector pICH47732, were assembled in the binary vector pAGM4723, resulting in CRISPR-1 final vector.

For CRISPR-2, only one sgRNA targeting both AtNRG1A and AtNRG1B (GTGGAAAGCTGGTCTGAAG[nGG]) was used. A plant codon-optimized allele of Cas9 containing a potato IV2 intron was assembled with AtRPS5a promoter and Pisum sativum rbcS E9 terminator in pICH47811. There were cloned in binary vector pICSL4723 in a similar fashion as CRISPR-1, resulting in CRISPR-2 final vector.

For CRISPR-3, one sgRNA targeting NbNRG1 (CAGTATTCGATGACATCGAG[nGG]) was used. A human codon-optimized allele of Cas9 was assembled with 35S promoter and Mas terminator in Golden Gate Level 1 vector pICH47742. The two vectors, along with a glufosinate resistance plant selectable marker in Level 1 vector pICH47732, were assembled in pICSL4723, resulting in CRISPR-3 final vector.

For CRISPR-4, two sgRNAs targeting WRR4B in Arabidopsis Ws-2 (AATCGCTTCCGTGAGACAGCTC[nGG] and TACATAGTGTACTATAAA[nGG]) were used. A plant codon optimized allele of Cas9 was assembled with AtUBI10 promoter and Ocs terminator in Golden Gate Level 1 vector pICH47742. The two vectors, along with a FAST-Red selectable marker in Level 1 vector pICH47732, were assembled in pAGM4723, resulting in CRISPR-4 final vector.

For CRISPR-2 sgRNA specifically, 67 bp of the AtU6-26 terminator was included by PCR at the sgRNA 3’ end. Expression of CRISPR-1 in Arabidopsis Ws-2 resulted in a c.1153delG mutation in NRG1A and a c.1159_1160insT mutation in NRG1B. Expression of CRISPR-2 in Arabidopsis Col-0 resulted in a c.1153delG mutation in NRG1A and a c.1159_1160insTATTTTTGGTCCTC mutation NRG1B. Expression of CRISPR-3 in N. benthamiana resulted in a c.638_644delAAAGAAA mutation in NbNRG1. Expression of CRISPR-4 in Arabidopsis Ws-2 resulted in a c.181_182insA mutation in WRR4B. All these mutations cause a codon reading frame shift and early stop codons in of before the NB-ARC domain encoding region. Progenies of the mutants were analysed and I selected lines
without T-DNA and nrg1 or wrr4b mutation at homozygous state. I refer to these lines as Ws-2_nrg1a-nrg1b, Col-0_nrg1a-nrg1b, N. benthamiana_nrg1 and Ws-2_wrr4b.

5.2.3 NRG1 is required for RRS1/RPS4-mediated HR but not bacterial resistance in Ws-2

In Arabidopsis Ws-2, RRS1-R and RPS4 comprise a TIR-NLR pair that recognises the effectors PopP2 from Ralstonia solanacearum and AvrRps4 from Pseudomonas syringae. Their paralogs RRS1B and RPS4B also recognise AvrRps4 but not PopP2. RRS1/RPS4 and RRS1B/RPS4B activation results in HR, defence gene induction and bacterial resistance (Narusaka et al., 2009; Saucet et al., 2015). I tested if these responses require NRG1. I delivered the effector AvrRps4 using Pseudomonas syringae pv. tomato strain DC3000 and assessed cell death (Figure 5-3). I observed fewer leaves displaying cell death in Ws-2_nrg1a-nrg1b than in Ws-2 WT. Cell death was not completely abolished in Ws-2_nrg1a-nrg1b. However, some combinations indicate that DC3000 can induce AvrRps4-independent cell death, such as DC3000 with empty vector in wild type, or with AvrRps4 in an rrs1-rps4-rps4b mutant. DC3000 carries a diverse set of effectors that influence immunity and could mask the AvrRps4-induced HR.
Figure 5-3: Cell death assay using *Pseudomonas syringae* strain DC3000
Arabidopsis leaves were infiltrated with *Pseudomonas syringae* pv tomato strain DC3000 expressing indicated effectors, OD600 = 0.05 (a.) or 0.1 (b.). Half-leaf was infiltrated. Pictures were taken 24 hours post infiltration. The red stars show a host tissue collapse. The numbers indicate the number of plants displaying cell death out of the total number of plants tested. OD: optical density at 600 nm. hpi: hour(s) post-inoculation.

To test for HR after delivery of a single effector, I used the *Pseudomonas fluorescens* strain Pfo-1 carrying a type III secretion system (Pfo-1 EtHAn) (Thomas et al., 2009). With the Pfo-1 system, AvrRps4 induces a specific HR in Ws-2 and Col-0, due to the activation of RRS1/RPS4 and RRS1B/RPS4B (Saucet et al., 2015). The RRS1/RPS4 and RRS1B/RPS4B-mediated HR is completely lost in both Col-0_nrgr1a_nrgr1b and Ws-2_nrgr1a_nrgr1b backgrounds (Figure 5-4). This observation confirms that the cell death observed with
DC3000_AvrRps4 in Ws-2_nrg1a-nrg1b, Ws-2_eds1 and Ws-2_WT is independent of AvrRps4 (Figure 5-3). However, AvrRpt2, AvrRpm1 and AvrPphB, which activate the CC-NLRs RPS2, RPM1 and RPS5 respectively (Debener et al., 1991; Kunkel et al., 1993; Simonich & Innes, 1995), induce HR after delivery from Pf0-1 even in an nrg1a-nrg1b mutant background. Both AtNRG1A and AtNRG1B can complement Ws-2_nrg1a-nrg1b loss of TIR-NLR-mediated HR, indicating that they are redundant for this function (Figure 5-3).

**Figure 5-4: NRG1 is required for RRS1/RPS4 mediated HR**

a. Arabidopsis leaves were infiltrated with *Pseudomonas fluorescens* strain Pf0-1 expressing indicated effectors, OD600 = 0.2. For Ws-2, the whole leaf was infiltrated. For Col-0, half-leaf was infiltrated. Pictures were taken 24 hours post inoculation. The red arrow shows a hypersensitive response (HR), indicating NLR activation by cognate effector. The numbers indicate the number of plants displaying HR out of the total number of plants tested. PopP2- and AvrRps4-mediated HR is lost in Ws-2_nrg1a-nrg1b and Col-0_nrg1a-nrg1b. AvrRps4KRYY (AvrRps4KRYYAAAA) and PopP2C321A are mutant alleles unable to trigger HR in WT plants, used as negative control for HR in
WT. They did not trigger HR indeed. AvrRpt2, AvrRpm1 and AvrPphB are CC-NLR-activating effectors used as positive control for HR in nrg1. They all trigger HR in both WT and nrg1 mutant lines. HR is recovered in lines complemented with either NRG1A or NRG1B, tagged with mCherry, GFP, V5 of HF.

I conducted an electrolyte leakage assay to test whether the HR caused by the CC-NLRs RPS2, RPM1 and RPS5 was quantitatively reduced in Col-o_nrg1a-nrg1b (Figure 5-5). Indeed, RPS2- and RPM1-, but not RPS5-mediated HR, as assayed by ion leakage assay, are also partially reduced at an early time point. This reduction is not seen at later time points for RPM1. These data indicate that NRG1 is fully required for the HR mediated by the TIR-NLRs RRS1/RPS4 and RRS1B/RPS4B, partially required for the HR mediated by the CC-NLRs RPM1 and RPS2, and not required for the HR mediated by the CC-NLR RPS5.

Figure 5-5: NRG1 is partially required for RPS2- and RPM1-mediated ion leakage
HR was quantified by the electrolyte leakage assay. Col-o WT or Col-o_nrg1a-nrg1b plants were infiltrated in the same conditions as for the pictures. 16 discs were collected, rinsed and immersed in 10 ml of water. The electroconductivity was measured at 4, 10 and 24 hpi. Boxplots represent nine data points (three biological replicates x three technical replicates). Colours indicate biological replicates (plants grown at different times); shapes indicate technical replicates (different plants inoculated with the same inoculum). Significance was calculated with t-tests and the p-value is indicated by n.s. (non-significant, > 0.05), * (<0.05), ** (<0.01) or *** (<0.001). WT: wild type. HR: hypersensitive response. hpi: hour(s) post inoculation.

I used the Pseudomonas syringae pv. tomato strain DC3000 to test for bacterial resistance (Xin & He, 2013). Bacterial growth on Ws-2 is reduced for DC3000 expressing AvrRps4, due to recognition by RRS1/RPS4 and RRS1B/RPS4B. Surprisingly, this AvrRps4-dependent reduced growth is unaltered in an nrg1a-nrg1b double mutant (Figure 5-6). Although DC3000 can trigger AvrRps4-independent cell death, the AvrRps4-dependent HR is lost in nrg1a-nrg1b (Figure 5-3, Figure 5-4, Figure 5-5). Thus, loss of NRG1 function abolishes HR
from activation of RRS1/RPS4 and RRS1B/RPS4B but does not compromise activation of disease resistance to bacteria in Arabidopsis. The maintenance of RPS2-mediated HR in the absence of NRG1 correlates with resistance to DC3000 expressing AvrRpt2 in Col-0_nrg1a-nrg1b (Figure 5-6).

**Figure 5-6: NRG1 is not required for bacterial resistance**

Bacterial growth assay in Arabidopsis using *Pseudomonas syringae* pv tomato strain DC3000 (DC3000). One leaf per plant was infiltrated with DC3000 expressing or not (a) AvrRps4 or (b) AvrRpt2 at OD600 = 0.001. Bacteria quantification was done just after infiltration (0 dpi) and 3 days post inoculation (3 dpi). Each dot represents one individual plant. Colours and shapes represent three biological replicates. a. Letters above the bars indicate significant differences (P<0.05) as determined by a one-way ANOVA followed by post hoc Tukey HSD analysis. b. The growth of DC3000_AvrRpt2 was not different in Col-0 WT and Col-0_nrg1a-nrg1b, as indicated by a non-significant (n.s) t-test (p-value = 0.418). Rep: Biological replicate. CFU: colony forming unit.

Activation of RRS1/RPS4 also results in induction of salicylic acid (SA)-related genes such as ICS1 and PR1 (Sohn et al., 2014). SA plays a crucial role as a signalling molecule that activates plant immunity (Durner et al., 1997). I measured changes in SA levels, the expression of SA-responsive PR1 and SA-biosynthetic gene ICS1 in response to Pf0-1_AvrRps4, and as a control, to Pf0-1_AvrRps4krvy, a Pf0-1 strain carrying non-recognized allele of AvrRps4 (Sohn et al., 2012). PR1 and ICS1 both show RPS4/RRS1/AvrRps4-dependent gene induction even in the absence of NRG1 (Figure 5-7a and c). SA is weakly induced by Pf0-1_AvrRps4krvy, likely due to PTI, and strongly induced by Pf0-1_AvrRps4
SA induction is not reduced in Ws-2_nrg1a-nrg1b. Thus, NRG1 is not required for RRS1/RPS4-mediated elevation of SA levels and induction of defence genes.
Figure 5-7: NRG1 is not required for salicylic acid pathway activation

**a. and c.** Induction of ICS1 and PR1 upon Pf0-1 strain infiltration. Arabidopsis leaves were infiltrated with bacteria at OD600 = 0.2. Samples were collected 4 hpi for ICS1 expression and 8 hpi for PR1 expression. Values represent the expression level as compared to EF1α, using the ‘double delta Ct’ method. Three lines were used and bulked for each treatment. Error bars represent standard error of three technical replicates. Three biological replicates all showed induction of ICS1 and PR1 by AvrRps4 in both WT and Ws-2_nrg1a-nrg1b.

**b.** Induction of salicylic acid upon Pf0-1 strain infiltration. Arabidopsis leaves were infiltrated with bacteria at OD600 = 0.05. Samples were collected 24 hpi. SA was extracted from 10 mg of dry weight and SA was quantified by Ultra high-pressure liquid chromatography. Colours and shapes represent three biological replicates, each represented by two independent extractions from the same set of infiltrated leaves. Significance was calculated with t-tests and the p-value is indicate by n.s. (> 0.05), * (<0.05), ** (<0.01) or *** (<0.001). hpi: hour(s) post inoculation. n.i.: non-infiltrated

5.2.4 Arabidopsis nrg1 mutants show impaired TIR-NLR-dependent resistance to oomycete pathogens

Arabidopsis accession Col-0 resists *Hyaloperonospora arabidopsidis* (downy mildew or Hpa) races Emoy2 and Cala2 via the TIR-NLRs RPP2 and RPP4 respectively (Van Der Biezen et al., 2002; Sinapidou et al., 2004). In Ws-2, Hpa race Cala2 is resisted via the TIR-NLR RPP1_WsA (Botella et al., 1998). I tested whether Cala2 and Emoy2 resistance requires NRG1. I found that Cala2 hyphal growth in cotyledons was more extensive in Ws-2_nrg1a-nrg1b than in Ws-2 WT for Cala2 (Figure 5-8).
Figure 5-8: Ws-2_nrg1a-nrg1b mutant is partially compromised in resistance to RPP1-mediated resistance to Cala2

7-day old plants were sprayed inoculated with *Hyaloperonospora arabidopsidis* Cala2, resisted by RPP1-Ws-1A in Ws-2. Plants were phenotyped 7 days after inoculation. Plants were stained with trypan blue to visualise hyphae growth. Six cotyledons for each genotype were examined under microscope. **a-f.** Six cotyledons from Ws-2_nrg1a-nrg1b inoculated plants. **g-h.** Six cotyledons from Ws-2 WT inoculated plants. Images were annotated using the software Affinity Photo. Hyphae and sporangiophores are coloured in red and green. The original pictures are showed on the bottom.
Moreover, more spores were produced on nrg1a-nrg1b infected plants, but not to the level of eds1 (enhanced disease resistance 1, an Arabidopsis mutant strongly affected in resistance mediated by TIR-NLRs (Parker et al., 1996)) (Figure 5-9). Some spores were observed in the resistant WT plant; these could either be spores that persist from the inoculation or fresh spores from rare sporangiophores sometimes observed in cotyledons of WT plant (Figure 5-8g). I conclude that RPP1- and RPP2-mediated resistance to Cala2 and RPP4-mediated resistance to Emoy2 partially requires NRG1.

Figure 5-9: nrg1 mutants are partially compromised in resistance to downy mildew 7-day old plants were sprayed inoculated with Hyaloperonospora arabidopsidis race Emoy2 or Cala2. Plants were phenotyped 7 days after inoculation. a, b, c. Full resistance was observed in Col-0 and Ws-2 WT, full susceptibility was observed in Col-0 eds1 (eds1-2) and Ws-2 eds1 (eds1-1) and intermeditated phenotype in Col-0 nrg1a-nrg1b and Ws-2 nrg1a-nrg1b. Ler-0 and Oy-0 were used as susceptible controls for Cala2 and Emoy2 infections respectively. (*) stars indicate significant differences (t-test at 0.05 significance) between WT and nrg1 mutant plants. Data are expressed in spores per plant.

Arabidopsis Col-0 and Ws-2 also contain TIR-NLRs that confer resistance to the oomycete Albugo candida (Ac), the cause of white rust. For instance, the TIR-NLR WRR4A confers resistance to the white rust race AcEm2 (Borhan et al., 2008). I expressed an allele of WRR4A from Arabidopsis accession Oy-0 in the AcEm2-compatible accession Ws-2 and observed full resistance indeed (Figure 5-10). However, Ws-2 nrg1a-nrg1b transformed with WRR4A was partially susceptible, showing that WRR4A requires NRG1 to activate full resistance to white rust. All transgenic lines expressed WRR4A, at different levels within
the same range (Figure 5-10). Thus, the partial loss of resistance in Ws-2_nrg1a-nrg1b_WRR4A is not due to low expression of WRR4A but rather to the loss of NRG1.

Figure 5-10: nrg1 mutants are partially compromised in resistance to white rust
3- to 5-week old plants were sprayed inoculated with Albugo candida. Plants were phenotyped 12 days after inoculation. Abaxial and adaxial picture of the same leaf are showed. Numbers indicate
the number of individual plants showing similar phenotype out of the number of plants tested. **a.** Two independent lines for Ws-2 WT and two independent lines for Ws-2_nrg1α-nrg1β were tested for *Albugo candida* race AcEm2 resistance. 6 plants were tested for each line. Ws-2 is susceptible, while WRR4A transgenic lines are resistant. Ws-2_nrg1α-nrg1β expressing WRR4A are still partially susceptible to AcEm2. Oy-0 allele of WRR4A was used here. **b.** RT-qPCR was conducted in WRR4A transgenic lines. Two lines were tested for each independent transformants. Error bars indicate standard error of three technical replicates. WRR4A is expressed in those lines at a higher level than in the fully resistant Col-0 WT. **c.** Eight or twelve plants for each genotype were tested with *Albugo candida* races Ac2V, Ac7V and AcBoT. Ws-2_eds1 (eds1-1) are fully susceptible to *Albugo candida* races Ac2V and Ac7V, while Ws-2_WT are resistant. Ws-2_nrg1α-nrg1β and Ws-2_wrr4b are partially susceptible to Ac2V and Ac7V. AcBoT resistance is not affected in Ws-2_nrg1α-nrg1β and Ws-2_wrr4b mutants, while Ws-2_eds1 is partially susceptible. Red arrows indicate presence of white pustules, resulting from production of zoospores. **T2:** line from the second generation after transformation. #: reference number of the independent transformant line.

Ws-2 lacks WRR4A but contains a paralog, called WRR4B, which confers resistance to the white rust races Ac2V, Ac7V and AcBoT (Cevik et al., 2019). I expressed a Cas9 construct targeting WRR4B with two sgRNAs and recovered a homozygous mutant line. This line is partially susceptible to Ac2V and Ac7V, but still resists AcBoT (Figure 5-10). Similarly, Ac2V and Ac7V, but not AcBoT can complete their life cycle on Ws-2_nrg1α-nrg1β. Thus, both WRR4B and NRG1 are required for full Ac2V and Ac7V resistance, and the phenotypes of both loss-of-function mutants are similar, suggesting that WRR4B also requires NRG1 to confer white rust resistance.

### 5.2.5 NRG1 is required for SOC3/CHS1- and CSA1/CHS3-induced HR in *Nicotiana benthamiana*

CSA1 and CHS3 are two adjacent TIR-NLR-encoding genes from Arabidopsis. CHS3 carries an integrated C-terminal LIM domain. chs3-2D is a gain-of-function allele that confers a CSA1-dependent autoimmune phenotype in Arabidopsis (Xu et al., 2015). Similarly, SOC3 and CHS1 are a TIR-NLR/TIR-NB (TIR, NB-ARC but no LRR) pair encoded by two adjacent genes in Arabidopsis. They interact to modulate immunity (Zhang et al., 2016c). I transiently expressed CSA1 and chs3-2D or SOC3 and CHS1 in leaves of *N. benthamiana* and in the Nb_nrg1 mutant. I found that CSA1 and chs3-2D can induce HR in *N. benthamiana* but not in Nb_nrg1 (Figure 5-11).
I also observed an HR triggered by SOC3 and CHS1 alone or in combination, which is lost in the absence of NRG1. Although the CHS1/SOC3 and CHS3/CSA1 phenotypes are weak in WT, they never induce HR in Nb_nrg1. The CC domain of the CC-NLR MLA from barley and the D505V ‘autoimmune’ allele of the CC-NLR RPM1 from Arabidopsis activates HR in N. benthamiana (Gaoa et al., 2011; Maekawa et al., 2011). Transient expression of these alleles results in HR in both N. benthamiana and Nb_nrg1. I tested the expression of CHS1 and SOC3 and found that the proteins are still expressed in Nb_nrg1 (Figure 5-12). As CHS3-2D
and CSA1 are not tagged, I tested the expression of their mRNA and found that there were also expressed in Nb_\_nrg1.

**Figure 5-12: Expression of CHS1/SOC3 and CHS3/CSA1 in Nicotiana benthamiana**

CHS1-V5, SOC3-HF, CHS3-2D and CSA1 were transiently expressed in *Nicotiana benthamiana*. For CHS1-V5 and SOC3-HF, proteins were extracted 3 dpi and revealed by Western Blot. Both are detected at their expected size, i.e. 50.39 kDa for CHS1-V5 and 124.66 kDa for SOC3-HF. Two bands are observed for SOC3-HF. CHS3-2D and CSA1 are not tagged and cannot be revealed by immunoblot. Their expression was measured by RT-PCR. RNA was extracted 3 dpi. RT-PCR was conducted to amplify a 176 bp fragment in CHS3-2D and a 643 bp fragment in CSA1. They were both expressed in *N. benthamiana* WT and nrg1 mutant after infiltration.

I conclude that NRG1 is required for the HR initiated by the TIR-NLR pair CSA1-CHS3 and the TIR-NLR/TIR-NB pair SOC3-CHS1 but not by the CC-NLRs MLA7 and RPM1 in *N. benthamiana*.

### 5.3 Discussion

**5.3.1 NRG1 is a conserved clade of RPW8-NLR that contributes to TIR-NLR signalling**

NRG1 was originally identified in *N. benthamiana* as required for resistance mediated by the TIR-NLR N (Peart et al., 2005). NRG1 is broadly conserved within angiosperms, forming the so-called NRG1 clade (Collier et al., 2011). NRG1 could be a helper clade for many NLR sensors. Interestingly, there is a correlation of NRG1 and TIR-NLR occurrence within plant genomes; the NRG1 lineage is missing in monocots and Lamiales, which also lack TIR-NLRs (Collier et al., 2011). On the other hand, the NLRome of *Amborella trichopoda*, an early
diverging lineage among angiosperms, contains both TIR-NLRs and NRG1 (Shao et al., 2016). This observation suggests that NRG1 could be specifically required for all TIR-NLR-mediated immunity. To test this hypothesis, I mutated NRG1 from N. benthamiana and Arabidopsis and tested for loss or reduction of TIR-NLR function. N. benthamiana and Arabidopsis carry one and two copies of NRG1 respectively (NbNRG1 and AtNRG1A/AtNRG1B). I found that NbNRG1 is required for the HR triggered by the TIR-NLR pairs SOC3/CHS1 and CSA1/CHS3-2D (Figure 5-11). Signalling of the CC-NLRs MLA7 and RPM1 still results in an HR in Nb_nrg1 that is indistinguishable from WT. In Arabidopsis, CHS3 signals via NRG1 but not CHS1, which signals via ADR1 (Wu et al., 2018b). It is possible that NbADR1 is too diverged from the three Arabidopsis ADR1 alleles to support CHS1 function and complement the loss of NRG1.

In Arabidopsis accession Ws-2, transgenic expression of WRR4A confers NRG1-dependent resistance to the A. candida race AcEm2 (Figure 5-10). Moreover, resistance to A. candida races Ac2V and Ac7V is strongly reduced in an nrg1a-nrg1b double mutant (Figure 5-10). This resistance is mediated in part by the TIR-NLR WRR4B, suggesting that the WRR4B paralog of WRR4A can also signal via NRG1. I also found that resistance mediated by the TIR-NLRs RPP1, RPP2 and RPP4 is partially dependent on NRG1 in Arabidopsis (Figure 5-9). However, the requirement of NRG1 by RPP1, RPP2 and RPP4 for downy mildew resistance, although significant, was weak. In addition, RPP4 function was not affected in an independent experiment (Wu et al., 2018b). Thus, other helpers, such as ADR1, already reported to contribute to RPP4 function (Bonardi et al., 2011), likely also contribute to RPP1, RPP2 and RPP4 function.

Using the Pf0-1 system, I found that the HR induced by RRS1/RPS4 and RRS1B/RPS4B is fully dependent on NRG1 in Arabidopsis (Figure 5-4). Surprisingly, I observed a weak reduction of the HR induced by AvrRpt2 (via CC-NLR RPS2) and AvrRpm1 (via CC-NLR RPM1) (Figure 5-5). The HR induced by AvrPphB (via CC-NLR RPS5) is unaltered in Col-0_nrg1a-nrg1b. It was previously reported that another CC-NLR, Rx2 from potato, also requires NRG1 redundantly with ADR1 to confer resistance to Potato Virus X (Collier et al., 2011).

NRG1 is indeed required for full function of all tested TIR-NLRs. However, the function of some CC-NLRs is also compromised in the absence of NRG1. Arabidopsis genomes contains two NRG1 copies in tandem while N. benthamiana has only one copy. This tandem duplication must have a biological relevance. The HR mediated by
RRS1-RPS4 is NRG1-dependent and both alleles are sufficient (Figure 5-4). It indicates that NRG1A and NRG1B are redundant for this function. However, NRG1A and NRG1B share less than 72% amino acid identity. It could be that NRG1A and NRG1B are not redundant for recognition of other NLRs. In addition, the polymorphism can enable to avoid targeting of both copies by the same effector. In both cases, the duplication event resulted in one more node in the NRG1 helper network, making it more robust.

5.3.2 Disease resistance is possible without HR

Despite the absence of AvrRps4-induced HR, RRS1/RPS4 still confers bacterial resistance in a Ws-2_nrg1a-nrg1b background (Figure 5-4, Figure 5-6). The HR is a form of cell death associated with plant resistance, and is strongly correlated with activation of a host R protein by a pathogen avirulence factor (Morel & Dangl, 1997). For instance, the recognized effectors AvrRpm1 and AvrRpt2 trigger both HR and bacterial resistance in Arabidopsis, due to recognition by the R proteins RPM1 and RPS2 (Mackey et al., 2002). However, not all avirulence factors induce HR. For example, the effectors HopZ5 and HopPsyA from P. syringae pv actinidiae and P. syringae pv syringae induce an HR-independent resistance in Arabidopsis accession Col-0 (Gassmann, 2005; Jayaraman et al., 2017). Similarly, the HR-induction by the CC-NLR Rx is not required to confer extreme resistance against potato virus X (Bendahmane et al., 1999). Here I found that NRG1 is required for HR but not resistance mediated by the TIR-NLR pair RRS1/RPS4. It was previously reported that AvrRps4 activates cell death from the cytoplasm and resistance from the nucleus, and that these two functions are independent (Heidrich et al., 2011). Whether NRG1 plays a specific role in this cytoplasmic cell death pathway is unknown. Moreover, RPP1-mediated HR in N. benthamiana fully requires NbNRG1 (Brendolise et al., 2018; Qi et al., 2018), while RPP1 function is partially retained in Ws-2_nrg1a-nrg1b (Figure 5-9). The uncoupling of HR and resistance highlights the ambiguous relationship between these two defence outputs.

In contrast, NRG1 is required for Roq1-mediated HR and Roq1-dependent bacterial resistance in N. benthamiana (Qi et al., 2018). I speculate that ADR1 proteins, from the sister clade of NRG1, might partially complement the loss of NRG1s for some TIR-NLRs, e.g. RRS1/RPS4 and RPP1 but not Roq1.
5.3.3 TIR-NLRs activate in parallel an SA-dependent pathway and an NRG1-dependent pathway that results in HR

ADR1 proteins are required for RPS2/AvrRpt2-induced SA production in Arabidopsis (Bonardi et al., 2011). I tested if the same is true for NRG1 proteins. I measured the induction of SA and the expression of the SA-biosynthetic gene ICS1 (Isochorismate Synthase 1) and the SA-responsive gene PR1 (Pathogenesis-Related1) in Ws-2_nrg1a-nrg1b (Dong, 2004). In this assay, I compared induction after infiltration of Pf0-1_AvrRps4WT (inducing both PTI and ETI via RRS1/RPS4 and RRS1B/RPS4B) and infiltration of a Pf0-1_AvrRps4KRY (inducing PTI only) (Figure 5-7). In Ws-2_rrs1-rps4-rps4b and Ws-2_eds1, ICS1 and PR1 expression are not induced by AvrRps4 or its inactive allele AvrRps4KRKY. However, AvrRps4 specifically enhances ICS1 and PR1 expression in both WT and nrg1a-nrg1b mutant lines. This is consistent with the production of SA, which is maintained at similar levels after induction in both WT and nrg1a-nrg1b. Although NRG1 is required for RRS1/RPS4 mediated HR, it is dispensable for the SA pathway activation. In contrast, ADR1 is required for SA induction during ETI (Bonardi et al., 2011). Retention of SA production upon RRS1/RPS4 activation could explain the wild-type-like bacterial resistance in Ws-2_nrg1a-nrg1b, despite the loss of HR. These results indicate that SA-dependent pathway and NRG1-dependent pathway (resulting in HR) can be activated in parallel by TIR-NLRs. Of the genes differentially regulated during XopQ-induced ETI (via the NRG1-dependent TIR-NLR Roq1), 80% depend on NRG1 (Qi et al., 2018). Some of these genes are involved in photosynthesis, RNA processing and protein degradation. Thus, it is likely that NRG1 regulates response pathways in addition to HR. In contrast, as ~20% of the ETI genes are still normally regulated in nrg1, SA-signalling is likely not the only immune response still active in nrg1 mutants.

5.3.4 Some TIR-NLRs signal via NRG1 only, some signal via both ADR1 and NRG1

RPW8-NLRs are a monophyletic anciently diverged clade. They do not show extensive diversification or expansion, and are conserved in angiosperms, gymnosperms and bryophytes (Zhong & Cheng, 2016). RPW8-NLRs diverged into ADR1 and NRG1 sub-clades before the monocot/dicot separation (Collier et al., 2011). ADR1s have been characterised in Arabidopsis as helpers for the R proteins RPP2, RPP4 and RPS2 (Bonardi et al., 2011). Additionally, an adr1-triple (adr1-adr1-l1-adr-l2) mutant can partially suppress the autoimmune phenotype of slh1-9, which is an Arabidopsis line with an auto-active allele of RRS1 (Dong et al., 2016). Thus, RPS2, RPP2, RPP4 and RRS1/RPS4 can signal via ADR1s. Here
I found that full RPP2-, RPP4- and RRS1/RPS4-mediated immune responses require NRG1s. Therefore, these three immune receptors signal via both ADR1s and NRG1s. Unlike slh1-9, the chs3-1 autoimmune phenotype is totally independent of ADR1 (Dong et al., 2016). Here I found that CHS3-induced HR requires NbNRG1. Moreover, the chs3-2D autoimmune phenotype is completely lost in an nrg1a-nrg1b-nrg1c mutant of Arabidopsis (Wu et al., 2018b). Thus, CHS3 signals only via NRG1s.

Conversely, RPS2-mediated macroscopic HR does not require NRG1 (Figure 5-5), while it requires ADR1 (Bonardi et al., 2011). Thus, RPS2 mainly signals via ADR1. It also partially requires NRG1 as indicated by a reduction of ion leakage in the absence of NRG1 (Figure 5-5).

AvrRpm1-induced macroscopic HR in Arabidopsis is still active in the adr1-triple mutant (Bonardi et al., 2011) and in the nrg1a-nrg1b mutant backgrounds (Figure 5-4). RPM1 either signals independently of ADR1 and NRG1 or can achieve full function dependent on either ADR1 or NRG1, like the CC-NLR Rx2 (Collier et al., 2011).

In conclusion, some R proteins signal via NRG1 only (e.g. CHS3) and some via both ADR1 and NRG1 (e.g. RRS1/RPS4, RPP2, RPP4, RPS2, Rx2, CHS1, SNC1) (Figure 5-13). Some others may signal via ADR1 only. A recently generated adr1-adr1-L1-adr1-L2-nrg1a-nrg1b-nrg1c sextuple mutant showed redundancy between NRG1s and ADR1s for TIR-NLR function.

The requirement of RPW8-NLRs by multiple sensor NLRs suggest that they form an NLR signalling network similar to the NRC signalling network (Wu et al., 2017, 2018a). In the NRC model, sensors and helpers evolved from a single pair. Through duplication events, the sensor clade expanded massively to enable effector recognition diversification. The helper clade underwent limited diversification (Wu et al., 2018a). Similarly, RPW8-NLR encoding genes are conserved and in low copy number in plant genomes (Shao et al., 2016, 2018). But in contrast to the NRC-sensor clade, the RPW8-NLR-dependent sensors does not fall in a defined phylogenetic clade. A more comprehensive definition of the sensor NLRs from the RPW8-NLR network will help to propose an evolutionary model for this network. In addition, RPW8-NLR and NRC co-exist in Asterids. It will be interesting to test if they form hermetic networks or if some sensor NLRs could signal via both NRC and RPW8-NLRs, adding another level of complexity, thus of robustness, in NLR signalling network.
As there are multiple copies of RPW8-NLRs in plant genomes, their characterisation was difficult with classic genetic tools. CRISPR method enabled to generate variation in RPW8-NLRs and confirm the hypothesis that they are required for multiple NLR functions (Qi et al., 2018; Wu et al., 2018b). In the future, unravelling the mechanism of RPW8 (Xiao et al., 2005; Chae et al., 2014) and RPW8-containing NLRs will push forward our understanding of the plant immune system and may ultimately be applied to deploy wisely genetic resistance in crops.

**Figure 5-13: CC-NLRs and TIR-NLRs can signal via NRG1 and/or ADR1**

NLRs are immune surveillance devices. Specific activation by pathogen effectors triggers resistance. For full response, they require RPW8-containing proteins NRG1 and ADR1. All the TIR-NLRs can signal via NRG1 and some can also signal via ADR1. CC-NLRs Signal via ADR1 and partially via NRG1. Some NLR may signal independently of ADR1 and NRG1. ADR1 and NRG1 are conserved within Angiosperms, suggesting that this model can be applied for all the flowering plants. RPM1, RPS2, RRS1/RPS4, RPP1, RPP2, RPP4, WRR4A, CSA1/CHS3 and SOC3/CHS1 signal via NRG1 (this thesis). RPS2, RPP2 and RPP4 signal via ADR1 (Bonardi et al., 2011). RPP4, RRS1 but not CHS3, signal via ADR1 (Dong et al., 2016). Roq1 signal via NRG1 (Brendolise et al., 2018). Roq1 and RPP1 signal via NRG1 (Qi et al., 2018). N signal via NRG1 (Peart et al., 2005). Rx2 signal via both NRG1 and ADR1 (Collier et al., 2011). SNC1 signals via both ADR1 and NRG1, CHS3 signals specifically via NRG1 (Wu et al., 2018b). TMV: tobacco mosaic virus. PVX: potato virus X.
Chapter 6 : General discussion

6.1 History of plant-microbe interaction genetics

Pioneering work by Gregor Mendel revealed the inheritance of genetic factors from both parents, resulting in predictable ratios for heritable phenotypes in their progenies (Mendel, 1865). These factors are encoded on the chromosomes (Morgan et al., 1915). Later, Harold Flor studied the genetic basis of the resistance to and avirulence in the flax rust fungus *Melampsora lini* (Flor, 1955, 1971). He was particularly interested in the genetics of avirulence in the pathogen. By crossing two races of rust, he determined that avirulence factors are inherited following the Mendel model. More importantly, he observed that the recognition of a single avirulence factor by flax is monogenic and proposed the gene-for-gene model. For each of the R-genes in flax, e.g. K, L6, M, N, P, there is a cognate avirulence factor in rust, e.g. AvrK, AvrL6, AvrM, AvrN, AvrP. Later, the diversity in compatibility within other pathosystems was used to identify more gene-for-gene pairs: RPS2-AvrRpt2 (Kunkel et al., 1993), RPM1-AvrRpm1 (Debener et al., 1991), RRS1-AvrRps4 (Hinsch & Staskawicz, 1995), RPS5-AvrPphB (Simonich & Innes, 1995) using the Arabidopsis-*P. syringae* pathosystem and RPP1-ATR1 (Rehmany et al., 2005), RPP2-ATR2 (Sinapidou et al., 2004), RPP4-ATR4 (Asai et al., 2018), RPP5-ATR5 (Bailey et al., 2011), RPP13-ATR13 (Allen et al., 2004) using the Arabidopsis-*Hpa* pathosystem.

The gene-for-gene terminology suggests that a single R-protein recognises a single effector. However, there are several examples of R-proteins that recognise many effectors, sometimes from highly diverged pathogens. WRR4A from Arabidopsis recognises nine avirulence factors from the oomycete *Albugo candida* (A. Redkar et al., in prep). RRS1 recognises three avirulence factors from the bacteria *Pseudomonas* and *Ralstonia* and the fungus *Colletotrichum* (Narusaka et al., 2009). Also, the same effector can be recognized by multiple different R-genes. For instance, the effector Avr2 from *P. infestans* is recognised by R2 in the Mexican *Solanum* species and by Rpi-mcq1 in Peruvian *Solanum* species (Aguilera-Galvez et al., 2018). The hunt for gene-for-gene pairs initiated by Flor has led to the discovery of many components of the plant immune system, unravelling many aspects of its mechanisms.
6.2 Limit to R-gene discovery

Any genetic study relies on phenotypic diversity. In the case of plant-microbe interactions, mapping of resistance loci requires a susceptible and a resistant parent that can produce a fertile progeny. This condition is not fulfilled for several crop-infecting races of *A. candida* in Arabidopsis: all the accessions tested are resistant to the brown mustard-infecting race Ac2V, the *Brassica rapa*-infecting race Ac7V and the cabbage-infecting race AcBoT. Thus, natural accessions cannot be used to identify the underlying R-genes.

6.3 Meeting the challenge of lack of natural diversity

6.3.1 Successful approaches

When susceptible parents are not available in the wild, susceptibility can be induced by mutagenesis. For instance, Col-0 seeds have been randomly mutagenized with EMS (ethyl methanesulfonate). A line that lacks recognition of AvrRpt2, called *rps2-201*, was identified from this mutant collection (Kunkel *et al.*, 1993). A backcross of *rps2-201* with the No-0 resistant parent enabled to map and clone *RPS2*.

Transgressive segregation in MAGIC Arabidopsis (Kover *et al.*, 2009) enabled the identification of two lines susceptible to Ac2V and Ac7V, and subsequently one line susceptible to AcBoT. They were used to map and clone novel R-genes to Ac2V, Ac7V and AcBoT (Cevik *et al.*, 2019).

6.3.2 Unsuccessful approach

While the “transgressive segregation” project (Cevik *et al.*, 2019) was ongoing in the Jones’ laboratory, several groups reported that CRISPR could be used to generate targeted mutations *in vivo* (Cong *et al.*, 2013; Mali *et al.*, 2013; Nekrasov *et al.*, 2013). At that time, no MAGIC line susceptible to AcBoT was identified. I thus tried to tackle this project by using genome editing rather than transgressive segregation. As Col-0 *eds1* is susceptible to AcBoT, I speculated that the gene(s) involved in AcBoT resistance encode(s) TIR-NLR(s). I assembled 25 sgRNAs supposedly targeting all the 96 TIR-NLR encoding genes in Col-0, with a Cas9 expression cassette. This construct did not show any CRISPR activity at the loci tested in Arabidopsis. For this reason, I investigated different parameters of the construct and by a process of hypothesis-testing combined with trial-and error, developed an optimized system for CRISPR in Arabidopsis (*Chapter 4*). In the
meantime, an AcBoT-compatible MAGIC line was identified so I stopped trying to knock-out all the TIR-NLR encoding genes in Col-0.

6.4 Beyond R-gene discovery: characterization of NLR signalling components

6.4.1 A black box between plant NLR and actual resistance

Hundreds of R-genes have been mapped and cloned in the last decades (Kourelis & Van Der Hoorn, 2018). However, little is known about the way they get activated and the signalling they trigger. In animals, it is well described that some NLRs form an oligomeric inflammasome upon activation (Lamkanfi & Dixit, 2017). The inflammasome imposes oligomerisation of the NLR N-terminal domain, recruiting and activating caspases 4, 5 and 11 upstream of caspase 1 activation, finally resulting in gasdermin D cleavage, membrane permeabilization and pyroptosis (a form of cell death). Plant NLRs can activate cell death but it is not known if they oligomerise, nor how they signal to fulfil this function. Random mutagenesis enabled identification of some general components of the plant immune system, such as EDS1, ICS1 and NDR1 (Century et al., 1995; Parker et al., 1996; Wilde et al., 2001). Conceivably, the most important regulators of immunity are encoded by redundant genes, thus unlikely to be revealed by EMS mutagenesis. This is the case for EDS1, which exists in two copies in Col-0 and many other accesses but was found in EMS screens in Ler-0 and Ws-2, that only have one copy. Similarly, the NRC genes encode for helper NLRs redundantly required for signalling of many CC-NLRs in Asterids (Wu et al., 2017).

6.4.2 RPW8-NLRs are candidates to be major ETI components

RPW8-NLRs are a characteristic sub-class of plant NLRs (Shao et al., 2016). Angiosperms carry two types of RPW8-NLRs: ADR1 and NRG1. In N. benthamiana, the single copy of NRG1 is required for the TIR-NLR sensor N function (Peart et al., 2005). As RPW8-NLRs are conserved and not expanded, they were proposed to be helper NLRs by Peter Moffett’s laboratory (Collier et al., 2011). The group tested and proved its hypothesis for the CC-NLR Rx2 (Collier et al., 2011).

In Arabidopsis, there are three copies of ADR1. They are genetically unlinked, so a triple mutant could be obtained from crosses using the single mutants. Such adr1-triple mutant
is at least partially impaired in signalling of many NLRs including RPS2, RPP4, RRS1 and UNI-1D (Bonardi et al., 2011; Dong et al., 2016).

There are two copies of NRG1 in Arabidopsis but they are in tandem in the genome, rendering it very difficult to get a double mutant from a cross between two single mutants. Although the hypothesis that NRG1s are helpers for TIR-NLRs was formulated in 2011, it was difficult to address that question before the availability of CRISPR-based targeted mutagenesis.

### 6.4.3 RPW8-NLRs are indeed major ETI regulators, as proven using CRISPR

In 2018/2019, three groups reported that they could use CRISPR to knock-out NRG1 in Arabidopsis and in N. benthamiana (Qi et al., 2018; Wu et al., 2018b) (Chapter 5). They found that, in addition to N, NRG1 is also a helper for many more TIR-NLRs: RPP1, Roq1, SNC1 and CHS3. Furthermore, I found that the TIR-NLRs RRS1/RPS4, RPP2, RPP4, WRR4A, CHS1 partially require NRG1. Although NRG1s and TIR-NLRs have co-evolved, NRG1 can also be a helper for CC-NLRs, as showed for Rx2 (Collier et al., 2011), RPS2 and RPM1 (Chapter 5).

NRG1 and ADR1 have mostly been studied independently. The only two studies inactivating NRG1 and ADR1 at the same time showed that they are redundantly required for full function of the CC-NLR Rx2 (Collier et al., 2011) and the TIR-NLR pair RRS1/RPS4 (Wu et al., 2018b). Conceivably, RPW8-NLRs are major ETI regulators, fully required in all angiosperms for function of many, if not all, other NLRs.

### 6.5 How do RPW8-NLRs regulate ETI?

#### 6.5.1 What can we learn from RPW8-NLR homologues in fungi and animals?

There is a similarity in cell death induced by the mammalian NLR NLRP3 and the fungal NLR NWD2 in fungi (Cai et al., 2014). Upon activation, NLRP3 interacts via its N-terminal domain with the protein ASC, inducing its active toxic prion form. Activation of cell death in fungus by NWD2 is similar. The active form of NWD2 recruits and converts the protein HET-S into its prion form. PNT1/HELPP is a pair homologous to NWD2/HET-S in the fungus Chaetomium globosum (NWD2/HET-S is described in the fungus Podospora anserina), which is thought to also cause cell death via prionisation (Daskalov et al., 2016).

Similar mechanisms have not been observed with plant NLRs. However, there is homology between NWD2, PNT1 and NRG1. All three are NLRs with a glycine zipper motif
at their N-termini (Daskalov et al., 2016). Whether RPW8-NLRs recruit or form a prion forming protein via their N-terminal domain to trigger cell death is not known.

6.5.2 What do we currently know about RPW8-induced immunity?

The two hypotheses proposed based on knowledge about animal and fungal NLR are (i) NLR activation induces oligomerisation of an RPW8-NLR and (ii) an N-terminal glycine zipper in RPW8-NLRs is involved in recruitment and activation of a cell death executor.

The evidence of self-association of NRG1 are contradictory. The Staskawicz laboratory observed self-interaction of NbNRG1 in the absence of activation (Qi et al., 2018). Xin Li’s laboratory could not see self-association of NRG1A or association between NRG1A and NRG1B even after activation (Wu et al., 2018b). Moreover, NRG1 does not interact with its sensor NLR N (Peart et al., 2005). Those data do not support an oligomerization model involving RPW8-NLRs.

To investigate the function of NRG1, I complemented a Ws-2_nrg1a-nrg1b mutant with NRG1A-mCherry and NRG1B-GFP (Figure 5-4). I could detect fluorescence under a confocal microscope upon activation (eight hours after infiltration of Pf0-1 carrying AvrRps4) of both NRG1A-mCherry and NRG1B-GFP. However, the resolution was too low to determine the sub-cellular localisation of the signal. Transient expression of these fluorescent-tagged Arabidopsis alleles of NRG1 in N. benthamiana suggest a cytoplasmic and endoplasmic reticulum localisation. Since these Arabidopsis alleles do not complement the loss of HR in an N. benthamiana_nrg1 mutant, I did not interpret those data any further. Consistently, Xin Li’s laboratory found that NRG1 primarily localises in the cytosol and at the ER (endoplasmic reticulum) in Arabidopsis and N. benthamiana (Wu et al., 2018b). The sub-cellular localisation and re-localisation upon activation will give clues about the NRG1 mechanism and I am planning to investigate this question.

RPW8.2, involved in powdery mildew resistance, is specifically targeted to the extra-haustorial membrane (EHM) during infection (Zhang et al., 2015). Surprisingly, a predicted trans-membrane domain (TMD) is not required for the EHM targeting (Wang et al., 2013b). It is possible that this domain is sufficient for EHM targeting but not required, if RPW8.2 forms a complex with other RPW8 or RPW8-NLR proteins having a functional TMD. Indeed, RPW8.2 N-terminal domain (that contains the predicted TMD) is conserved between other RPW8 alleles and also with NRG1 and the fungal cell death activators HET-S and HELPP (Daskalov et al., 2016). Particularly, this region contains a highly conserved
glycine zipper motif, required for plasma-membrane targeting and cell death activation for HELPP. I expressed two glycine zipper motif mutants of NRG1A (G8I and G10I) in Ws-2_nrg1a-nrg1b. These alleles did not complement the loss of AvrRps4-triggered HR, unlike the WT allele of NRG1A. However, I do not know if the mutated proteins were expressed. If those preliminary data get confirmed, it will indicate that the N-terminal predicted TMD of NRG1A is required for function.

6.5.3 Can we formulate a unifying hypothesis for RPW8-NLR function?

NLRs emerged independently as cell death activator in animals, fungi and plants. It is tempting to speculate that RPW8-NLRs regulate cell death in a similar fashion as do fungal and animal NLRs. But the current data suggest that it is not the case, and RPW8-NLRs cause cell death via a different mechanism. Previous data indicate that overexpression of the RPW8 domains of NRG1 and ADR1 causes cell death in tobacco and N. benthamiana leaves. One hypothesis is that ADR1 and NRG1 activate a cell death executor via their N-terminal RPW8 domain. The best candidate to be this cell death executor is the small protein RPW8 itself.

The relevance of RPW8-NLRs in ETI has recently been emphasized using genetics. The next step is now the investigation of their mechanism of action. Unravelling the mode of action of RPW8 and RPW8-NLRs will contribute to our knowledge in the plant immune system, for wise deployment of genetics to protect crop plants from pests.

6.6 What to expect from CRISPR?

6.6.1 A powerful genetic tool for biology

In this thesis, I exemplify the usefulness of CRISPR to interrogate a 7-year-old hypothesis. Peter Moffett’s laboratory proposed that NRG1s are helpers for many TIR-NLRs in plants. The two copies of NRG1 are in tandem in the genome of the model plant Arabidopsis. Thus, CRISPR was indeed the best technology to investigate Peter Moffett’s hypothesis by loss-of-function. CRISPR has been used to describe other gene families than NRG1, such as the yellow genes involved in pigmentation of the butterfly Agrotis ipsilon (Chen et al., 2018), or the CDR genes that confers antifungal drug resistance in Candida albicans (Vyas et al., 2015). Genome engineering is not limited to biological investigations: it can be used in biotechnology for production of biofuel, improved food or biomaterials; and in medicine to develop drugs or gene surgery (Hsu et al., 2014).
6.6.2 Deployment of knock-in breeding

Alteration of targeted genomic sequence is the most probable outcome of CRISPR-mediated DNA cleavage, from imprecise repair by NHEJ. Such alteration often affects the codon reading frame of a gene, resulting in a loss-of-function allele. Double-strand breaks induced by CRISPR can also stimulate HDR. Co-delivery of a CRISPR nuclease with an operator-specified sequence, homologous to the target site, can result in replacement of the target by the delivered sequence. This can be applied for in-vivo gene tagging for fundamental research or replacement of a gene by an allele of interest for breeders. Introgression of an allele of interest, e.g. an R-gene, into a selected crop can be achieved by classic crossing. This method is effective but slow, involving successive back-crossing and genotyping. Gene replacement mediated by CRISPR, also called knock-in breeding (KIB), can be achieved in one generation and requires only one back-cross to eliminate transgenic sequences used during the process. Conditional excision of the transgene or transgene-free delivery (e.g. bombardment of pre-assembled ribonucleoprotein complexes into plant cells) are alternative methods to circumvent a back-cross step. KIB presents several advantages compared to classic breeding. The resulting product will conserve all its genetic information, while products from classic breeding are likely to contain linked fragments of the donor genome. KIB also enables breeders to introgress genes from wild relative species that are sexually incompatible with the selected crop.

The optimization of KIB is currently ongoing. Substantial progresses have recently been reported (Čermák et al., 2015; Dahan-Meir et al., 2018; Huang & Puchta, 2019; Vu et al., 2019).

6.6.3 Deployment of base editing

The nuclease function of CRISPR nucleases can be inactivated. Delivery of an inactive protein at a target site is not valuable per se, unless implemented with a novel function. The base editors are one remarkable example of engineering an innovative re-purpose of Cas9, by targeting specific nucleotides using deaminase enzymes. The cytidine deaminase APOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 1) and a modified adenosine deaminase TadA can excise the amine group of cytosine or adenosine, converting them to uracil or inosine respectively. Uracil and inosine are processed as thymine and guanine during replication, causing targeted C>T and A>G mutation. Fusions of Cas9 with APOBEC or TadA, called base editors, have been applied
in rice to generate semi-dwarf plants, by inducing a SNP mutation in the SLR1 gene (Lu & Zhu, 2017).

CRISPR-Cas9 was successfully used to reconstruct the domestication of tomato using NHEJ-mediated repair (Li et al., 2018; Zsögön et al., 2018). Domestication of wild species can be further enhanced with base editors. Indeed, some traits relative to crop domestication are explained by single amino acid substitution, such as the naked kernel in maize caused by a Lys->Asn mutation in the transcription factor TGA1. Conceivably, base editors could be used to generate novel traits in existing crops.

6.6.4 Perspective for agriculture

Inactivation of deleterious genes, KIB and base editors are novel tools available for breeders to face the current and future agricultural challenges. Blast resistant rice, powdery mildew resistant wheat and tomato, virus resistant cucumbers and many others improved crops have been already generated using CRISPR (Langner et al., 2018; Ramasamy et al., 2018). As the final products are identical to similar products that would have been obtain with classic breeding or mutagenesis, they would be expected to be under the same regulation. In fact, the regulation of CRISPR engineered crops was discussed by policy makers of the European Union. Their conclusion was that CRISPR plants should be considered as GM (genetically modified). Thus, they are subjected to the same regulation as GM crops. It is understandable that public may perceive CRISPR plants as less “natural”. However, CRISPR-induced mutations are indistinguishable from naturally occurring mutations. Moreover, CRIPSR plants including plants resisting pathogen with reduced need for chemical sprays. This can be highly valuable for farmers and customers. Further discussions involving breeders, farmers, customers, policy makers and scientists are required to adopt smarter regulations for engineering, farming and commercialisation of CRISPR crops, so that the promise of this innovative new technology can be realized.
References


Bonardi V, Cherkis K, Nishimura MT, Dangl JL. **2012.** A new eye on NLR proteins: Focused


**Callis J, Fromm M, Walbot V. 1987.** Introns increase gene expression in cultured maize


Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH,


Cook DE, Mesarich CH, Thomma BPHJ. 2015. Understanding Plant Immunity as a


Glazebrook J, Rogers EE, Ausubel FM. 1997. Use of Arabidopsis for genetic dissection of


van der Hoorn R a L, Kamoun S. 2008. From Guard to Decoy: a new model for perception


Johal GS, Briggs SP. 1992. Reductase activity encoded by the HM1 disease resistance gene


Lechtenberg BC, Mace PD, Riedl SJ. 2014. Structural mechanisms in NLR inflammasome


Albugo candida, a generalist parasite. eLife 4: 0.7554/eLife.04550.


Morgan T, Sturtevant AH, Muller HJ, Bridges CB. 1915. The mechanisms of Mendelian heredity.


Peterson BA, Haak DC, Nishimura MT, Teixeira PJPL, James SR, Dangl JL, Nimchuk ZL.


Pidskalny RS, Rimmer SR. 1985. Virulence of Albugo Candida from turnip rape (Brassica campestris) and mustard (Brassica juncea) on various crucifers. Canadian Journal of Plant Pathology 7: 283–286.


Raitskin O, Schudoma C, West A, Patron NJ. 2018. Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: An expanded toolkit for precision genome


Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ,


Waibel F, Filipowicz W. 1990. U6 snRNA genes of Arabidopsis are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. *Nucleic acids research** **18**: 3451–3458.


