Evolutionary trade-offs at the Arabidopsis *WRR4A* resistance locus underpin alternate *Albugo candida* race recognition specificities

Baptiste Castel[1,2], Sebastian Fairhead[1,3], Oliver J. Furzer[1,4], Amey Redkar[1,5], Shanshan Wang[1], Volkan Cevik[1,6], Eric B. Holub[3], Jonathan D. G. Jones[1*]

1 The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, NR4 7UH, Norwich, United Kingdom

2 Department of Biological Sciences, National University of Singapore, Singapore 117558

3 Warwick Crop Centre, School of Life Sciences, University of Warwick, CV35 9EF, Wellesbourne, United Kingdom

4 Department of Biology, University of North Carolina, Chapel Hill, NC 27599, USA

5 Department of Genetics, University of Cordoba, 14071 Cordoba, Spain

6 The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, BA2 7AY Bath, United Kingdom

* For correspondence (email: jonathan.jones@tsl.ac.uk)

Orcid:

Baptiste Castel: 0000-0002-2722-0228

Sebastian Fairhead: 0000-0002-0716-5698

Oliver J. Furzer: 0000-0002-3536-9970

Amey Redkar: 0000-0001-5171-8061

Shanshan Wang: 0000-0002-5819-9265

Volkan Cevik: 0000-0002-3545-3179

Eric B. Holub: 0000-0003-3341-3808

Jonathan D. G. Jones: 0000-0002-4953-261X

Running head:

Allelic trade-off at a white rust resistance locus

Key-words:

immunity, resistance gene, NLR, natural variation, evolution, effector recognition, crop protection, Arabidopsis thaliana, camelina.

Corresponding author details:

Postal address: Jonathan DG Jones, The Sainsbury Laboratory, Norwich Research Park, NR4 7UH, Norwich, United Kingdom

Email address: jonathan.jones@tsl.ac.uk

Summary:

The oomycete *Albugo candida* causes white rust of Brassicaceae, including vegetable and oilseed crops, and wild relatives such as *Arabidopsis thaliana*. Novel *White Rust Resistance* (*WRR*)-genes from Arabidopsis enable new insights into plant/parasite co-evolution. *WRR4A* from Arabidopsis accession Col-0 provides resistance to many but not all white rust races, and encodes a nucleotide-binding (NB), leucine-rich repeat (LRR) (NLR) immune receptor. Col-0 *WRR4A* resistance is broken by AcEx1, an isolate of *A. candida*. We identified an allele of *WRR4A* in Arabidopsis accession Oy-0 and other accessions that confers full resistance to AcEx1. $WRR4A^{Oy\text{-}0}$ carries a C-terminal extension required for recognition of AcEx1, but reduces recognition of several effectors recognized by the $WRR4A^{Col\text{-}0}$ allele. $WRR4A^{Oy\text{-}0}$ confers full resistance to AcEx1 when expressed in the oilseed crop *Camelina sativa*.

Introduction:

Plants have evolved powerful defence mechanisms that can arrest attempted colonization by microbial pathogens. Timely defence activation requires perception of pathogen-derived molecules by cell-surface pattern-recognition receptors (PRRs) and intracellular Nucleotide-binding (NB), Leucine-rich repeat (LRR), or NLR, immune receptors (Jones and Dangl, 2006). Extensive NLR genetic diversity within plant populations is associated with robustness of NLR-mediated immunity (Baggs *et al.*, 2017), and plant NLR sequences reveal diversifying selection on NLR genes compared to other genes (Meyers *et al.*, 1998; Kuang *et al.*, 2004;

Monteiro and Nishimura, 2018). To investigate NLR diversity, next-generation sequencing technologies were combined with sequence capture to develop *Resistance (R)*-gene enrichment sequencing (RenSeq) (Jupe *et al.*, 2013). This method has shed new light on NLR repertoires in several plant genomes including tomato, potato and wheat (Andolfo *et al.*, 2014; Steuernagel *et al.*, 2016; Witek *et al.*, 2016). A comparison of 64 *Arabidopsis thaliana* (Arabidopsis) accessions using RenSeq documented NLR sequence diversity within a single species, revealing the Arabidopsis "pan-NLRome" (Van de Weyer *et al.*, 2019). Each Arabidopsis accession contains 150-200 NLR-encoding genes. About 60% are found in clusters (within 200 kb from each other) that show copy number variation (Lee and Chae, 2020). From all the NLRs of the 64 accessions, 10% are singletons and the rest are distributed between 464 orthogroups. Each accession contains a unique subset comprising, on average, 25% of the orthogroups.

NLRs vary in their intramolecular architecture. Plant NLR proteins usually display either a “Toll, Interleukin-1, *R*-gene” (TIR), a “Coiled-Coil” (CC) or “Resistance to Powdery mildew 8” (RPW8) N-terminal domain, a central NB domain and a C-terminal LRR domain. Some NLRs also comprise additional C-terminal domains. For example, RRS1 is an Arabidopsis TIR-NLR with a WRKY domain required to detect the effectors AvrRps4 (from the bacterium *Pseudomonas syringae*) and PopP2 (from the bacterium *Ralstonia solanacearum*). The integrated WRKY is called a decoy as it mimics the authentic AvrRps4 and PopP2 effector targets (Le Roux *et al.*, 2015; Sarris *et al.*, 2015). Several other integrated decoy domains have been described (Cesari, 2017). Analysis of NLR integrated domains can potentially reveal novel effector targets (Kroj *et al.*, 2016).

RPP1 and Roq1, two TIR-NLRs from Arabidopsis and *Nicotiana benthamiana* respectively, form tetrameric resistosomes upon activation (Ma *et al.*, 2020; Martin *et al.*, 2020). In this structure, a C-terminal jelly-roll/Ig-like domain (C-JID) physically binds the cognate effector, along with the LRR domain. The C-JID corresponds to previously described motifs found after the LRR of many TIR-NLRs, called post-LRR motifs (Van Ghelder and Esmenjaud, 2016; Saucet *et al.*, 2021). We will refer this domain as C-JID in the rest of the text.

A. candida causes white blister rust in Brassicaceae and serious annual yield losses in brassica crops such as oilseed mustard (*Brassica juncea*) in India (Gupta *et al.*, 2018). It comprises several host-specific groups, which include race 2 from *B. juncea*, race 7 from *B. rapa*, race 9 from *B. oleracea* and race 4 from wild relatives (*e.g., Capsella bursa-pastoris, Arabidopsis spp.* and *Camelina sativa*) (**Table S1**) (Jouet *et al.*, 2018; Pound and Williams, 1963). They have been proposed to evolve by rare recombination events that occurred between the races, followed by clonal propagation on susceptible hosts (McMullan *et al.*, 2015). The Arabidopsis Columbia (Col-0) allele of *WRR4A* can confer resistance to isolates

of all four races (Borhan *et al.*, 2010; Borhan *et al.*, 2008). The allele encodes a canonical TIR-NLR and belongs to an orthogroup of three genes in Col-0 at the same locus. The accession Ws-2 (susceptible to *A. candida* race 4) lacks *WRR4A* but contains the two other paralogs, illustrating intra-species copy number variation within clusters. Interestingly, one of these paralogs, *WRR4B*, also confers resistance to the Ac2V isolate of race 2 (Cevik *et al.*, 2019). In addition, the CC-NLR-encoding *BjuWRR1*, which confers resistance to several *A. candida* isolates collected on *B. juncea,* was mapped and cloned from the European accession of *B. juncea* Donskaja-IV (Arora *et al.*, 2019).

Several Col-0-virulent isolates of *A. candida* race 4 have been collected from naturally infected Arabidopsis plants. They were used to identify an alternative source of broad-spectrum white rust resistance. One of these isolates, AcEx1, was used to reveal a source of resistance in Oy-0 that mapped to the *WRR4* locus (Fairhead, 2016; Castel, 2019). We set out to clone the gene conferring AcEx1 resistance in Oy-0, and characterise the corresponding pathogen effector(s).

AcEx1 was collected from *Arabidopsis halleri* in Exeter, UK. It is also virulent in *Camelina sativa,* an emerging oilseed crop which has been engineered to provide an alternative source of long chain omega-3 polyunsaturated fatty acids (LC-PUFAs) (Ruiz-Lopez *et al.*, 2014; Petrie *et al.*, 2014). Transgenic camelina oil is equivalent to fish oil for salmon feeding and for human health benefits (Betancor *et al.*, 2018; West *et al.*, 2019). Despite challenges to distribute a product derived from a genetically modified crop (Napier *et al.*, 2019), an increase in camelina cultivation can be expected in the near future. Fields of *C. sativa* will inevitably be exposed to *A. candida* and early identification of *R*-genes will enable crop protection. Furthermore, AcEx1 can suppress Arabidopsis non-host resistance to the potato late blight pathogen *Phytophthora infestans* (Belhaj *et al.*, 2017; Prince *et al.*, 2017), and also to downy mildews (Cooper *et al.*, 2008) emphasizing the importance of protecting camelina fields from white rust.

In this study we identified two alleles of *WRR4A* conferring full resistance to AcEx1 from Arabidopsis accessions Oy-0 and HR-5. They both encode proteins with a C-terminal extension compared to the Col-0 *WRR4A* allele. This extension enables recognition of at least one effector from AcEx1. We propose that *WRR4A*$^{Oy\text{-}0}$ is the ancestral state, and that in the absence of AcEx1 selective pressure, an early stop codon in *WRR4A* generated the Col-0-like allele, enabling more robust recognition of other *A. candida* races while losing recognition of AcEx1. Finally, we successfully transferred *WRR4A*$^{Oy\text{-}0}$-mediated resistance to AcEx1 from Oy-0 into *Camelina sativa*.

Results:

Resistance to AcEx1 is explained by **WRR4A** *alleles of HR-5 and Oy-0*

AcEx1 growth on Col-0 results in chlorosis that is not seen in the fully susceptible accession Ws-2 (**Figure 1a**). Since *WRR4A* confers resistance to all other *A. candida* races tested and Ws-2 lacks *WRR4A*, we tested if the chlorotic response could be explained by *WRR4A*, by testing a Col-0_*wrr4a-6* mutant, and found that it shows green susceptibility to AcEx1. We also tested Ws-2 transgenic lines carrying *WRR4A* from Col-0 and observed chlorotic susceptibility (**Figure 1b**). Thus, *WRR4A* from Col-0 weakly recognises AcEx1 and provides partial resistance. However, AcEx1 is still able to complete its life cycle on Col-0, which is therefore considered susceptible.

In a search for more robust sources of AcEx1 resistance, we tested 283 Arabidopsis accessions (**Table S2**). We identified 57 (20.1%) fully resistant lines, including Oy-0 and HR-5. We phenotyped 278 Recombinant Inbred Lines (RILs) between Oy-0 (resistant) and Col-0 (susceptible) and conducted a quantitative trait locus (QTL) analysis that revealed one major QTL on chromosome 1 and two minor QTLs on chromosomes 3 and 5 (**Figure S1a**). All loci contribute to resistance, with a predominant contribution of the QTL on chromosome 1 (see Figure 3.7 of Fairhead, 2016). We did not investigate the minor QTL on chromosome 5. Fine mapping on chromosome 1 and 3 QTLs refined the QTL boundaries (**Figure S2 and S3**, see Experimental Procedures). Based on sequence identity between the QTL in Col-0 and in an Oy-0 RenSeq dataset (Van de Weyer *et al.*, 2019), we identified four NLRs associated with the QTLs in Oy-0: three TIR-NLR paralogs on chromosome 1 (*WRR4A*, *WRR4B* and one absent in Col-0 that we called *WRR4D*) and a CC-NLR absent in Col-0 on chromosome 3 (that we called *Candidate to be WRR11*, *CWR11*) (**Figure S1bc**).

We expressed these genes, with their own promoters and terminators, in the fully susceptible accession Ws-2. Only $WRR4A^{Oy\text{-}0}$ conferred full resistance (**Figure 1b**). *CWR11*, the only NLR from the *WRR11* locus, does not confer AcEx1 resistance. The gene underlying *WRR11* locus resistance remains unknown.

We conducted a bulk segregant analysis using an F2 population between HR-5 (resistant) and Ws-2 (susceptible). RenSeq on bulked F2 susceptible segregants revealed a single locus on chromosome 1, that maps to the same position as the chromosome 1 QTL in Oy-0 (**Figure S4a**). Since $WRR4A^{Oy\text{-}0}$ confers resistance to AcEx1, we expressed its HR-5 ortholog, in genomic context, in the fully susceptible accession Ws-2, and found that $WRR4A^{HR\text{-}5}$ also confers full resistance to AcEx1 (**Figure 1b**).

In conclusion, *WRR4A* from Col-0 can weakly recognise AcEx1 but does not provide full resistance. We identified two *WRR4A* alleles, in Oy-0 and HR-5, that confer full AcEx1 resistance.

WRR4A^{Col-0} carries an early stop codon compared to WRR4A^{Oy-0}

To understand why the Oy-0 and HR-5 alleles of *WRR4A* confer full resistance to AcEx1, while the Col-0 allele does not, we compared the gene and protein sequences (**Figure 2**). First, we defined the cDNA sequence of *WRR4A*$^{Oy-0}$. The splicing sites are identical between the two alleles. There are 46 polymorphic amino acids between Col-0, HR-5 and Oy-0. Col-0 shares 96.03% amino acid sequence identity with Oy-0 and 96.23% with HR-5, while Oy-0 and HR-5 share 97.15% amino acid sequence identity. *WRR4A*$^{Col-0}$ carries a 156-nucleotide insertion in the first intron compared to Oy-0 and HR-5. A more striking polymorphism is a TGC->TGA mutation in *WRR4A*$^{Col-0}$, resulting in an early stop codon compared to *WRR4A*$^{Oy-0}$ and *WRR4A*$^{HR-5}$ (**Figure 2**), located 178 amino acids after the C-JID, resulting in an 89 amino acid extension in WRR4A^{Oy-0} and WRR4A^{HR-5}. The nucleotide sequence for this extension is almost identical between HR-5, Oy-0 and Col-0 (two polymorphic sites). Thus, by mutating TGA to TGC in Col-0, we could engineer an allele with the extension, that we called *WRR4A*$^{Col-0_LONG}$ (**Figure 3a**). By mutating TGC to TGA in Oy-0, we could engineer an Oy-0 allele without the extension, that we called *WRR4A*$^{Oy-0\_SHORT}$. We expressed these alleles, as well as the WT Col-0 and Oy-0 alleles, with their genomic context, in the AcEx1-compatible accession Ws-2. For unknown reasons, none of the *WRR4A*$^{Col-0\_LONG}$ and *WRR4A*$^{Oy-0\_SHORT}$ transgenic seeds germinated. We tried to generate Arabidopsis Col-0 lines with *WRR4A*$^{Col-0-STOP}$ using CRISPR adenine base editor (see Experimental Procedures). Out of 24 transformed plants, none displayed editing activity at all. Thus, we did not generate stable *WRR4A* stop codon mutants in Arabidopsis. We therefore cloned these alleles under the control of the *35S* promoter and the *Ocs* terminator for transient overexpression in *N. tabacum* (**Figure 3**).

Since many TIR-NLRs carry a C-JID, we conducted a Hidden Markov Model (HMM) search and found one in WRR4A (www.ebi.ac.uk/Tools/hmmer/search/hmmsearch on *Arabidopsis thaliana* using HMM previously reported (Ma *et al.*, 2020), e-value = 5.7e-14). This C-JID is present in Oy-0, HR-5 and Col-0 alleles (**Figure 2b**). The C-terminal extension in WRR4A^{Oy-0} relative to WRR4A^{Col-0} does not show homology with known protein domains.

Extension in WRR4A *confers specific recognition of AcEx1 candidate effectors*

In order to identify AcEx1 effectors specifically recognised by WRR4A^{Oy-0}, we tested for a hypersensitive response (HR), a typical phenotype upon NLR activation, after transient expression of WRR4A^{Oy-0} along with AcEx1 candidate effectors in *N. tabacum* leaves. Secreted CxxCxxxxxG (CCG) proteins are expanded in the genomes of *Albugo* species and

198 are effector candidates (Kemen *et al.*, 2011, Furzer *et al.*, 2021). We identified 55 CCGs in
199 the AcEx1 genome (Jouet *et al.*, 2018, Redkar *et al.*, 2021), and PCR-amplified and cloned
200 21 of them, prioritizing those that showed allelic variation with other races. From them, CCG39
201 induces a WRR4A$^{Oy\text{-}0}$-dependent HR (**Figure 3**) and explains AcEx1 resistance in Oy-0.
202 WRR4A$^{Col\text{-}0_LONG}$ can also recognise CCG39, but WRR4A$^{Oy\text{-}0_SHORT}$ cannot. Hence, the C-
203 terminal extension fully explains the acquisition of recognition of CCG39. In addition,
204 WRR4A$^{Col\text{-}0_LONG}$ recognises CCG35 (**Figure 3b**). Recognition of CCG35 is not explained
205 solely by the C-terminal extension (as WRR4A$^{Oy\text{-}0}$ does not recognise it) or by the core region
206 of the Col-0 allele (as WRR4A$^{Col\text{-}0}$ does not recognise it).

207 WRR4A$^{Col\text{-}0}$ can recognise eight CCG effectors from other races of *A. candida* (Redkar *et al.*,
208 2021). We found that WRR4A$^{Oy\text{-}0}$ is able to recognise CCG28, CCG40 and CCG104, but not
209 CCG30, CCG33, CCG67, CCG71 and CCG79 (**Figure 3c**). WRR4A$^{Col\text{-}0_LONG}$ recognises all
210 the CCGs indistinguishably from WRR4A$^{Col\text{-}0}$, indicating no influence of the C-terminal
211 extension on their recognition.

212 In conclusion, we identified one AcEx1 effector specifically recognised by WRR4A$^{Oy\text{-}0}$. The C-
213 terminal extension is required and sufficient for its recognition. We also found that WRR4A$^{Oy\text{-}0}$
214 does not recognise several of the Col-0-recognised CCG from other races.

215

216 **WRR4A *alleles carrying a C-terminal extension are associated with AcEx1 resistance***

217 The NLR repertoire of 64 Arabidopsis accessions has been determined using resistance gene
218 enrichment Sequencing (RenSeq) (Van de Weyer *et al.*, 2019). We found 20 susceptible and
219 5 resistant genotypes that belong to the 64 accessions (**Table S2**). We retrieved *WRR4A* from
220 these 25 accessions (http://ann-nblrrome.tuebingen.mpg.de/apollo/jbrowse/). The read
221 coverage was insufficient to resolve *WRR4A* sequence in Bur-0 (susceptible) and Mt-0
222 (resistant). *WRR4A* is absent from the *WRR4* cluster in Ws-2, Edi-0 and No-0. Consistently,
223 these accessions are fully susceptible to AcEx1. From the DNA sequence of the 20 other
224 accessions, we predicted the protein sequence, assuming that the splicing sites correspond
225 to those in Col-0 and Oy-0 (**Figure 4 and Dataset 1**). There are two well-defined groups of
226 *WRR4A* alleles. One includes WRR4A$^{Col\text{-}0}$; the other includes WRR4A$^{Oy\text{-}0}$. The Col-0-like and
227 Oy-0-like groups are also discriminated in a phylogeny constructed based on predicted protein
228 sequences (**Figure S5 and Dataset S2**). All alleles from the Col-0 group carry TGA (apart
229 from ULL2-5, TGC, but *WRR4A* is pseudogenised in this accession), while all alleles from the
230 Oy-0 group carry TGC, at the Col-0 stop codon position. Several alleles from both groups,
231 including Bay-0, ULL2-5, Wil-2, Ler-0, Ws-0 and Yo-0, carry an early stop codon (*i.e.* upstream
232 of the Col-0 stop codon position), so the resulting proteins are likely not functional.

Consistently, all the accessions from the Col-0 group and all the accessions carrying an early stop codon are susceptible to AcEx1. The only exception is Kn-0, that carries an Oy-0-like allele of *WRR4A* but is susceptible to AcEx1. Otherwise, the presence of an Oy-0-like C-terminal extension associates with resistance.

AcEx1 resistance can be transferred from Arabidopsis to Camelina

AcEx1 can grow on *Camelina sativ*a **(Figure 5)**, which like Arabidopsis, can be transformed using the floral dip method (Liu *et al.*, 2012). We generated a *WRR4A*$^{Oy\text{-}0}$-transgenic camelina line. We obtained four independent transformants, including two with a single-locus T-DNA insertion (**Figure S6**). From these lines we obtained five and four lines showing no symptoms upon AcEx1 inoculation, from which we obtained one bag of homozygous seeds (the others giving either no seeds or segregating seeds). We tested the single homozygous resistant line obtained for stable resistance to AcEx1 (**Figure 5**). Out of twelve individuals, eight showed resistance without symptoms, three showed resistance with a chlorotic response (likely WRR4A-mediated HR) and one showed susceptibility (white pustule formation caused by sporulation of *A. candida*). All twelve WT camelina control plants showed mild to severe white rust symptoms. This indicates that WRR4A$^{Oy\text{-}0}$ can confer resistance to AcEx1 in *C. sativa*.

Discussion:

Col-0 and HR-5 *WRR4A* alleles recognise effectors from AcEx1

A screen for novel sources of resistance to AcEx1 identified accessions HR-5 and Oy-0 as worthy of further investigation. Positional cloning from Oy-0 and then allele mining in HR-5 showed that this immunity is mediated by alleles of *WRR4A* in HR-5 and Oy-0 with distinct recognition capacities compared to the Col-0 allele. In Oy-0, two additional dominant loci, *WRR11* on chromosome 3 and *WRR15* on chromosome 5 contribute resistance to AcEx1 but the molecular basis of these resistances was not defined. Further investigation on *WRR11* was conducted but did not reveal the causal gene (Castel, 2019, Chapter 3).

WRR4A$^{Oy\text{-}0}$ recognizes at least one AcEx1 effector that is not recognized by WRR4A$^{Col\text{-}0}$ (**Figure 3**). Conceivably, *WRR4A*$^{Oy\text{-}0}$ could be combined with *WRR4A*$^{Col\text{-}0}$ and *WRR4B*$^{Col\text{-}0}$ to expand the effector recognition spectrum of a stack of *WRR* genes that could be deployed in *B. juncea* or *C. sativa* (Pedersen, 1988).

WRR4A alleles fall into two groups that can or cannot confer AcEx1 resistance

Analysis of *WRR4A* allele diversity in Arabidopsis revealed WRR4A$^{Oy\text{-}0}$-like and WRR4A$^{Col\text{-}0}$-like alleles. Since WRR4A$^{Col\text{-}0}$- like alleles show near-identity to WRR4A$^{Oy\text{-}0}$- like alleles in nucleotide sequence after the premature stop codon, the latter are likely to be ancestral, and the WRR4A$^{Col\text{-}0}$- like early stop codon occurred once, in the most recent common ancestor of Sf-2 and Col-0. Other early stop codons, resulting in loss-of-function proteins, occurred randomly in both Oy-0- and Col-0-containing groups. About a third of the investigated accessions contain another early stop codon resulting in a likely non-functional allele (**Figure 4**). The full-length Oy-0-like alleles are associated with resistance to AcEx1, while the Col-0-like alleles are associated with susceptibility (**Figure 4**). The only exception is Kn-0, which displays a full length Oy-0-like allele but is susceptible to AcEx1. Susceptibility in Kn-0 could be explained by SNPs, lack of expression or mis-splicing of *WRR4A$^{Kn\text{-}0}$*.

The Col-0 allele C-terminal truncation correlates with gain of recognition for some CCGs and loss of recognition for others, suggesting an evolutionary trade-off

A. candida isolates identical or almost identical to AcEx1 are broadly distributed, at least across Europe (Jouet *et al.*, 2018). Similarly, the WRR4A$^{Col\text{-}0}$ allele is not associated with a geographic location, indicating that it is maintained by a non-climatic factor (**Figure S7**). We propose that, in the absence of AcEx1 selection pressure, the Col-0-like early stop codon occurred to provide a new function, along with the loss of AcEx1 effector recognition. This new function enables recognition of additional CCGs from other *A. candida* races.

By combining the C-terminal extension on WRR4A$^{Oy\text{-}0}$ with the core region of WRR4A in Col-0 (**Figure 3e**), recognition of additional AcEx1 CCGs was enabled. Furthermore, Arabidopsis natural accessions carrying the core region of the Col-0-like allele also lack the C-terminal extension (**Figure 4 and S5, Dataset S1 and S2**). This could be an example of intramolecular genetic suppression (Kondrashov *et al.*, 2002; Schülein *et al.*, 2001; Brasseur *et al.*, 2001; Davis *et al.*, 1999). The combination between the core region of the Col-0 allele with the C-terminal extension may form a hyper-active WRR4A allele with excessive fitness cost for the plant, which may explain why no transgenic Arabidopsis could be recovered that carry WRR4A$^{Col\text{-}0_LONG}$. The early stop codon may have occurred in Col-0 to compensate for hyper-activation of an ancestral WRR4A allele. Hyper-activation of the immune system is deleterious, as shown for example by hybrid incompatibility caused by immune receptors (Wan *et al.*, 2021).

Many TIR-NLRs contain conserved post-LRR motifs (Meyers *et al.*, 2002; Van Ghelder and Esmenjaud, 2016), that cover a functional C-JID motif involved in effector binding (Ma *et al.*, 2020, Martin *et al.*, 2020). We found that WRR4A also contains this domain. Both WRR4A^{Col-0}- and WRR4A^{Oy-0}- carry the C-JID, so it does not explain the unique CCG recognition of each allele. Instead, the polymorphism that explains AcEx1 recognition is a short sequence, particularly enriched in negatively charged residues (Glu and Asp, **Figure 2**), located after the C-JID. Polymorphism within the C-JID between RPP1 and Roq1 contributes to effector recognition specificity (Ma *et al.*, 2020; Martin *et al.*, 2020). In the case of WRR4A, it seems that polymorphism after the C-JID also contributes to specific effector recognition. Biochemical studies of WRR4A should provide more insights into the mechanism of CGG recognition.

Arabidopsis *WRR4A* resistance to AcEx1 can be transferred to the crop camelina

Camelina sativa was recently engineered to produce LC-PUFAs, an essential component in the feed used in fish farming (Petrie *et al.*, 2014). Currently, fish farming uses wild fish-derived fish oil. Fish oil-producing camelina offers a solution to reduce the need for wild fish harvesting, potentially reducing pressure on world marine fish stocks (Betancor *et al.*, 2018). There are challenges in delivering products derived from transgenic crops but fish oil-producing crops could reduce the environmental impact of fish farming. White rust causes moderate symptoms on camelina. Moreover, *A. candida* is capable of immunosuppression (Cooper *et al.*, 2008). *A. candida*-infected fields constitute a risk for secondary infection of otherwise incompatible pathogens. To safeguard camelina fields against white rust, both chemical and genetic solutions are possible. Genetic resistance offers the advantage of a lower cost for farmers and reduces the need for fungicide release in the environment. Since the first report of white rust on camelina in France in 1945, no genetic resistance has been characterised. All the strains collected on camelina can grow equally irrespective of camelina cultivar (Séguin-Swartz *et al.*, 2009). This absence of phenotypic diversity precludes discovery of resistance loci using classic genetic tools. We found that WRR4A^{Oy-0} confers resistance to AcEx1 in camelina (**Figure 5**). Arabidopsis WRR4A resistance is functional in *B. juncea* and *B. oleracea* (Cevik *et al.*, 2019), suggesting that the mechanism of activation and the downstream signalling of WRR4A is conserved, at least in Brassicaceae.

In conclusion, we found a novel example of post-LRR polymorphism within an NLR family, associated with diversified effector recognition spectra. By investigating the diversity of WRRA, we identified an allele that confers white rust resistance in the camelina crop.

Experimental procedures:

Plant material and growth conditions

Arabidopsis thaliana (Arabidopsis) accessions used in this study are Øystese-0 (Oy-0, NASC: N1436), HR-5 (NASC: N76514), Wassilewskija-2 (Ws-2, NASC: N1601) and Columbia (Col-0, NASC: N1092). Col-0_*wrr4a-6* mutant is published (Borhan *et al.*, 2008). 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, 5-weeks old plants were transferred under long-day condition: 21°C, with 16 hours of light and 8 hours of dark, 75% humidity. For *Nicotiana tabacum* (cultivar Petit Gerard) and *Camelina sativa* (cultivar Celine), seeds were sown directly on compost and plants were grown at 21°C, with cycles of 16 hours of light and 8 hours of dark, at 55% humidity.

Albugo candida infection assay

For propagation of *Albugo candida*, zoospores from the infected leaf inoculum 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 to promote spore germination. Infected plants were kept under 10-hour light (20 ˚C) and 14-hour dark (16˚C) cycles. Phenotypes were monitored 14 to 21 days after inoculation.

QTL analysis

QTL mapping of the bipartite F8 Oy-0 x Col-0 population (470 Recombinant Inbreed Lines, RILs, publiclines.versailles.inra.fr/page/27) (Simon *et al.*, 2008) was performed on a genetic map of 85 markers across the five linkage groups that accompanied the population using RQTL (Broman *et al.*, 2003). Standard interval mapping using a maximum likelihood estimation under a mixture model (Lander and Botstein, 1989) was applied for interval mapping. Analysis revealed two major QTLs: on chromosome 1 and on chromosome 3.

Chromosome 1 QTL is located between 20,384 Mb and 22,181 Mb (**Figure S2**). It includes the TIR-NLR cluster *WRR4* and the CC-NLR cluster *RPP7*. Six RILs (three resistant and three susceptible) recombine within the QTL and were used for fine mapping. We designed a Single Nucleotide Polymorphism (SNP, 21,195 Mb, Fw: TCAGATTGTAACTGATCTCGAAGG, Rv: CCATCAAGCACACTGTATTCC, amplicon contains two SNPs, Oy-0: A and G, Col-0: G and C) and an Amplified Fragment Length Polymorphism (AFLP, 21,691 Mb, Fw:

AAGGCAATCAGATTAAGCAGAA, Rv: GCGGGTTTCCTCAGTTGAAG, Oy-0: 389 bp, Col-0: 399 bp) markers between *WRR4* and *RPP7*. Four lines eliminate *RPP7* from the QTL. The only NLR cluster in chromosome 1 QTL is *WRR4*.

Chromosome 3 QTL is located between 17,283 Mb and 19,628 Mb (**Figure S3**). It includes the atypical *resistance*-gene cluster *RPW8*, the CC-NLR *ZAR1* and the paired TIR-NLRs *At3g51560-At3g51570*. Six RILs (three resistant and three susceptible) recombine within the QTL and were used for fine mapping. We designed an AFLP (18,016 Mb, Fw: gctacgccactgcatttagc, Rv: CCAATTCCGCAACAGCTTTA, Oy-0: 950 bp, Col-0: 1677 bp) and a Cleaved Amplified Polymorphic Sequence (CAPS, 18,535 Mb, Fw: TCAAGCCTGTTAAGAAGAAGAAGG, Rv: GCCCTCCACAAAGATTCTGAAGTA, enzyme: DdeI, Oy-0: uncleaved, Col-0: cleaved) markers between the QTL border and *RPW8*. We designed a CAPS marker (18,850 Mb, Fw: TCTCGGGGAAAATATGATTAGA, Rv: GGTTGATTTTTATTGTGGTAGTCGT, enzyme: SwaI, Oy-0: cleaved, Col-0: uncleaved) between *RPW8* and *ZAR1*. We designed a SNP (18,937 Mb, Fw: CCACAAGGTCGGAATCTGTAGC, Rv: TGCACAGAAGTAACCCACCAAC, Oy-0: C, Col-0: T) and a CAPS (19,122 Mb, Fw: ACCACCACCTCGATGCATTTC, Rv: CCTTCCCTGCGAAAGACACTC, enzyme: BsrI, Oy-0: uncleaved, Col-0: cleaved) markers between *ZAR1* and the TIR-NLR pair. Three recombinants eliminate the TIR-NLR pair, two eliminate *ZAR1* and one eliminates *RPW8*. None of the Col-0 NLR clusters orthologs are present in the QTL. The gene underlying chromosome three resistance is located between the border of the QTL and RPW8.

Bulk segregant analysis and RenSeq

We generated an F2 population from a cross between HR-5 (resistant) and Ws-2 (susceptible). We phenol/chloroform extracted DNA from 200 bulked F2 lines fully susceptible to AcEx1. The bulked DNA sample was prepared as an Illumina library and enriched using the Arabidopsis v1 RenSeq bait library (Arbor Bioscience, MI, USA) (**Table S3**), as described by (Jupe et al., 2013). The sample was sequenced in a pooled MiSeq run (data available on request). Firstly, reads were aligned with BWA mem (Li and Durbin, 2009) to the Col-0 reference genome and SNPs called with Samtools (Li *et al.*, 2009). The genome was scanned for regions of high linkage with the next generation mapping tools at http://bar.utoronto.ca/ngm/ (Austin *et al.*, 2011). Secondly, the reads were mapped using BWA to the RenSeq PacBio assembly generated for HR-5 (Van de Weyer *et al.*, 2019). Highly linked regions were confirmed visually with the integrated genome viewer (Robinson *et al.*, 2017).

Gene cloning

Vectors were cloned with the USER method (NEB) following the manufacturer recommendations. For expression of resistance gene candidates in Arabidopsis, genes were cloned with their natural 5' and 3' regulatory sequences into LBJJ233-OD (containing a FAST-Red selectable marker, pre-linearized with PacI and Nt. BbvcI restriction enzymes). For overexpression in *Nicotiana tabacum*, genes were cloned into LBJJ234-OD (containing a FAST-Red selectable marker and a 35S / Ocs expression cassette, pre-linearized with PacI and Nt. BbvcI restriction enzymes). Primers, template and vectors are indicated in (**Table S4)**. *WRR4A*$^{Col\text{-}0}$ is published (Cevik *et al.*, 2019). CCGs recognised by WRR4A$^{Col\text{-}0}$ are published (Redkar *et al.*, 2021).

All the plasmids were prepared using a QIAPREP SPIN MINIPREP KIT on *Escherichia coli* DH10B thermo-competent cells selected with appropriate antibiotics. Positive clones (confirmed by size selection on electrophoresis gel and capillary sequencing) were transformed in *Arabidopsis thaliana* via *Agrobacterium tumefaciens* strain GV3101. Transgenic seeds were selected under fluorescent microscope for expression of the FAST-Red selectable marker (Shimada *et al.*, 2010).

CRISPR adenine base editor

An sgRNA targeting *WRR4A* stop codon in Col-0 (TTCTGAgaagcattcgaaag[nGA]) was assembled by PCR to a sgRNA backbone and 67 bp of the *U6-26* terminator. It was then assembled with the *AtU6-26* promoter in the Golden Gate compatible level 1 *pICH47751*. We designed a mutant allele of a plant codon optimized Cas9 with a potato intron (Addgene: 117515) with D10A (nickase mutant) and R1335V/L1111R/D1135V/G1218R/ E1219F/A1322R/T1337R, to change the PAM recognition from NGG to NG (Nishimasu *et al.*, 2018). We assembled this Cas9 (golden gate compatible BpiI: GACA-GCTT) along with a barley codon optimized TadA module (golden gate compatible BpiI: AATG-GCTT) in a level 0 vector *pICH41308*. It was then assembled with the YAO promoter (Addgene: 117513) and the E9 terminator (Addgene: 117519) in a level 1 vector *pICH47811* (with expression in reverse orientation compared to the other level 1 modules). It was then assembled with a FAST-Red selectable marker (Addgene: 117499) and the sgRNA level 1 cassette into a level 2 vector *pICSL4723*, using the end-linker *pICH41766*. Level 0 vector was cloned using BpiI enzyme and spectinomycin resistance. Level 1 vectors were cloned using BsaI enzyme and carbenicillin resistance. Level 2 vector was cloned using BpiI enzyme and kanamycin resistance. It was expressed via *Agrobacterium tumefaciens* strain GV3101 in Arabidopsis Oy-0. In the first generation after transformation, we did not detect any mutant from 24

independent transformants. It indicates an absence of activity of the construct. It can be explained by the Cas9 mutations that were not tested before on this specific allele nor in combination with TadA.

Transient expression in* N. tabacum *leaves

A. tumefaciens strains were streaked on selective media and incubated at 28 °C for 24 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 infiltration buffer (10 mM $MgCl_2$, 10 mM MES, pH 5.6) at OD_{600} = 0.4 (2 x10^8 cfu/ml). For co-expression, each bacterial suspension was adjusted to OD_{600} = 0.4 for infiltration. The abaxial surface of 4-weeks old *N. tabacum* were infiltrated with 1 ml needle-less syringe. Cell death was monitored three days after infiltration.

Resolution of* WRR4A$^{Oy\text{-}0}$ *cDNA sequence

RNA was extracted from Oy-0 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). PCR was conducted using Fw: TCTGATGTCCGCAACCAAAC (in the first exon) and Rv: GTCCTCTTCGGCCATATCTTC (in the last exon) with the Taq Polymerase enzyme (NEB) following the manufacturer protocol. The 2848 nt amplicon sequence, corresponding to the cDNA sequence (*i.e.* with already spliced introns) was resolved by capillary sequencing. It indicates that the splicing sites are identical between *WRR4A*$^{Oy\text{-}0}$ and the splicing sites reported in the database TAIR10 for *WRR4A*$^{Col\text{-}0}$.

Gene expression by RT-PCR

RNA was extracted from leaf tissue 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). PCR was conducted using primers indicated in **Table S4** with the Taq Polymerase enzyme (NEB) following the manufacturer protocol.

Protein extraction and western blot

Proteins were extracted from leaf tissue using TruPAGE LDS Sample Buffer (Sigma-Aldrich) following the manufacturer's recommendations. They were separated by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and analysed by immunoblotting. After electrophoresis, separated proteins were transferred to Immunobilon-P PVDF (Merck Millipore) membranes for immunoblotting. Membranes were blocked for 2 h in 5% nonfat milk, probed with horseradish peroxidase (HRP)-conjugated antibodies overnight and imaged.

Generation of WRR4A$^{Oy\text{-}0}$ *transgenic* Camelina sativa *line*

We transformed a *WRR4A*$^{Oy\text{-}0}$ construct, under native promoter and terminator transcriptional regulation and with a FAST-Red selectable marker (Shimada *et al.*, 2010) (see cf paragraph "Gene cloning") in *Camelina sativa* cv. Celine using the floral dip method (Liu *et al.*, 2012). We obtained four independent T1 lines (**Figure S6**). Transgenic expression was measured in T1 plants using RT-qPCR for *WRR4A*$^{Oy\text{-}0}$, with *EF1a* as a housekeeping reference gene. We extracted RNA using the RNeasy Plant Mini Kit (QIAgen) and treated with RNase-Free DNase Set (QIAGEN). Reverse transcription was carried out using SuperScript IV Reverse Transcriptase (ThermoFisher, Waltham, MA, USA). qPCR was performed using a CFX96 Touch Real-Time PCR Detection System. Primers for qPCR analysis of *WRR4A*$^{Oy\text{-}0}$ are GCAAGATAGCGAGCTCCAGA and GCAAGAAACATACAAGTCCTCCA. Primers for qPCR analysis of *EF1a* are CAGGCTGATTGTGCTGTTCTTA and GTTGTATCCGACCTTCTTCAGG. Data were analysed using the double delta Ct method (Livak & Schmittgen, 2001). We measured the segregation of FAST-Red in T2 seeds. Two lines are segregating 15:1 indicating a dual loci T-DNA insertion and two are segregating 3:1 indicating a single locus insertion. From the two single-locus insertion lines, we obtained five and four AcEx1 resistant lines, without any symptoms of infection. From these nine lines, one produced a bag of homozygous T3 seeds (the others producing no seeds or 3:1 segregating seeds). Twelve plants from this line were tested with AcEx1, 11 showed resistance (eight without symptoms and three with a chlorotic response) and one showed susceptibility (**Figure 5**).

Accession numbers:

Arabidopsis thaliana (Arabidopsis) accessions used in this study are Øystese-0 (Oy-0, NASC: N1436), HR-5 (NASC: N76514), Wassilewskija-2 (Ws-2, NASC: N1601) and Columbia (Col-0, NASC: N1092).

Data availability statement:

All relevant data can be found within the manuscript and its supporting materials. The sequences of the genomic clones of $WRR4A^{Oy-0}$ and $WRR4A^{HR-5}$ are deposited at NCBI GenBank as MW533532 and MW533533 respectively.

Acknowledgements:

We thank the Gatsby Foundation (UK) for funding to the Jones lab. We thank Mark Youles in TSL Synbio for his excellent support with Golden Gate cloning and for providing modules. This research was supported in part by the NBI Computing infrastructure for Science (CiS) group and Dan MacLean's group by providing computational infrastructure. B.C., S.F., O.F. and V.C. were supported by Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/L011646/1. A.R. was supported by EMBO LTF (ALTF-842- 2015). B.C, S.W and J.D.G.J. were supported in part by ERC Advanced Investigator grant to JDGJ 'ImmunityByPairDesign' Project ID 669926.

Author contributions:

BC, SF, OF, AR, VC, EH and JJ designed research; BC, SF, OF, AR, SW, VC performed research; BC, SF, OF, AR, VC, EH and JJ analysed data; and BC and JJ wrote the paper. All authors read and approved the final manuscript.

Conflicts of interest statement:

The authors declare that they have no conflict of interests

Supporting materials:

Figure S1: Detailed map of candidate loci in Oy-0

Figure S2: Fine mapping of chromosome 1 QTL

Figure S3: Fine mapping of chromosome 3 QTL

Figure S4: Detailed map of candidate loci in HR-5

Figure S5: Phylogeny of WRR4A based on protein sequences

Figure S6: Selection of a stable WRR4A$^{Oy\text{-}0}$ transgenic *Camelina sativa* line for resistance to AcEX1

Figure S7: Map of Arabidopsis accessions used in this study

Table S1: Summary of the *A. candida* isolates mentioned in this study

Table S2: Phenotype of 283 Arabidopsis accessions in response to AcEx1 infection

Table S3: Arabidopsis v1 RenSeq bait library (Arbor Bioscience, MI, USA) as described by (Jupe *et al.*, 2013)

Table S4: Primers used in this study

Dataset S1: Alignment of *WRR4A* nucleotide sequences used to generate the phylogenic tree on Figure 4. File is on fasta format and was generated using the software MEGA10, with using the recommended parameters of MUSCLE.

Dataset S2: Alignment of WRR4A amino acid sequence used to generate the phylogenic tree on Figure S5. File is on fasta format and was generated using the software MEGA10, with using the recommended parameters of MUSCLE.

References:

Andolfo, G., Jupe, F., Witek, K., Etherington, G.J., Ercolano, M.R. and Jones, J.D.G. (2014) Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC Plant Biol.*, **14**, 120. Available at: http://www.biomedcentral.com/1471-2229/14/120.

Arora, H., Padmaja, K.L., Paritosh, K., Mukhi, N., Tewari, A.K., Mukhopadhyay, A., Gupta, V., Pradhan, A.K. and Pental, D. (2019) BjuWRR1, a CC-NB-LRR gene identified in Brassica juncea, confers resistance to white rust caused by Albugo candida. *Theor. Appl. Genet.*, **132**, 2223–2236. Available at: https://doi.org/10.1007/s00122-019-03350-z

Austin, R.S., Vidaurre, D., Stamatiou, G., et al. (2011) Next-generation mapping of

Arabidopsis genes. *Plant J.*, **67**, 715–725.

Baggs, E., Dagdas, G. and Krasileva, K. V. (2017) NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Curr. Opin. Plant Biol.*, **38**, 59–67. Available at: http://dx.doi.org/10.1016/j.pbi.2017.04.012.

Belhaj, K., Cano, L.M., Prince, D.C., et al. (2017) Arabidopsis late blight: infection of a nonhost plant by *Albugo laibachii* enables full colonization by *Phytophthora infestans*. Cell. Microbiol., **19**, e12628.

Betancor, M.B., Li, K., Bucerzan, V.S., et al. (2018) Oil from transgenic Camelina sativa containing over 25 % n-3 long-chain PUFA as the major lipid source in feed for Atlantic salmon (Salmo salar). *Br. J. Nutr.*, **119**, 1378–1392.

Borhan, M.H., Gunn, N., Cooper, A., Gulden, S., Tör, M., Rimmer, S.R. and Holub, E.B. (2008) WRR4 encodes a TIR-NB-LRR protein that confers broad-spectrum white rust resistance in Arabidopsis thaliana to four physiological races of Albugo candida. *Mol. Plant-Microbe Interact.*, **21**, 757–768.

Borhan, M.H., Holub, E.B., Kindrachuk, C., Omidi, M., Bozorgmanesh-Frad, G. and Rimmer, S.R. (2010) WRR4, a broad-spectrum TIR-NB-LRR gene from Arabidopsis thaliana that confers white rust resistance in transgenic oilseed brassica crops. *Mol. Plant Pathol.*, **11**, 283–291.

Brasseur, G., Rago, J.P. Di, Slonimski, P.P. and Lemesle-Meunier, D. (2001) Analysis of suppressor mutation reveals long distance interactions in the bc1 complex of Saccharomyces cerevisiae. *Biochim. Biophys. Acta - Bioenerg.*, **1506**, 89–102.**Broman, K.W., Wu, H., Sen, Ś. and Churchill, G.A.** (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, **19**, 889–890.

Castel, B. (2019) Natural and CRISPR-induced genetic variation for plant immunity. University of East Anglia, PhD thesis. Available at: https://ueaeprints.uea.ac.uk/id/eprint/71447/.

Cesari, S. (2017) Multiple strategies for pathogen perception by plant immune receptors. *New Phytol.*, **219**, 17–24. Available at: http://doi.wiley.com/10.1111/nph.14877.

Cevik, V., Boutrot, F., Apel, W., et al. (2019) Transgressive segregation reveals mechanisms of Arabidopsis immunity to Brassica-infecting races of white rust (Albugo candida). *Proc. Natl. Acad. Sci.*, **116**, 2767–2773.

Cooper, A.J., Latunde-Dada, A.O., Woods-Tör, A., Lynn, J., Lucas, J.A., Crute, I.R. and

Holub, E.B. (2008) Basic compatibility of Albugo candida in Arabidopsis thaliana and Brassica juncea causes broad-spectrum suppression of innate immunity. *Mol. Plant-Microbe Interact.*, **21**, 745–756.

Davis, J.E., Voisine, C. and Craig, E.A. (1999) Intragenic suppressors of Hsp70 mutants: Interplay between the ATPase- and peptide-binding domains. *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 9269–9276.

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. (2009) Golden gate shuffling: A one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One*, **4**, e5553.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D.G., Patron, N.J. and Marillonnet, S. (2014) A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.*, **3**, 839–843.

Fairhead, S. (2016) Translating genetics of oomycete resistance from *Arabidopsis thaliana* into Brassica production. University of Warwick, PhD thesis. Available at: http://wrap.warwick.ac.uk/90258

Furzer, O.J., Cevik, V., Fairhead, S., Bailey, K., Redkar, A., Schudoma, C., MacLean, D., Holub, E.B., Jones, J.D.G. (2021) PacBio Sequencing of the *Albugo candida* Ac2V genome reveals the expansion of the “CCG” class of effectors. bioRxiv.

Gupta, A.K., Raj, R., Kumari, K., Singh, S.P., Solanki, I.S. and Choudhary, R. (2018) Management of Major Diseases of Indian Mustard Through Balanced Fertilization, Cultural Practices and Fungicides in Calcareous Soils. *Proc. Natl. Acad. Sci. India, Sect. B - Biol. Sci.*, **88**, 229–239.

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

Jouet, A., Saunders, D., McMullan, M., et al. (2018) Albugo candida race diversity, ploidy and host-associated microbes revealed using DNA sequence capture on diseased plants in the field. *New Phytol.*, **221**, 1529–1543. Available at: http://doi.wiley.com/10.1111/nph.15417.

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

Kemen, E., Gardiner, A., Schultz-Larsen, T., et al. (2011) Gene Gain and Loss during

622 Evolution of Obligate Parasitism in the White Rust Pathogen of Arabidopsis thaliana.
623 *PLoS Biol.*, **9**, e1001094. Available at: http://dx.plos.org/10.1371/journal.pbio.1001094.

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

627 **Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E. and Michelmore, R.W.** (2004) Multiple
628 genetic processes result in heterogeneous rates of evolution within the major cluster
629 disease resistance genes in lettuce. *The Plant Cell*, **16**, 2870–2894.

630 **Lander, E.S. and Botstein, D.** (1989) Mapping mendelian factors underlying quantitative
631 traits using RFLP linkage maps. *Genetics*, **121**, 185–199. Available at:
632 http://www.ncbi.nlm.nih.gov/pubmed/2563713.

633 **Lee, R.R.Q. and Chae, E.** (2020) Plant Communications Patterns of NLR Cluster Variation
634 in Arabidopsis thaliana Genomes. *Plant Commun.*, **1**, 100089. Available at:
635 https://doi.org/10.1016/j.xplc.2020.100089.

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

638 **Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,**
639 **Abecasis, G. and Durbin, R.** (2009) The Sequence Alignment/Map format and
640 SAMtools. *Bioinformatics*, **25**, 2078–2079.

641 **Liu, X., Brost, J., Hutcheon, C., et al.** (2012) Transformation of the oilseed crop Camelina
642 sativa by Agrobacterium-mediated floral dip and simple large-scale screening of
643 transformants. *Vitr. Cell. Dev. Biol. - Plant*, **48**, 462–468.

644 **Livak, K.J. and Schmittgen, T.D.** (2001) Analysis of Relative Gene Expression Data Using
645 Real- Time Quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ Method. *Methods*, **25**, 402-408.

646 **Ma, S., Lapin, D., Liu, L., et al.** (2020) Direct pathogen-induced assembly of an NLR
647 immune receptor complex to form a holoenzyme. *Science*, **370**, eabe3069. Available at:
648 https://dx.doi.org/10.1126/science.abd9993.

649 **Martin, R., Qi, T., Zhang, H., Liu, F., King, M., Toth, C. and Staskawicz, B.J.** (2020)
650 Structure of the activated Roq1 resistosome directly recognizing the pathogen effector
651 XopQ. *Science*, **370**, eabd9993. Available at: https://doi.org/10.1126/science.abd9993.

652 **McMullan, M., Gardiner, A., Bailey, K., et al.** (2015) Evidence for suppression of immunity
653 as a driver for genomic introgressions and host range expansion in races of Albugo

654 candida, a generalist parasite. *Elife*, **4**, e04550. Available at:
655 http://elifesciences.org/lookup/doi/10.7554/eLife.04550.

656 **Meyers, B.C., Shen, K.A., Rohani, P., Gaut, B.S. and Michelmore, R.W.** (1998) Receptor-
657 like genes in the major resistance locus of lettuce are subject to divergent selection.
658 *The Plant Cell*, **10**, 1833–1846.

659 **Meyers BC, Morgante M, Michelmore RW.** (2002) TIR-X and TIR-NBS proteins: Two new
660 families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsis
661 and other plant genomes. *The Plant Journal*, **32**, 77–92.

662 **Monteiro, F. and Nishimura, M.T.** (2018) Structural , Functional , and Genomic Diversity of
663 Plant NLR proteins : An Evolved Resource for Rational Engineering of Plant Immunity.
664 *Annu. Rev. Phytopathol.*, **56**, 12.1-12.25.

665 **Napier, J.A., Haslam, R.P., Tsalavouta, M. and Sayanova, O.** (2019) The challenges of
666 delivering genetically modified crops with nutritional enhancement traits. *Nat. Plants*, **5**,
667 563–567. Available at: http://dx.doi.org/10.1038/s41477-019-0430-z.

668 **Napier, J.A., Usher, S., Haslam, R.P., Ruiz-Lopez, N. and Sayanova, O.** (2015)
669 Transgenic plants as a sustainable, terrestrial source of fish oils. *Eur. J. Lipid Sci.*
670 *Technol.*, **117**, 1317–1324.

671 **Nishimasu, H., Shi, X., Ishiguro, S., et al.** (2018) Engineered CRISPR-Cas9 nuclease with
672 expanded targeting space. *Science.*, **361**, 1259–1262.

673 **Pedersen, W.L.** (1988) Resistance To Maintain Residual Effects. *Annu. Rev. Phytopathol.*,
674 **26**, 369–378.

675 **Petrie, J.R., Shrestha, P., Belide, S., et al.** (2014) Metabolic engineering Camelina sativa
676 with fish oil-like levels of DHA. *PLoS One*, **9**, e85061.

677 **Pound G.S. and Williams P.H**. (1963) Biological races of *Albugo candida. Phytopathol*., **53**,
678 1146-1149.

679 **Prince, D.C., Rallapalli, G., Xu, D., et al.** (2017) Albugo-imposed changes to tryptophan-
680 derived antimicrobial metabolite biosynthesis may contribute to suppression of non-host
681 resistance to Phytophthora infestans in Arabidopsis thaliana. *BMC Biol.*, **15**, 20.
682 Available at: http://bmcbiol.biomedcentral.com/articles/10.1186/s12915-017-0360-z.

683 **Redkar, A., Cevik, V., Bailey, K., Furzer, O.J., Fairhead, S., Borhan, M.H., Holub, E.B.,**
684 **Jones, J.D.G.** (2021) The Arabidopsis WRR4A and WRR4B paralogous NLR proteins
685 both confer recognition of multiple *Albugo candida* effectors. bioRxiv.

Robinson, J.T., Thorvaldsdóttir, H., Wenger, A.M., Zehir, A. and Mesirov, J.P. (2017) Variant review with the integrative genomics viewer. *Cancer Res.*, **77**, e31–e34.

Roux, C. Le, Huet, G., Jauneau, A., et al. (2015) A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling of Transcription Factors to Immunity. *Cell*, **161**, 1074–1088. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867415004420.

Ruiz-Lopez, N., Haslam, R.P., Napier, J.A. and Sayanova, O. (2014) Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. *Plant J.*, **77**, 198–208.

Sarris, P.F., Duxbury, Z., Huh, S.U., et al. (2015) A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. *Cell*, **161**, 1089–1100. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867415004419.

Saucet, S.B., Esmenjaud, D. and Ghelder, C. Van (2021) Integrity of the post-LRR domain is required for TNLs' function. *Mol. Plant-Microbe Interact.*, **34**, 286-296.

Schülein, R., Zühlke, K., Krause, G. and Rosenthal, W. (2001) Functional Rescue of the Nephrogenic Diabetes Insipidus-causing Vasopressin V2 Receptor Mutants G185C and R202C by a Second Site Suppressor Mutation. *J. Biol. Chem.*, **276**, 8384–8392.

Séguin-Swartz, G., Eynck, C., Gugel, R.K., et al. (2009) Diseases of Camelina sativa (false flax). *Can. J. Plant Pathol.*, **31**, 375–386.

Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.*, **61**, 519–528. Available at: http://doi.wiley.com/10.1111/j.1365-313X.2009.04060.x.

Simon, M., Loudet, O., Durand, S., Bérard, A., Brunel, D., Sennesal, F.X., Durand-Tardif, M., Pelletier, G. and Camilleri, C. (2008) Quantitative trait loci mapping in five new large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide polymorphism markers. *Genetics*, **178**, 2253–2264.

Steuernagel, B., Periyannan, S.K., Hernández-Pinzón, I., et al. (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.*, **34**, 652–655.

Van Ghelder, C., Esmenjenaud, D. (2016) TNL genes in peach: Insights into the post-LRR domain. *BMC Genomics*, **17**, 317. Available at: http://doi.org/10.1186/s12864-016-2635-0

Wan, W., Kim, S., Castel, B., Charoennit, N. and Chae, E. (2021) Genetics of autoimmunity in plants: an evolutionary genetics perspective. *New Phytol.*, **229**, 1215–1233.

Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, **6**, e16765.

West, A.L., Miles, E.A., Lillycrop, K.A., Han, L., Sayanova, O., Napier, J.A., Calder, P.C. and Burdge, G.C. (2019) Postprandial incorporation of EPA and DHA from transgenic Camelina sativa oil into blood lipids is equivalent to that from fish oil in healthy humans. *Br. J. Nutr.*, **121**, 1235–1246.

Weyer, A.-L. Van de, Monteiro, F., Furzer, O.J., et al. (2019) A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana. *Cell*, **178**, 1260–1272. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0092867419308372.

Witek, K., Jupe, F., Witek, A.I., Baker, D., Clark, M.D. and Jones, J.D.G. (2016) Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.*, **34**, 656–660. Available at: http://dx.doi.org/10.1038/nbt.3540.

Figure legends:

Figure 1: Oy-0 and HR-5 alleles of WRR4A confer full resistance to AcEx1

5-week old plants were sprayed inoculated with AcEx1. Plants were phenotyped 14 days after inoculation. **a.** AcEx1 response in nature Arabidopsis accessions and mutants. Indicated genotypes always display this phenotype in response to AcEx1. **b.** AcEx1 response in transgenic Ws-2 expressing *WRR4A*$^{Oy-0}$, *WRR4B*$^{Oy-0}$, *WRR4D*$^{Oy-0}$, *WRR4A*$^{Col-0}$ or *CWR11*$^{Oy-0}$. Numbers indicate the number of independent transgenic lines showing similar phenotype out of the number of independent transgenic lines tested. Red arrows indicate a chlorotic response seen in susceptible lines containing *WRR4A*$^{Col-0}$ (*i.e.* Col-0 WT and Ws-2_*WRR4A*$^{Col-0}$ transgenic). Adaxial picture of the leaves has been added to illustrate the chlorotic response.

Figure 2: Allelic variation between Col-0, HR-5 and Oy-0 alleles of WRR4A

a. Nucleotide sequence alignment of *WRR4A* alleles. Plain yellow areas represent exons. Yellow lines represent introns. bp: base pair **b.** Amino acid alignment. a.a: amino acid. The C-terminal extension is framed in yellow for Col-0 to indicate that an early stop codon avoids translation of this sequence. **a.b.** Cartoons made with CLC Workbench Main. Green represents identity. Red represents polymorphism. Figures are on scale.

Figure 3: Recognition of CCG effectors by WT and stop codon mutant alleles of *WRR4A*

CCG effector candidates were transiently expressed in 4-week old *N. tabacum* leaves, under the control of the 35S promoter and Ocs terminator, alone or with WT or mutant alleles of *WRR4A*. Leaves were infiltrated with *Agrobacterium tumefaciens* strain GV3101 in infiltration buffer at OD_{600} = 0.4. Pictures were taken at 4 dpi. **a.** cartoon of the WRR4A alleles: (a) Col-0 WT, (b) Oy-0 WT, (c) Col-0 with TGA-TGC mutation, causing an Oy-0 like C-terminal extension, (d) Oy-0 with a TGC-TGA mutation causing a truncation of the C-terminal extension. **b.** AcEx1 CCG effector candidates alone (e) or with one of the four WRR4A alleles as shown in Fig 3a (a), (b), (c), (d). MLA7 CC domain was used as an HR positive control (f). Numbers indicate the number of positive HR observed out of the number of infiltrations conducted. **c.** Eight CCGs from other races of *Albugo candida* known to be recognized by Col-0 allele of WRR4A were tested with the three others WRR4A alleles. CCG effector candidates alone (e) or with one of the four WRR4A alleles as shown in Fig 3a (a), (b), (c), (d). For CCG28, CCG30, CCG33 and the control leaf, WRR4A^{Oy-0} (b) is infiltrated on the top right of the leaf, instead of WRR4A$^{Col-0_LONG}$. Numbers indicate the number of positive HR observed out of the number of infiltrations conducted. **d.** Expression of CCG35 (68 kDa) and CCG39 (72 kDa) showed by western blot with anti-V5 antibody. Expression of WRR4 alleles (not tagged) by RT-PCR (197 bp), using *NbActin* as a control (143 bp). **e.**

Summary of the CCG recognition by WRR4A alleles. Red nails: AcEx1 effectors, black nails: CCGs from other *A. candida* races.

Figure 4: An early stop codon in *WRR4A* is associated with AcEx1 susceptibility

WRR4A genomic sequence of 20 Arabidopsis accessions were extracted from http://ann-nblrrome.tuebingen.mpg.de/apollo/jbrowse/ (Van de Weyer *et al.*, 2019). Nucleotide sequences corresponding from ATG to TAA of the Oy-0 allele (including introns) were aligned using MUSCLE (software: MEGA10, the alignment is available as Dataset S1 on Supporting materials). WRR4B from Col-0 was used as outgroup. A phylogenetic tree was generated using the Maximum Likelihood method and a bootstrap (100 replicates) was calculated (software: MEGA10). The tree is drawn to scale (apart the two broken branches, whose length is indicated in parenthesis), with branch lengths measured in the number of substitutions per site. The resistance / susceptibility phenotypes are indicated. Cartoons on the right represent WRR4A predicted protein, on scale. TIR, NB-ARC, LRR and C-JID are indicated in the Col-0 allele. Dashed orange line represents the Col-0 stop codon. Dashed blue line represents the Oy-0 stop codon.

Figure 5: WRR4A confers resistance to AcEx1 in camelina crop

Five-week old camelina (cultivar Celine) plants were sprayed inoculated with AcEx1 race of the white rust oomycete pathogen *A. candida*. Pictures were taken 12 dpi (day post inoculation). **Top row:** Twelve wild-type plants all show mild to severe white rust symptoms. **Bottom row:** twelve lines transformed with $WRR4A^{Oy\text{-}0}$ were tested. One shows mild white rust symptoms, three show local chlorotic response and eight show complete green resistance. White dash line indicates sporulation; red dash line indicates a chlorotic response with no pustule formation.

Figure S1: Detailed map of candidate loci in Oy-0

a. Genome scans using the maximum likelihood algorithm (logarithm of odds (LOD) of 2.5 for both isolates at a 5% confidence interval). Numbers on x axis indicate chromosomes. Three QTLs are over the LOD-score threshold: *WRR4*, *WRR11* and *WRR15*. **b.** *WRR4* locus in Col-0 and two RenSeq contigs containing WRR4 paralogs from Oy-0 (Van de Weyer *et al.*, 2019). **c.** Unique RenSeq contig from Oy-0 sharing identify with the *WRR11* locus. It contains a CC-NLR absent from Col-0. The corresponding locus in Col-0 is displayed on top. Loci are on scale.

Figure S2: Fine mapping of chromosome 1 QTL

Six RILs recombines within the Chromosome 1 QTL. Three carry the Oy-0 allele (resistant): 19, 273 and 336. Three carry the Col-0 allele (susceptible): 91, 118, 298. Fine mapping was conducted using two markers between *WRR4* and *RPP7* NLR clusters. Dark grey indicates the region containing the gene. Light grey indicates the region excluding the gene. Dash grey indicates the region containing the Oy-0 / Col-0 recombination site. Numbers indicate the position in Mb on Chromosome 1. QTL: Quantitative Trait Locus (indicates the borders of the QTL before refining). SNP: Single Nucleotide Polymorphism. AFLP: Amplified Fragment Length Polymorphism. Chr1/C1: Chromosome 1. Figure is not on scale.

Figure S3: Fine mapping of chromosome 3 QTL

Six RILs recombines within the Chromosome 3 QTL. Three carry the Oy-0 allele (resistant): 155, 304, 403. Three carry the Col-0 allele (susceptible): 471, 472, 497. Fine mapping was conducted using five markers along the QTL. Dark grey indicates the region containing the gene. Light grey indicates the region excluding the gene. Dash grey indicates the region containing the Oy-0 / Col-0 recombination site. Numbers indicate the position in Mb on Chromosome 3. QTL: Quantitative Trait Locus (indicates the borders of the QTL before refining). SNP: Single Nucleotide Polymorphism. AFLP: Amplified Fragment Length Polymorphism. CAPS: Cleaved Amplified Polymorphic Sequence. Chr3/C3: Chromosome 3. Figure is not on scale.

Figure S4: Detailed map of candidate loci in HR-5

a. Homozygosity score of ~100 bulked F2 lines susceptible to AcEx1. Bulked lines were sequence upon *R*-gene enrichment (RenSeq). *WRR4* and *RPP7* cluster present a high degree of homozygosity. This panel was produced using the NGM system where the different coloured bands represent the density of SNPs at different allele frequency levels, used to assess the degree of linkage across the genome. **b.** *WRR4* locus in Col-0 and two RenSeq contigs containing WRR4 paralogs from HR-5 (Van de Weyer *et al.*, 2019).

Figure S5: Phylogeny of WRR4A based on protein sequences

WRR4A proteins sequence of 16 Arabidopsis accessions were predicted using Augustus (http://bioinf.uni-greifswald.de/augustus/). The C-terminal extension was not used for alignment. Proteins were aligned using MUSCLE (software: MEGA10, Dataset S2). Pseudogenised WRR4A alleles (Ws-0, Yo-0, Ler-0, Wil-2, ULL2-5 and Bay-0) were not used. WRR4B from Col-0 was used as an outgroup. A phylogenetic tree was generated using the Maximum Likelihood method and a bootstrap (100 replicates) was calculated (software: MEGA10). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

838

839 Figure S6: Selection of a stable WRR4A$^{Oy\text{-}0}$ transgenic *Camelina sativa* line for resistance to
840 AcEX1

841 *WRR4A$^{Oy\text{-}0}$* was transformed in *Camelina sativa* cultivar Celine using the floral dip method. Four
842 independent transformants were obtained. Transgene expression is indicated in fold difference as
843 compared to EF1a. Transgene segregation was estimated by counting the FAST-Red marker in T2
844 seeds, using the χ^2 method. Phenotyping was conducted 12 days upon inoculation of AcEx1.
845 Representative phenotypes are indicated in the blue frame. We identified one bag of 100% red seeds
846 from 9 resistant plants. This line was kept for further analyses (see figure 5).

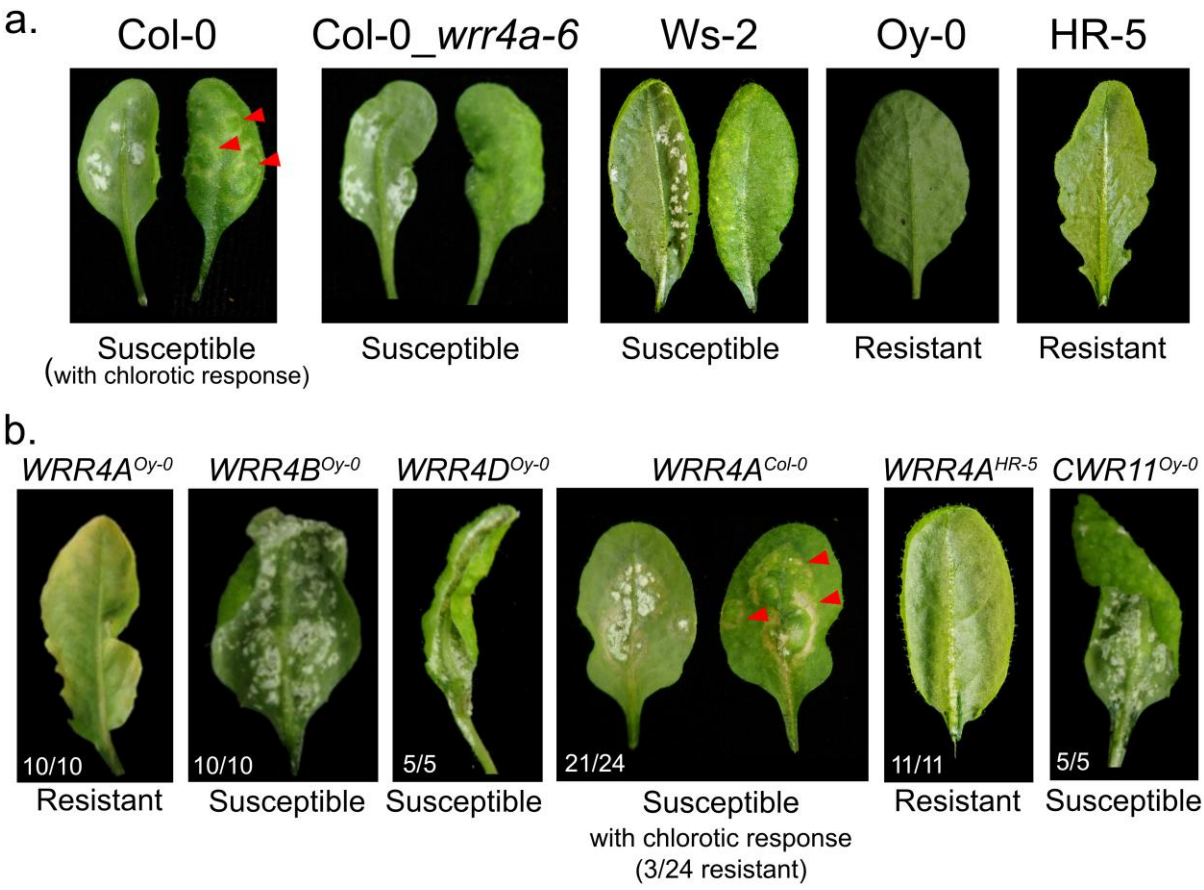
847

848 Figure S7: Map of Arabidopsis accessions used in this study

849 Green: *WRR4A* Oy-0 like (with C-terminal extension). Blue: *WRR4A* Col-0-like (without C-terminal
850 extension). Purple: *WRR4A* with earlier stop codon, likely not functional. Black: *WRR4A* missing. From
851 West to East: Yo-0 (USA) Col-0 (USA), Bur-0 (Ireland), C24 (Portugal), Ped-0 (Spain), Edi-0 (Scotland),
852 HR-5 (England), Sf-2 (Spain), Oy-0 (Norway), Hi-0 (Netherlands), Po-0 (Germany), Zu-0 (Switzerland),
853 Wu-0 (Germany), Bay-0 (Germany), No-0 (Germany), ULL2-5 (Sweden), Ler-0 (Poland), Mt-0 (Libya),
854 Kn-0 (Lithuania), Wil-2 (Lithuania), Ws-0 (Ukraine), Ws-2 (Ukraine, in black hidden behind Ws-0 in red),
855 Rsch-4 (Russia), Tsu-0 (Japan). Map created with the online tool “My Maps” (Google). Interactive map
856 available online:
857 https://drive.google.com/open?id=1MjXiwQvqzWjdNExQ7gHcBKhHUufM8zFk&usp=sharing.

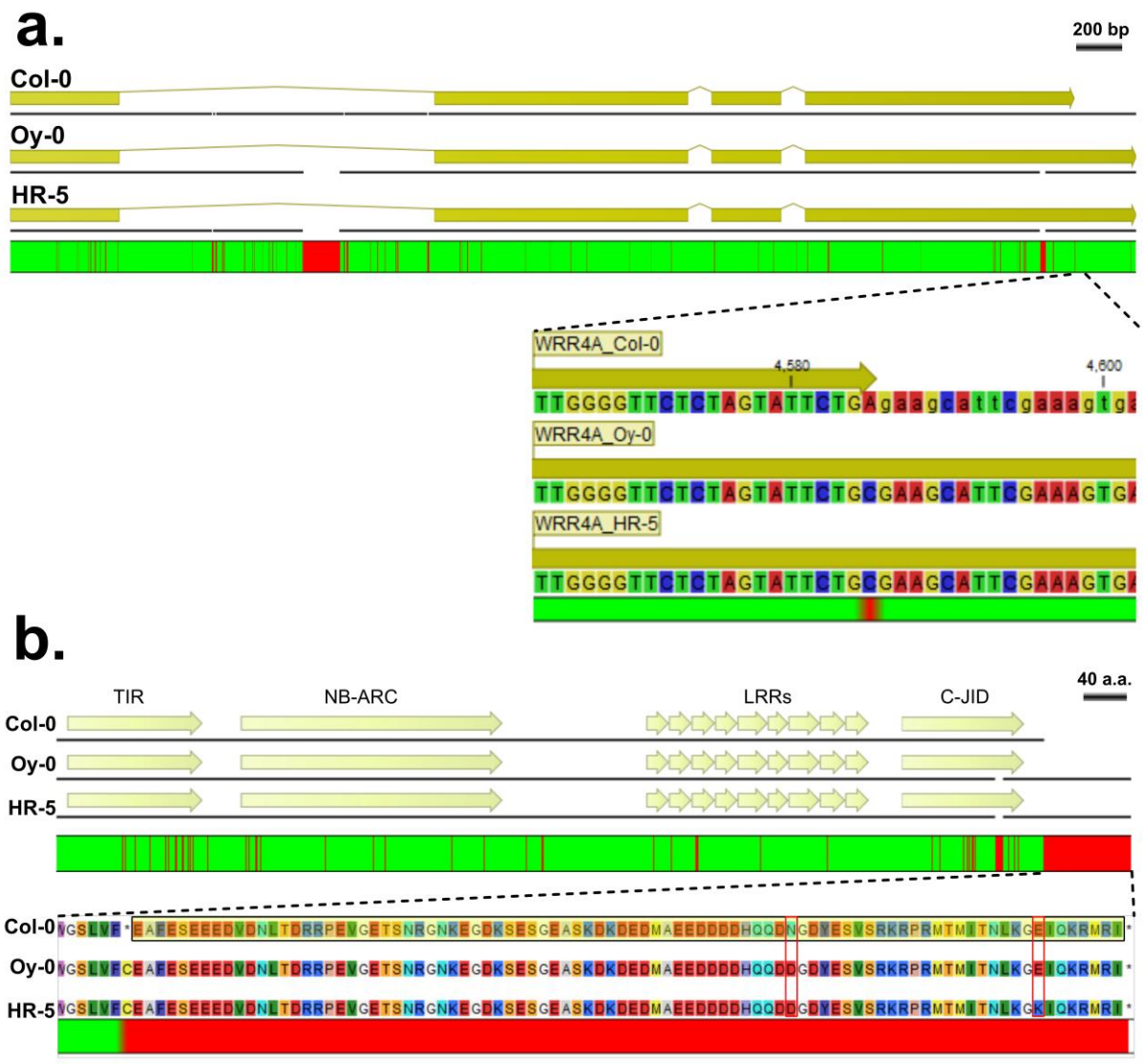
858

859



860

861 Figure 1



862

863 Figure 2

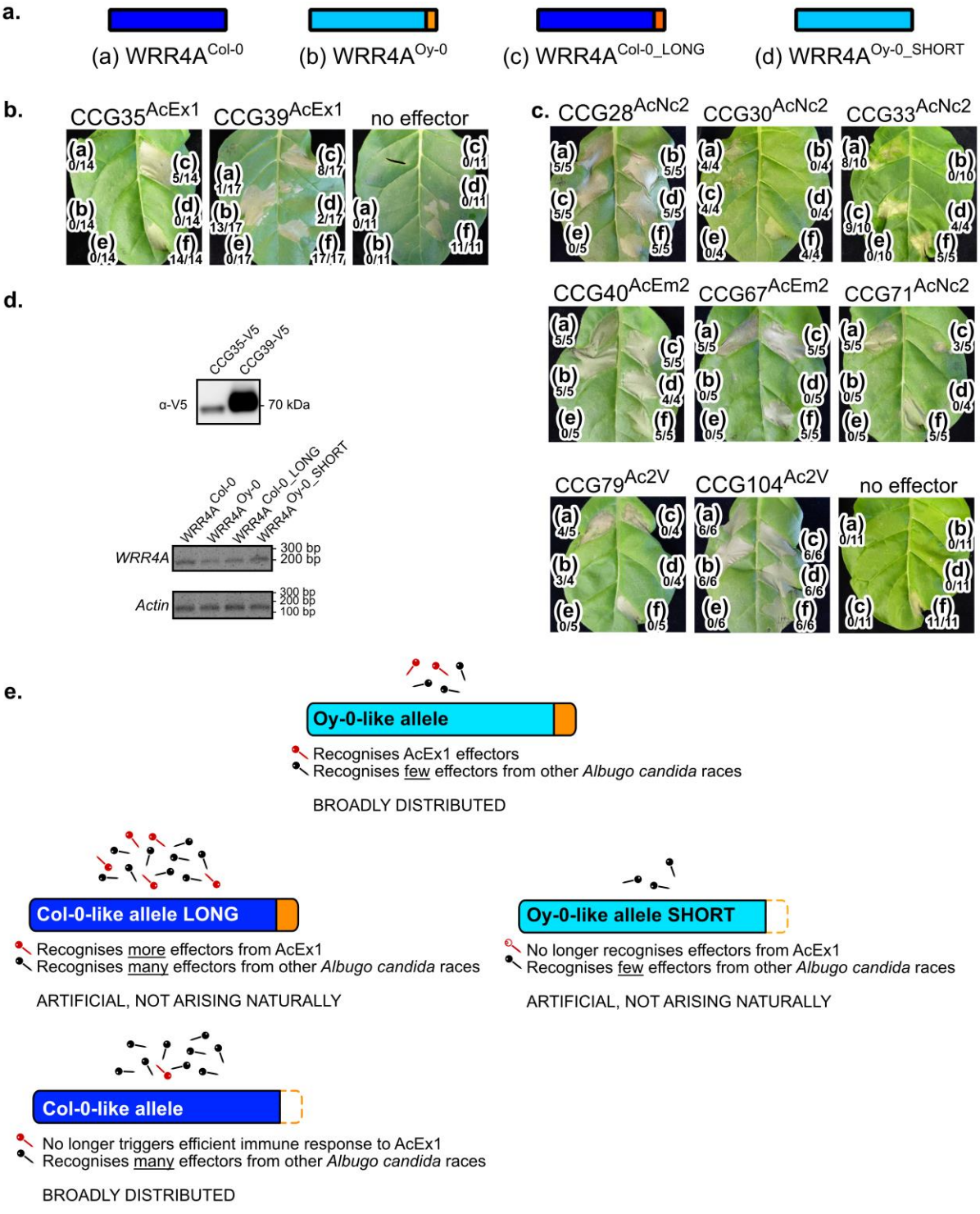
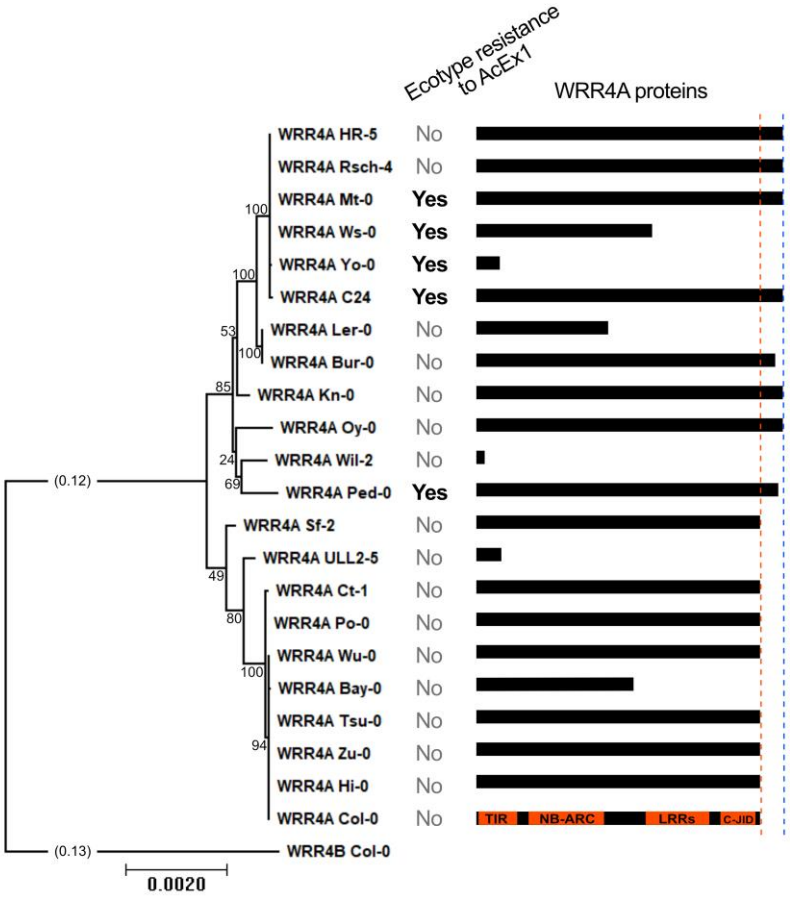

a.
(a) WRR4A Col-0
(b) WRR4A Oy-0
(c) WRR4A Col-0_LONG
(d) WRR4A Oy-0_SHORT
b.
CCG35 AcEx1
CCG39 AcEx1
no effector
c.
CCG28 AcNc2
CCG30 AcNc2
CCG33 AcNc2
CCG40 AcEm2
CCG67 AcEm2
CCG71 AcNc2
CCG79 Ac2V
CCG104 Ac2V
no effector
d.
CCG35-V5
CCG39-V5
α-V5
70 kDa
WRR4A Col-0
WRR4A Oy-0
WRR4A Col-0_LONG
WRR4A Oy-0_SHORT
WRR4A
Actin
300 bp
200 bp
100 bp
e.
Oy-0-like allele
Recognises AcEx1 effectors
Recognises few effectors from other Albugo candida races
BROADLY DISTRIBUTED
Col-0-like allele LONG
Recognises more effectors from AcEx1
Recognises many effectors from other Albugo candida races
ARTIFICIAL, NOT ARISING NATURALLY
Oy-0-like allele SHORT
No longer recognises effectors from AcEx1
Recognises few effectors from other Albugo candida races
ARTIFICIAL, NOT ARISING NATURALLY
Col-0-like allele
No longer triggers efficient immune response to AcEx1
Recognises many effectors from other Albugo candida races
BROADLY DISTRIBUTED

Figure 3

867

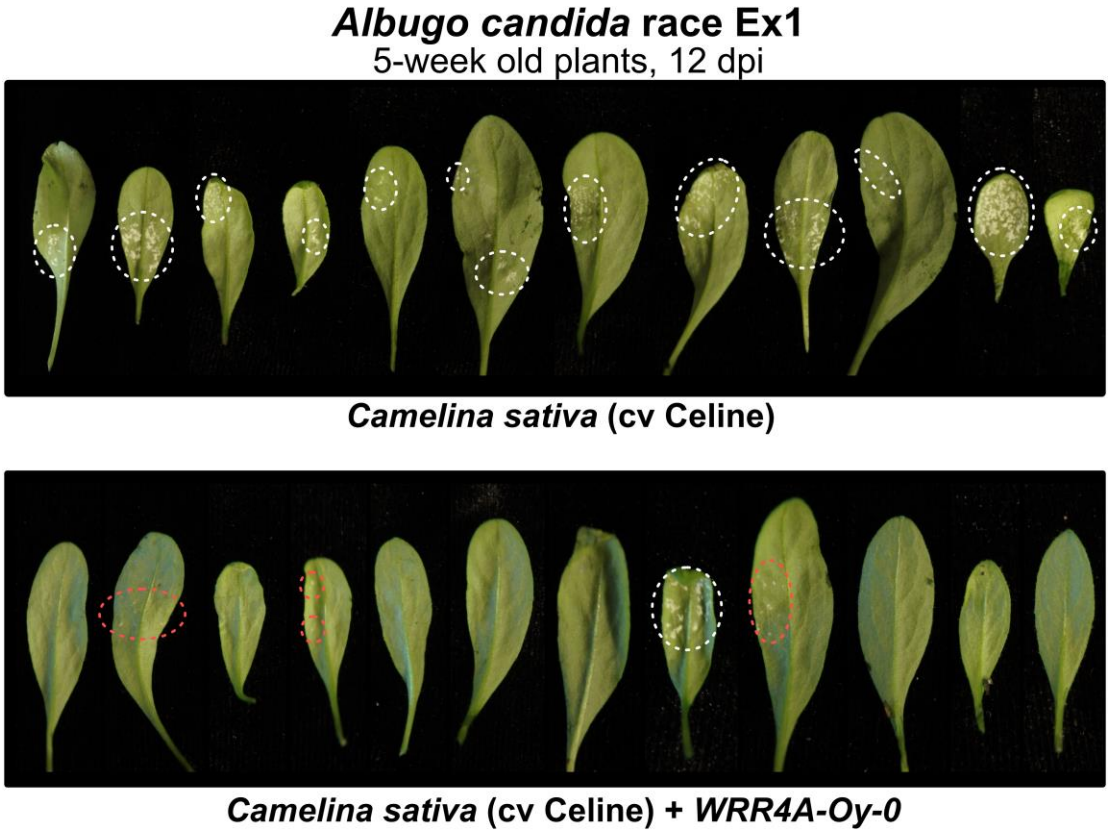


868

869 Figure 4

870

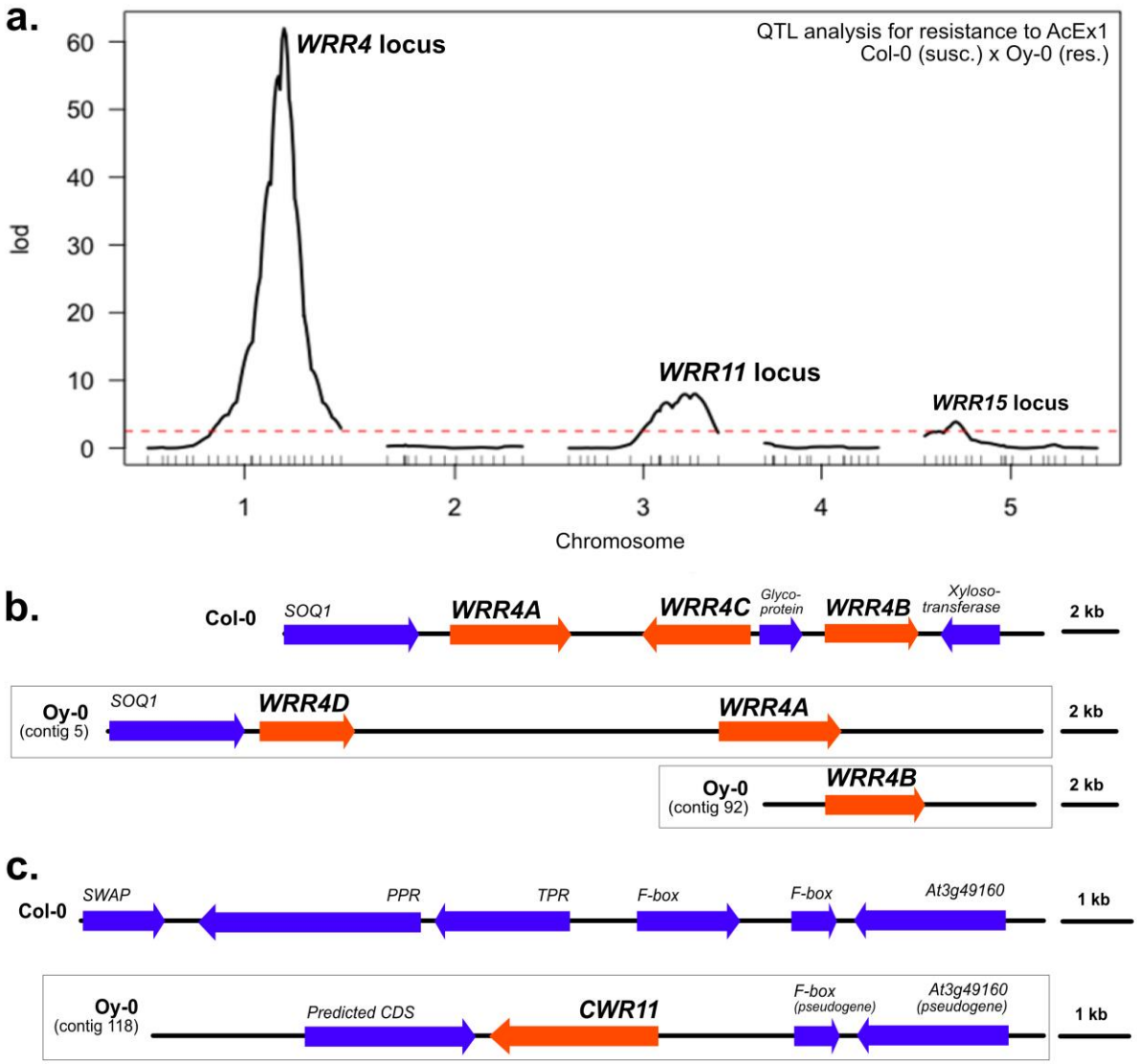
871



872

873 Figure 5

874

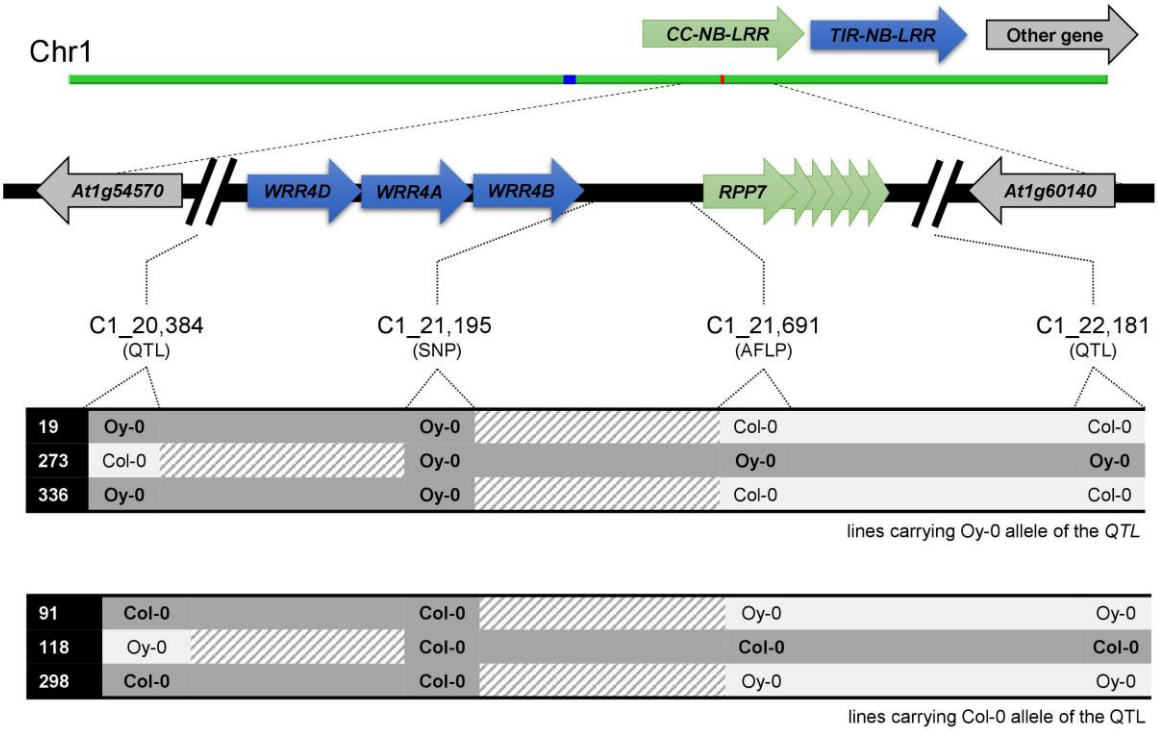


875

876 Figure S1

877

878

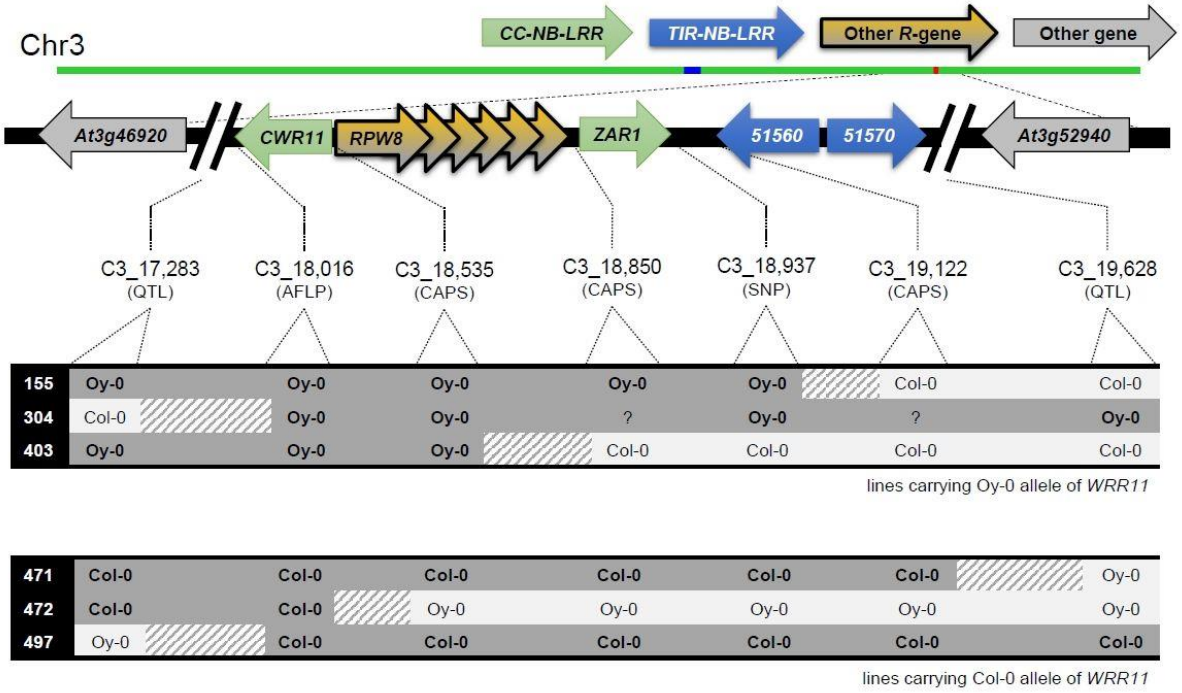


879

880 Figure S2

881

882

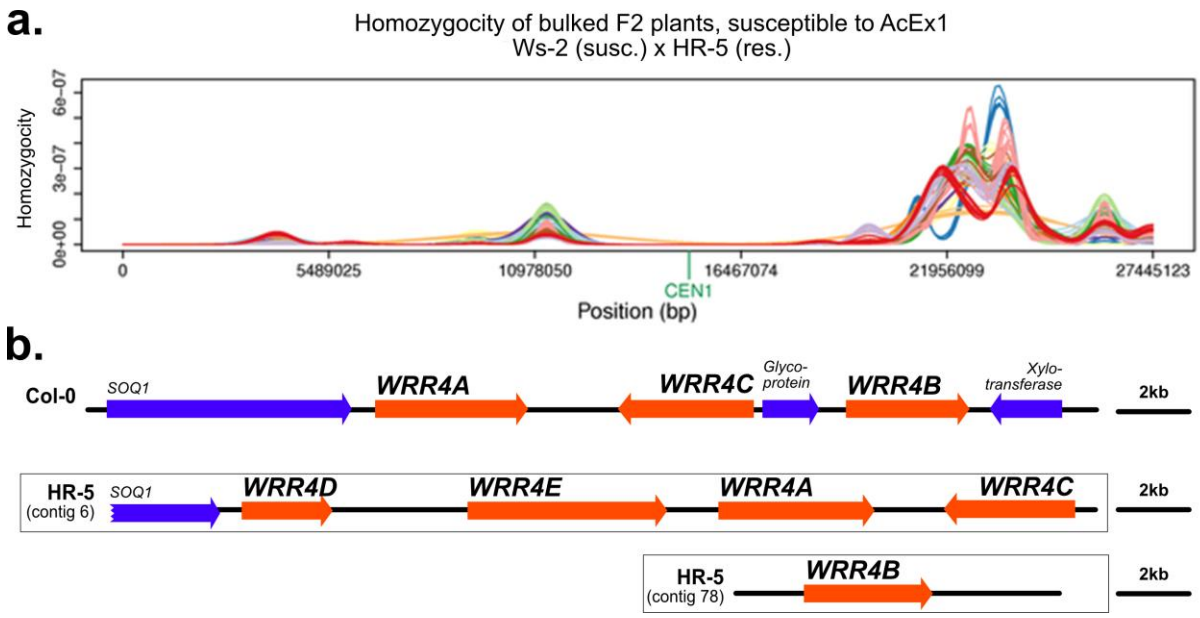


883

884 Figure S3

885

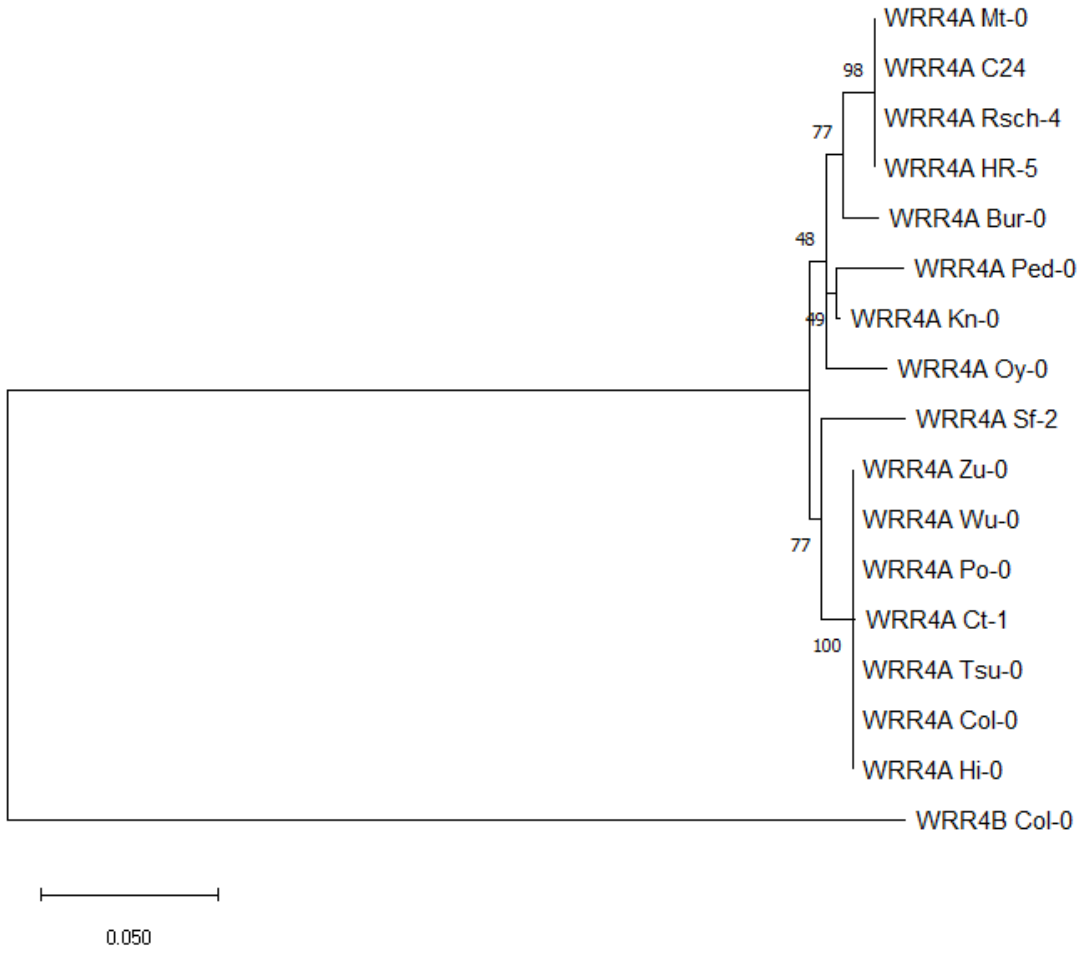
886



887

888 Figure S4

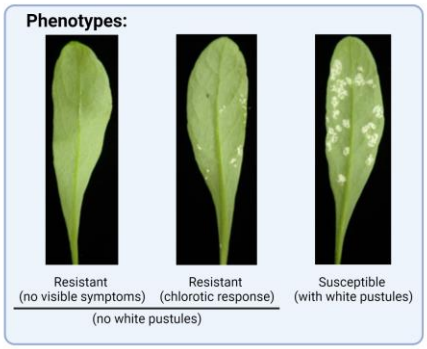
889



890

891 Figure S5

892



		Camelina sativa (cv Celine) (WT)	Line 1 (WRR4A$^{Oy\text{-}0}$)	Line 2 (WRR4A$^{Oy\text{-}0}$)	Line 3 (WRR4A$^{Oy\text{-}0}$)	Line 4 (WRR4A$^{Oy\text{-}0}$)
WRR4$^{Oy\text{-}0}$ expression level		0	5.95	12.58	6.01	7.23
Transgene segregation		N/A	15:1	3:1	15:1	3:1
AcEx1 response (in segregating T2)	Resistant (no visible symptoms):	2	0	**5**	0	**4**
	Resistant (chlorotic response):	0	2	2	1	3
	Susceptible (with white pustules):	8	8	2	9	3
Homozygous T3 (from of a resistant T2, from a single site T-DNA insertion T1)		N/A	N/A	0	N/A	1

Figure S6


Greenland
Iceland
Finland
Sweden
Russia
Canada
United States
Mexico
North Atlantic Ocean
Ukraine
Kazakhstan
Mongolia
China
South Korea
Turkey
Iraq
Iran
Afghanistan
Pakistan
India
Thailand
Algeria
Libya
Egypt
Saudi Arabia
Mali
Niger
Chad
Sudan
Nigeria
Ethiopia
Kenya
DRC
Tanzania
Angola
Namibia
Botswana
Madagascar
South Africa
South Atlantic Ocean
Indian Ocean
Indonesia
Australia
Venezuela
Colombia
Peru
Brazil
Bolivia
Chile
Argentina

Figure S7