Supporting Information

Nemri et al. 10.1073/pnas.0913160107

SI Text

Arabidopsis thaliana Genotypes and Growth. The Nordborg set of 96 *A. thaliana* accessions (1) was provided by C. Dean's group (JIC, Norwich, UK). Sets of *A. thaliana* (*At*) recombinant inbred lines (RILs) were kindly provided by I. Bancroft's group (JIC, Norwich, UK) (2), with exception of the Bay-0 × Shahdara population, which was kindly provided by O. Loudet (INRA, Versailles, France) (3). *A. thaliana* plants were grown as seeds and sown on a 12:1 mixture of Scott's M2 plus Intercept to grit, and they were stratified for 2 days at 4°C in the dark. Then, they were moved to growth chambers kept at 20–22 °C and 70% humidity under 10 h light and 14 h dark. Seeds were covered with a clear polystyrene lid for 2 days to promote germination. When necessary, accessions were vernalized in the dark for 2–3 months at 4 °C before germination.

- 1. Nordborg M, et al. (2005) The pattern of polymorphism in Arabidopsis thaliana. *PLoS Biol* 3:e196.
- O'Neill CM, et al. (2008) Six new recombinant inbred populations for the study of quantitative traits in Arabidopsis thaliana. Theor Appl Genet 116:623–634.

Hyaloperonospora arabidopsidis Materials and Infections. The isolates Emwa1, Hiks1, Emco5, Cala2, and Emoy2 were originally collected and provided by E. Holub (HRI, Warwick, UK). Noco2 was isolated and propagated in the Sainsbury Laboratory in the United Kingdom. Propagation, inoculumpreparation, infection, and incubation were conducted as described previously (4). After 7 days of incubation, the interaction was scored specifically on first true leaves as compatible, incompatible, or intermediate depending on the consistency of presence/absence of sporangiophores. Disease assessments on each genotype were done on 5–10 seedlings and repeated independently at least three times. In some accession × isolate interactions, variation in occurrence of sporulation was consistently observed within and between experiments, and this is referred to as intermediate.

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F (2002) Bay-0 x Shahdara recombinant inbred line population: A powerful tool for the genetic dissection of complex traits in Arabidopsis. *Theor Appl Genet* 104:1173–1184.

Reignault P, et al. (1996) Four Arabidopsis RPP loci controlling resistance to the Noco2 isolate of Peronospora parasitica map to regions known to contain other RPP recognition specificities. Mol Plant Microbe Interact 9:464–473.



Fig. S1. (Continued)



Fig. S1. (Continued)



Fig. S1. Geographic distribution of resistance phenotypes in a set of 96 *A*. *thaliana* accessions challenged with five isolates of *H*. *arabidopsidis* ex *parasitica* (*Hpa*). Blue dots indicate resistant accessions, and red dots indicate susceptible accessions. Intermediate interactions are not reported. Accessions from the same origin may appear as a single dot but generally showed the same resistance phenotype.

DNAS Nd



Fig. 52. Genome-wide linkage and association mapping of resistance to five isolates of *H. arabidopsidis* ex *parasitica* in a set of *A. thaliana* RIL populations and accessions. All five chromosomes are pictured in scale and numbered I–V. *Upper* displays the location of 19 resistance QTLs. *Lower* shows results of association mapping using the random forest method on the 250K SNP dataset and four treatments (i.e., Hpa isolate). Genome-wide association mapping (GWAM) results for Emoy2 resistance mapping are not presented. Details on the sources of the resistance QTLs and their recognition specificities can be found in Table 1. The QTL colors in *Upper* correspond to the following RIL population × isolate interactions: Cvi-0xAg-0/Emwa1 (1), Nok-3xGa-0/Hik1 (2), KonxBr-0/Emwa1 (3), Cvi-0xAg-0/noco2 (4), Bay-0xShah/Emco5 (6), Bay-0xShah/Emco5 (7), SorboxGy-0/Noco2 (8), SorboxGy-0/Emwa1 (9), and Nok-3xGa-0/Emwa1 (10).



Fig. S3. GWAM of downy mildew resistance in a set of 96 *A. thaliana* accession results in enrichment of defense-related genes. Enrichment of candidate SNPs is shown as function of (*A*) Fischer's exact test *P* values and (*B*) random forest number of top SNPs. The gene list was constructed by searching The Arabidopsis Information Resource (http://www.arabidopsis.org/) for genes with suitable annotation and literature searches; however, literature searches were not extensive. The gene list was constructed before looking at the association plots.



Fig. S4. (Continued)



Fig. S4. (Continued)



Fig. 54. Haplotypes analysis of selected polymorphisms. *Top* displays associations with two statistical tests, random forest (red circles) and Fisher's exact (blue circles). *Middle* displays the physical position of gene models in the genomic interval; exons are displayed as boxes, introns as lines, red filling as a translated region, and blank as an untranslated region. *Bottom* is an SNP matrix with each row as an accession and each column as an SNP. The accession (name and country of origin) is displayed on the right with a colored bar indicative of the resistance phenotype of the accessions. Below is the phenotype legend with 0 (blue) coding for susceptibility and 1 (red) coding for resistance. Within the SNP matrix, blue represents major alleles, whereas yellow represents minor alleles. The color gradient is proportional to the Fisher's exact score, with high scores shown as stronger colors. The left side of the SNP matrix is a principal component (PC) 1 and 2 are obtained by running Principal Component Analysis (PCA) on the displayed SNP matrix only, not the whole genome data. The PCA aims to visualize the correspondence between the genotype clustering and the phenotype clustering within this genomic region. (A) GWAM of resistance to Hiks1 on chromosome II between positions 21745097 and 21785097 within the translated regions of *RPP7* (At1g56802) shows high association compared with SNPs in introns. (C) GWAM of resistance to Hiks1 on chromosome V between positions 17474817 and 17484817 shows association upstream and within *RPP8* coding sequence.



Fig. 55. Haplotypes analysis at the *RPP13* locus in relation to Emco5 resistance and susceptibility. (A) Association mapping of Emco5 resistance was performed on a 580-bp sequence of the Leucine-rich repeat-encoding domain of *RPP13* characterized for each of the 96 accessions (36). The monogenic nature of the *RPP13* locus enabled this genetic characterization as opposed to the complex structure of most other known *RPP* loci. Very strong (P < 0.001; χ^2 test) associations were observed at the *RPP13* locus at multiple locations. The multiple peaks pattern may be caused by either causal polymorphism in different regions within the gene or linkage disequilibrium between the causal SNPs and other noncausal SNPs. (B) A phylogenetic tree was constructed using the predicted protein sequence encoded by a 580-bp sequence of the LRR-encoding domain of *RPP13* of 96 At accessions. The resistance/susceptibility to *Hpa* isolate Emco5 of the accessions is represented as red (resistant) and blue (susceptible) bars to the right of the accession names. Despite the extreme amino acid diversity in the RPP13 protein, two clades show a clear enrichment for susceptible accessions, whereas the clade containing Nd-1 (resistant to Emco5 through RPP13) contains only resistant accessions. The associations in *A* were highest at the nucleotides that separate these clades.

Table S1.	Previously characterized res	sistance genes that control downy mildew resistance in A. thaliana with	າ their
genome lo	cation, accessions that carry	y them, and recognition specificities	

					H. arabidopsis isolates detected											
Chr	MRC supercluster	R Locus	Col-0 orthologs	Mol st	Cala2	Emco5	Emoy2	Emwa1	Hiks1	Noco2						
I	B-2 B	RPP7 RPP27 RPP9	At1g58602	CNL					Col-0, Ler-0 Ler-0 Wei-0							
III	F-1 F-1 F-1	RPP1 _{WsA} RPP1 _{WsB} RPP1 _{WsC}	At3g44480	TNL TNL TNL	Ws-0		Ws-0 Ws-0, Nd-1		Nd-1	Ws-0 Ws-0 Ws-0						
IV	F-2 H-1	RPP13 RPP2a	At3g46530 At4g19510	CNL TNL	Col-0, Oy-0	Nd-1										
v	H-2 J-1 J	RPP5 RPP8 RPP31	At4g16860 At5g43470	TNL CNL		Ler-0 Col-0	Ler-0, Col-0 Ler-0	Ler-0, Col-0	Ler-0	Ler-0						

The table represents genetic determinants of incompatible interactions identified in various host accessions against six isolates of *H. arabidopsidis*. The lines represent the loci and when known, the genes controlling resistance. Their recognition specificities are represented in columns. The intersection of line (RPP locus) and columns (*Ha* isolate) indicates the accession(s) in which the locus confers resistance to the isolate. Chr, chromomsome; MRC, major recognition complex; Col-0 orthologs, gene references in Col-0; Mol st, molecular structure of protein encoded; TNL, Toll-like interleukin receptor-nucleotide-binding, leucine-rich repeat (NLR) encoding gene; CNL, Coiled-coil-NLR encoding gene.

PNAS PNAS

Table S2. Survey of natural variation in disease resistance interactions between 96 accessions of *A. thaliana* and five isolates of *H. arabidopsidis* ex *parasitica*

Host disease phenotype on	challenge with
H. arabidopsis is	olate

				,		
Accession	Stock number	Emco5	Emwa1	Emoy2	Hiks1	Noco2
Ag-0	CS22630	R	R	R	R	R
An-1	CS22626	S	R	S	S	S
Bay-0	CS22633	I	R	S	S	I
Bil-5	CS22578	I	R	R	R	R
Bil-7	CS22579	R	R	_	_	S
Bor-1	CS22590	R	R	R	R	R
Bor-4	CS22591	R	R	R	R	R
Br-0	CS22628	R	R	R	R	R
Bur-0	CS22656	S	S	S	R	R
C24	CS22620	R	R	R	R	R
CIBC-17	CS22603	S	S	R	—	S
CIBC-5	CS22602	R	R	R	R	R
Col-0	CS22625	S	R	R	R	S
CS22491	CS22621	R	R	R	R	R
Ct-1	CS22639	R	R	R	R	R
Cvi-0	CS22614	R	S	S	I I	S
Eden-1	CS22572	R	R	_	R	R
Eden-2	CS22573	R	I .	S	R	R
Edi-0	CS22657	S	_	S	S	S
Ei-2	CS22616	S	_	_	_	_
Est-1	CS22629	R	R	_	R	R
Fab-2	CS22576	R	R	R	R	R
Fab-4	CS22577	R	R	_	R	S
Fei-0	CS22645	R	_	S	R	S
Ga-0	CS22634	R	I	S	S	S
Got-22	CS22609	R	S	R	R	S
Got-7	CS22608	R	S	R	R	S
Gu-0	CS22617	_	S	S	S	R
Gy-0	CS22631	R	S	S	R	R
HR-10	CS22597	R	R	R	R	R
HR-5	CS22596	S	R	R	S	R
Kas-1	CS22638	R	R	S	R	I
Kin-0	CS22654	R	R	R	S	S
Knox-10	CS22566	R	S	S	S	S
Knox-18	CS22567	R	S	S	R	S
Kondara	CS22651	R	R	R	R	S
Kz-1	CS22606	S	S	S	S	S
Kz-9	CS22607	R	R	S	_	S
Ler-1	CS22618	R	R	R	R	R
LL-0	CS22650	R	S	R	R	S
Lov-1	CS22574	R	R	_	S	S
Lov-5	CS22575	R	R	S	S	_
Lp2-2	CS22594	R	S	S	R	R
Lp2-6	CS22595	R	S	_	S	R
Lz-0	CS22615	R	R	S	R	R
<i>M</i> _r -0	CS22640	R	R	_	R	R
Mrk-0	CS22635	R	I	S	S	S
Ms-0	CS22655	R	I	S	R	R
Mt-0	CS22642	R	R	R	S	R
Mz-0	CS22636	R	R	R	R	R
Nd-0	CS22619	R	R	S	R	S
NFA-10	CS22599	R	S	R	S	S
NFA-8	CS22598	S	S	S	S	R
Nok-3	CS22643	R	1	S	S	S
Omo2-1	CS22584	S	R	R	S	S
Omo2-3	CS22585	R	R	R	S	S
Oy-0	CS22658	R	S	S	S	R
Pna-10	CS22571	R	S	S	I	S

PNAS PNAS

PNAS PNAS

Accession	Stock number	Emco5	Emwa1	Emoy2	Hiks1	Noco2
Pna-17	CS22570	R	S	S	R	R
Pro-0	CS22649	S	S	R	S	S
Pu2-23	CS22593	1	R	R	R	R
Pu2-7	CS22592	R	_	_	_	_
Ra-0	CS22632	S	S	S	I	I
Ren-1	CS22610	R	R	R	R	R
Ren-11	CS22611	R	R	R	R	R
Rmx-A02	CS22568	_	_	_	_	_
Rmx-A180	CS22569	R	S	S	S	S
RRS-10	CS22565	R	R	S	S	S
RRS-7	CS22564	R	R	R	R	R
Se-0	CS22646	R	S	R	_	S
Shahdara	CS22652	R	R	_	S	R
Sorbo	CS22653	R	R	S	R	R
Spr1-2	CS22582	R	R	_	S	R
Spr1-6	CS22583	R	R	R	R	_
Sq-1	CS22600	1	R	R	S	R
Sq-8	CS22601	I.	S	S	R	S
Tamm-2	CS22604	I.	_	_	R	R
Tamm-27	CS22605	I.	S	_	R	S
Ts-1	CS22647	R	S	I	R	S
Ts-5	CS22648	R	R	R	S	S
Tsu-1	CS22641	S	S	S	R	S
Ull2-3	CS22587	R	S	S	S	R
Ull2-5	CS22586	R	S	—	S	R
Uod-1	CS22612	R	R	S	S	S
Uod-7	CS22613	R	R	R	R	R
Van-0	CS22627	S	R	R	R	R
Var2-1	CS22580	1	R	_	_	S
Var2-6	CS22581	R	R	_	R	R
Wa-1	CS22644	R	R	R	R	R
Wei-0	CS22622	S	S	S	S	S
Ws-0	CS22623	S	S	R	R	R
Ws-2	CS22659	R	S	S	S	R
Wt-5	CS22637	R	R	R	R	R
Yo-0	CS22624	R	S	_	S	S
Zdr-1	CS22588	R	R	R	R	R
Zdr-6	CS22589	R	R	R	R	R
-						

Host disease phenotype on challenge with *H. arabidopsis* isolate

Interaction scoring: R, incompatible; S, compatible; I, intermediate; —, nonassessed. The experiment was repeated four times on five seedlings per genotype. Disease was scored specifically on first leaves.

 Table S3.
 Summary of survey of natural variation in disease resistance interactions between 96 accessions of A. thaliana and five isolates of H. arabidopsidis ex parasitica

Phenotypes	Total		Emco5		Emwa1		Emoy2		Hiks1	Noco2	
R	265 (60.1%)	70	(74.5%)	53	(58.9%)	41	(52.6%)	52	(59.1%)	49	(53.8%)
S	156 (35.4%)	16	(17%)	32	(35.5%)	36	(46.2%)	33	(37.5%)	39	(42.9%)
I	20 (4.6%)	8	(8.5%)	5	(5.5%)	1	(1.2%)	3	(3.4%)	3	(3.3%)
_	31	2		6		18		8		5	

The summary scores of the occurrence of resistance, susceptibility, intermediate interaction, and nonassessed (n.a.) for each isolate and in total are presented with frequencies in percents between parentheses.

	Mapping intervals: resistance/recognition specificity		Bay-0/E1										Gy-0/N2							Bay-0/C2					Kon/E1							Shah/ E5-E1-C2								
	source of		Sorbo/E1										Ga-0/E5												Nok-3/E5															
R to	2	Noco2	*				++	*	*	*				I		+		*	*		*					*				*			++					*	*	
ation with	<i>Hpa</i> isolate	Emco5		++		+	• ++							÷	*				*			+	*		4	ĸ	*	I	I	I	I			Ι	I	I		I	I	
Associ	t t	Emwa1			+		*		*					I		++	*	*	*				*		-	ĸ		*	*		*			*	÷	*				
	Distance to	candidate (kbps)		0			0	10	2 0							17	m		0						c	0 4	n c	•					6	12	14					
	Associated	candidate gene	~	At1g56540 (TNL)	- - -	~	CW9 (RPP8-like)	At1060320 (TNI)	At1a61060 (RPP8-like)	,			·	~- ſ	~ ~	At3q43770 (TNL)	SALM	ć	RPP13 (CNL)		ذ	ć	۲.			At4g143/0 (INL) ^+^~1^610 (CNL)	At4014610 (CNL)	2	~	2	ć		RPP5a (TNL)	RPP5 cluster (TNL)	RPP5 cluster (TNL)	ذ		<i>د</i>	۲.	
		Linked marker	nga128								t12p18ind8-8	f5i1449495	k11j14ind16-16							msat3.21				t16k521877	ciw6							msat4.15					t6k21ind15-15			msat4.18
		Physical location	20633251 21079875	21187559	21657491	21669047	21908330	22283402	22492711	23270329	23704640	24379374	12534916	146/2505 15020165	15526753	15679305	16379891	16977709	17141912	17256068	17277073	18180258	18379314	18470089	7892620	8280/64 8275566	8380918	8646352	9058122	9277112	9314687	9360000	9478782	9577993	9579976	9680252	9965867	10275867	10977375	64916611
Ŧ		Chr, MRC	I, MRC-B										III, MRC-F												IV, MRC-H															
datase			(¥)										(B)											1	Ũ															

Table S4. Association mapping of resistance loci against *H. arabidopsidis* in mapping intervals overlapping with MRCs using a set of 96 *A. thaliana* accessions with the 2010 genotype

S A Z d

rvals: mition specificity												Sorbo/N2																
Mapping inte		Ag-0/E1-N2																										
SOLITCE		sBr-0/E1																										
2	Noco2		#	*	*		*		*												*	*				*		
<i>pa</i> isolate	Emco5		*						*	*	*		*	*	*	+			*		++		++		*	*	*	
H	Emwa1					*		*		*					+			*		*	*		*					
Distance to	candidate (kbps)					m			80				10		10			14		0	11	0	12				m	
Associated	candidate gene		~	~	د.	At5g39365 (DEFL)	<i>c</i> .	<i>د</i> .	GLIP1	د.	<i>~</i>		BKI1	ذ	RPP8 (CNL)	ذ		At5g44345 (RPP8-like)	<i>.</i>	At5g44870 (TNL)	RPS4 (TNL cluster)	At5g45510 (Probable R)	At5g45770 (RLP)		~	~	At5g48595 (DEFL)	
	Linked marker	aths0191										ciw9					mln1ind8-8							nga129				mq158636
	Physical location	15021915	15065018	15171770	15456895	15767914	15871814	16368204	16445959	16565622	16869955	17061229	17171242	17242293	17469049	17692511	17805928	17864866	18065657	18134361	18364871	18464494	18568481	19007000	19123616	19614150	19715720	19218085
	Chr, MRC	V, MRC-J																										
		<u>í</u>																										

on. Locations of associated markers correspond to the position of the first nucleotide on the sequenced fragment. All significant associations at the 5% limit within the surveyed genomic regions are reported following their significance level. Recognition specificities correspond to isolates: E1, Emwa1; E5, Emco5; N2, Noco2; C2, Cala2; 1H, Hiks1. Candidate genes: TNL, TIR-NLR; CNL, CC-NLR; RLK, receptor-like kinase; RLP, receptor-like genes: TNL, TIR-NLR; CNL, CC-NLR; RLK, receptor-like kinase; RLP, receptor-like genes: TNL, TIR-NLR; CNL, CC-NLR; RLK, receptor-like kinase; RLP, receptor-like protein; DEFL, defensin-like genee; GLP1, GDSL-lipase; BK11, brassinosteroid-insensitive11-interacting 1; SALM, S-adenosyl-L-methionine carboxyl methyltransferase. In bold are the loci closest to known RPP genes when located within a distance allowing linkage disequilibrium (LD) mapping. ${}^{*}P < 0.05.$ ${}^{\dagger}P < 0.01.$ ${}^{\ddagger}P < 0.001.$

Nemri et al. www.pnas.org/cgi/content/short/0913160107

Table S4. Cont.

PNAS PNAS

14 of 15

Table S5. Identification of candidate RPP loci from GWAM of resistance to five isolates of H. arabido	sidis
---	-------

Chromosome MRC	I-B	III-F	IV-H	V-J	Total
Left border	nga128 (20633251)	<i>k11j14ind16-16</i> (12564916)	ciw6 (7892620)	aths0191 (15021915)	
Right border	f5i1449495 (24379373)	t16k521877 (18470089)	t22a6ind10-10 (12603189)	mql58 636 (19218084)	
Size (Mbp)	3.72	5.97	4.7	4.78	19.17
Number of SNPs	8,734	10,783	11,280	11,054	41,851
Number of SNPs > 0.004	201 (2.3%)	207 (1.91%)	119 (1.05%)	181 (1.63%)	672 (1.6%)
>0.004 and <5 kbp	16	20	7	13	56
apart					
Primary candidates	4	3	2	4	13
	TNL (At1g57850)	RPP13 (At3g46530)	RPP5 (At4g16860)	TNL (At5g40920)	
	CNL (At1g58390)	EDS1 (At3g48080)	ACD6 (At4g14400)	CNL (At5g43740)	
	TNL (At1g61100)	CNL (At3g46710)		TNL (At5g46270)	
	TNL (At1g63740)			CNL (At5g47260)	
Secondary candidates	4	2	1	1	8
	DEFL (At1g58055)	DEFL (At3g42557)	IAA1 (AXR5, At4g14560)	Avirulence-responsive (At5g39730)	
	RLK (At1g26950)	RLK (At3g47110)			
	RbohD-interactor	_			
	(At1g64060)				
	DEFL (At1g64195)				
Unknown	8	15	4	8	34

The table reports loci where very strong and consistent associations were found on four genomic regions corresponding to MRC and quantitative trait loci (QTL) intervals with their corresponding left and right border markers (physical position in parentheses). The total number of tested SNPs contained in the intervals is given. Loci were first selected on the basis of strength of association (importance > 0.004) and further on the basis of consistency of association (i.e., SNPs showing association with resistance to at least two *Hpa* isolates occurred within a 5-kbp window). Primary candidates correspond to NLR encoding genes (TNL, TIR-NLR; CNL, CC-NLR). Secondary candidates include various homologs of defense-related genes or pathogen-induced genes. Loci corresponding to yet nonannotated genes or genes not clearly linked to defense were grouped as unknown.

VAS PNAS