



## Supporting Online Material for

### **Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within *A. thaliana***

David Bikard, Dhaval Patel, Claire Le Metté, Veronica Giorgi, Christine Camilleri, Malcolm Bennett, Olivier Loudet\*

\*To whom correspondence should be addressed. E-mail: [loudet@versailles.inra.fr](mailto:loudet@versailles.inra.fr)

Published 30 January 2009, *Science* **323**, 623 (2009)  
DOI: 10.1126/science.1165917

#### **This PDF file includes:**

Materials and Methods  
Figs. S1 to S3  
Table S1  
References

## Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities within *A. thaliana*

David Bikard<sup>1</sup>, Dhaval Patel<sup>2</sup>, Claire Le Metté<sup>1</sup>, Veronica Giorgi<sup>1</sup>, Christine Camilleri<sup>1</sup>, Malcolm Bennett<sup>2</sup> & Olivier Loudet<sup>1</sup>

<sup>1</sup> INRA – Genetics and Plant Breeding SGAP UR254 – F-78026 Versailles, France

<sup>2</sup> Division of Plant Sciences, Centre for Plant Integrative Biology, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, UK LE12 5RD

---

### Supporting Online Material:

#### Material and Methods

##### - Plant Material.

The Cvi x Col RIL set was described earlier (*S1*) and can be obtained (along with all genotypic and phenotypic information) from the VNAT website at <http://dbsgap.versailles.inra.fr/vnat/>. 367 RILs genotyped at 90 markers have been analysed thanks to the GGT software ([http://www.plantbreeding.wur.nl/UK/software\\_ggt.html](http://www.plantbreeding.wur.nl/UK/software_ggt.html)) taking advantage of the LD tool with the ‘-10log(p)’ statistics to detect unlinked pairs of markers not segregating independently from each other. A F2 population from the same cross was tested (500 individuals) to recapitulate allelic segregation between LD1.1 and LD1.5.

Many of the F6 RILs are still individually segregating for one specific region. We took advantage of this to generate nearly isogenic lines in the form of Heterogeneous Inbred Families (HIF) as described (*S2*). In this case, we were careful to choose RILs in the appropriate reciprocal genotypic combination to be able to reproduce the incompatibility. RIL140 is heterozygous for LD1.1 region, while fixed Cvi at LD1.5. Reciprocally, RIL016 is heterozygous at LD1.5, while fixed Col at LD1.1. Hence, descendants from these lines (respectively families HIF140 and HIF016) confirm the interaction by not segregating for all genotypes at their respective heterozygous region. RIL016 also confirms the weak-root phenotype when heterozygous individuals from a progeny testing are phenotyped in vitro.

Mutants in *HPA1* gene include *emb2196*, a T-DNA-insertion line (*S3*), *hpa1*, an EMS (Ethyl Methane Sulfonate) point-mutation allele (*S4*) and SALK\_089516, a T-DNA-insertion line from the Salk Institute collection and obtained from NASC. All mutants are in the Col background.

The *Ler/Cvi* RIL set was described before (*S5*) as well as the nearly-isogenic lines (NIL) set derived from it (*S6*), including NIL5-4 used to confirm LD1 segregation in the *Ler* background.

Accessions (lines derived from *A. thaliana* wild populations) were screened at the sequence and transcript level (RT-PCR) to analyze the species-wide divergence of the *HPA* genes and build groups of accessions with similar combinations of alleles. We used essentially a core-collection of 24 accessions defined earlier to represent an optimal set to explore the species’ known variation (*S7*), plus a few parental accessions from crosses we wanted to check incompatibility in (Col-0, *Ler*, Bay-0, Ber, Lip-0, Eden-1). See table S1 for the complete list of 30, their origin and stock references (<http://dbsgap.versailles.inra.fr/vnat/>).

For the quantitative complementation test, we crossed heterozygous individuals from HIF016 (segregating at LD1.5) to either *emb2196* or *hpa1* mutant heterozygous individuals. Crosses were performed with each genotype both as male and female parent to control for and avoid maternal effects. In the progeny of these crosses, individuals were genotyped and phenotyped in vitro (see below) to record information on the expected four allelic classes. Parental lines, Col-0 as well as *hpa1* homozygous mutant plants were included as controls. Significantly different means and interaction effects were tested by ANOVA.

- Fine-mapping.

HIF016 (respectively, HIF140) plants cannot become homozygous for the Cvi (Col) allele at LD1.5 (LD1.1). 1.800 (1.900) individuals descending from HIF016 (HIF140) heterozygous plants were screened with markers surrounding the segregating interval LD1.5 (LD1.1) to identify recombinants within the candidate region harbouring a Cvi (Col) genotype at one of the marker. These plants were then individually genotyped with additional markers across the interval to identify the limits of the Cvi (Col) region and exclude it from the candidate interval. Markers used were microsatellite or indel markers, then CAPS when previous types of markers were exhausted and finally direct sequencing of the SNPs to precisely localize recombination breakpoints.

- Plant growth and phenotyping.

Plants were grown in vitro on classical media, as described (S2), including 4 days of seed stratification and then 12 days of vertical growth before the plates were scanned for root length measurement using Optimas or mounted and observed under the microscope for mature cortical cell length measurements. Cell production rate was estimated by dividing root growth rate by cell length.

The biochemical complementation of the 'weak-root' phenotype with histidine (His) was obtained in vitro by adding increasing concentrations of sterile His in the media after autoclaving and directly sowing the plants (segregating HIF016 progeny) on this media. On soil, the biochemical complementation of the embryo lethality phenotype was obtained by watering flowering plants (heterozygous HIF016 and HIF140 individuals) every other day with 1 mM His; seeds were then harvested to test for the segregation of the three genotypic classes in vitro on His-media.

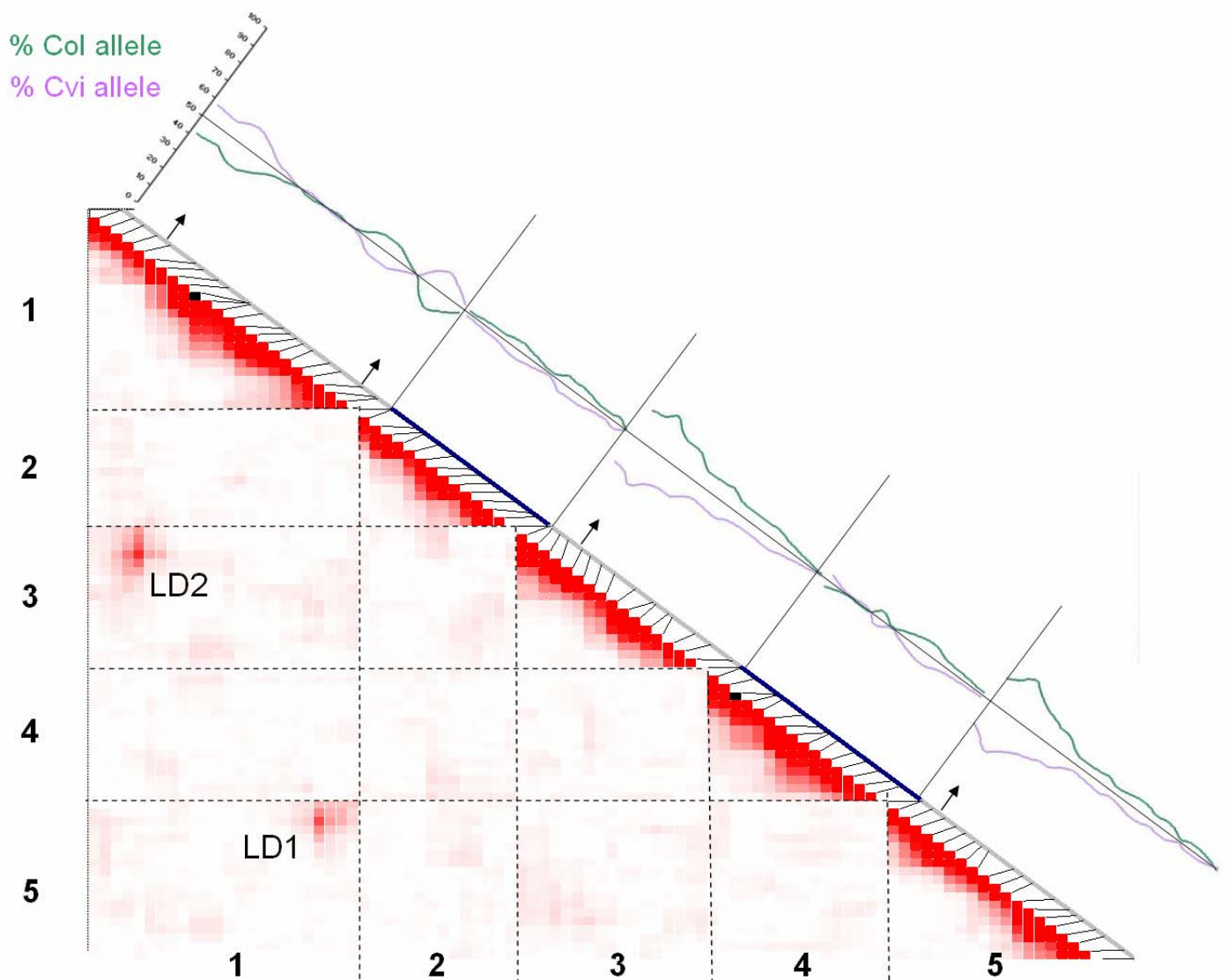
- Paralogs molecular analysis and genotyping.

Specific sequencing of the *HPA1* (At5g10330) coding sequence was performed with primers SEQ3bF (AGAAACCCCGATTTCAGCTT) and SEQ503246bR (TTTTCGTCGCACCACATTT). Specific sequencing of the *HPA2* (At1g71920) coding sequence was performed with primers SEQ3bF and SEQ1R (TGTTTTGTTTCTCACTTTGAATGAA). Sequencing the *HPA1* potential deletions was performed by amplifying specifically over the locus with the primers SEQ503246F (TCCAGGAAGGGTCCTGAGAT) and SEQ503254R (GTCTTTGCCAAATCCAGGAG) in Cvi; SEQ503246F and SEQ503252R (GCAGATGCAATCTCGTTTGA) in Ita-0, Can-0, Mt-0, Ber, Lip-0 and Eden-1.

For Col (as well as for most accessions according to *HPA1* / *HPA2* SNPs detected in the coding sequence), *HPA1* and *HPA2* mRNA can be distinguished with non-specific primers HPA-RT-3F (GAAAAACCCAAATGCATATTCC) and HPA-RT-2R (ACGAACCATCACCCCATTT) and digestion of the RT-PCR (or semi-quantitative RT-PCR) product with enzyme RsaI which cuts only the *HPA1* mRNA. Other accessions not bearing this SNP between the two paralogs have been analyzed with the same RT-PCR

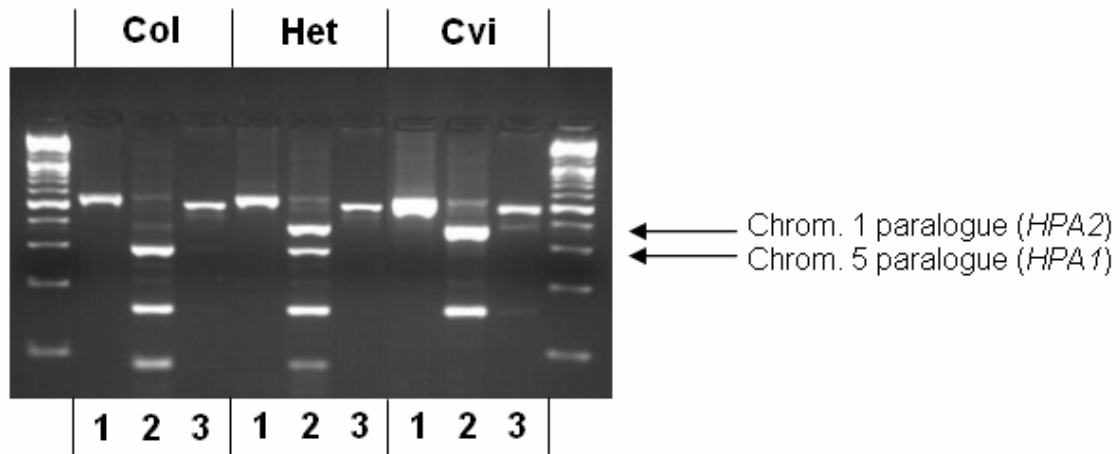
primers and with enzymes HpaII cutting only *HPA1* mRNA in B1-1; MaeIII cutting only *HPA2* mRNA in Can-0 and Kn-0.

Segregation of the genotypes in HIFs and within specific F2 or RIL populations was checked with polymorphic microsatellite markers at LD1.1 and LD1.5 loci, from the MSAT database at <http://www.inra.fr/vast>.



**Supporting Figure S1:** Trans-chromosomal linkage disequilibrium (pseudo-LD) and segregation distortion along chromosomes within Cvi-0 x Col-0 RIL set.

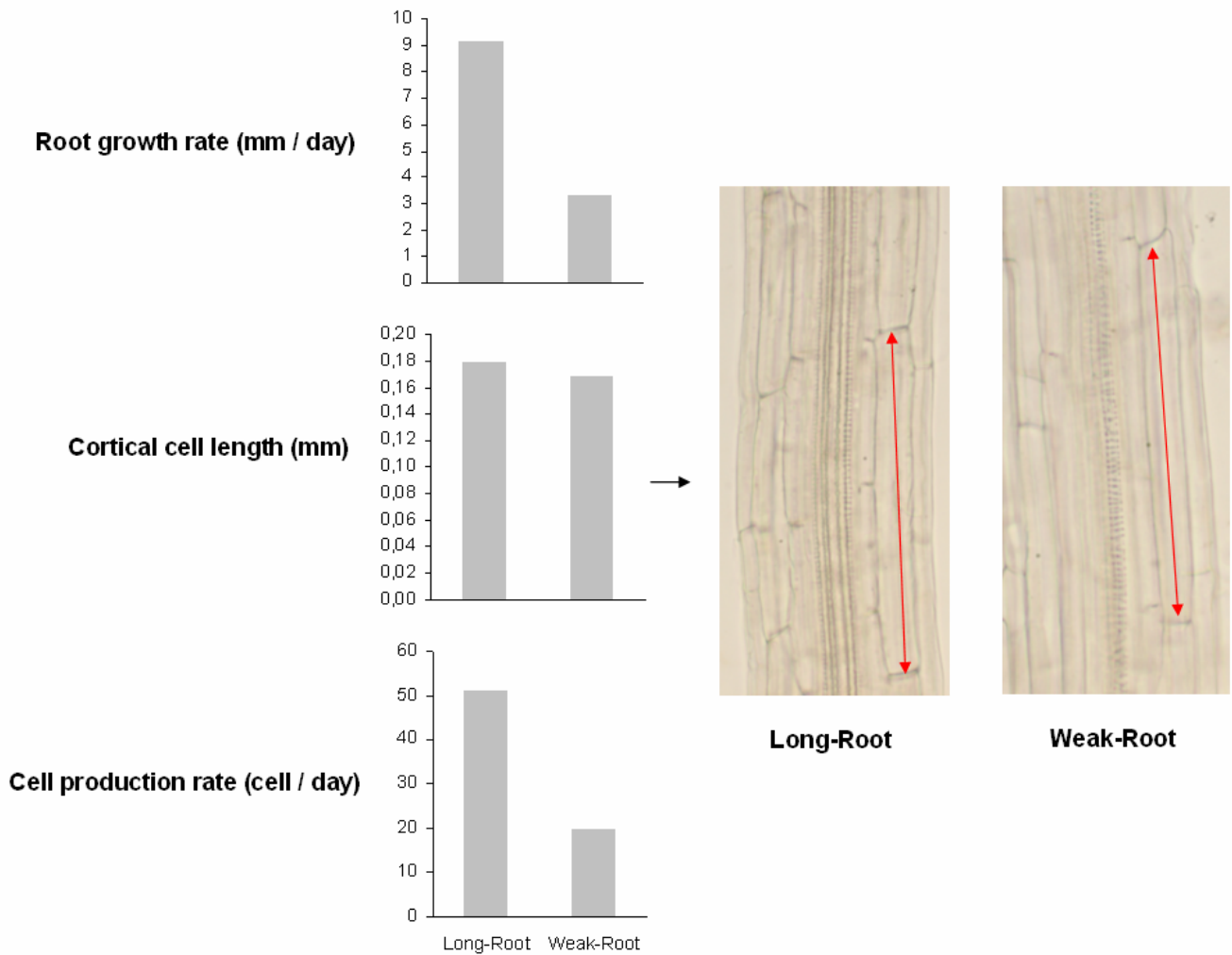
Below the diagonal we show a heatmap obtained from GGT32 software representing the statistics of LD ( $-10\log(p)$ ) where darker red spots indicate high LD between any pairs of markers. Two significant LD between physically unlinked genomic regions are indicated (LD1 and LD2) and represent pairs of loci which do not segregate independently from each other. Curves above the diagonal represent the respective percentage of each allele (green: Col; purple: Cvi) along chromosomes, among 367 RIL (excluding heterozygous and missing data). Significant distortions potentially explained by LD1 or LD2 are indicated by arrows along the curve.



1 : Non-digested RT-PCR  
 2 : Rsa-I digested RT-PCR  
 3 : RT-PCR Control (APT)

**Supporting Figure S2:** Expression of the respective *HPA* paralogs in Col, Cvi and a fully heterozygous genotype (Het).

Results from semi-quantitative RT-PCR are shown for each genotype as (1) undigested RT-PCR product and (2) after digestion with *RsaI* enzyme, which shows a specific digestion site within *HPA1* (chrom. 5) paralog's mRNA (as well as another digestion site common to all alleles and responsible for the band at ~150bp). APT gene expression is shown as a control (3). Identical results were obtained from shoot- or root-extracted mRNAs.



**Supporting Figure S3:** The weak-root phenotype is explained by variation in cell production rate.

For each variable, two genotypes were evaluated: one with normal root elongation (long-root phenotype) corresponding to plants homozygous Col at LD1.5 and one with a weak-root phenotype corresponding to plants heterozygous at LD1.5. Average root elongation per day (growth rate) represents a combination of cell production rate (number of cells produced per day) and cell elongation (measured as the mature length of the cortical cells) factors, of which only the first variable was found to vary between the genotypes and which explains the observed growth rate difference. Typical examples of magnified mature cortical cells from cleared roots of each genotype are shown to the right, with red lines highlighting cell length.

Supp. Table S1

Accession	Country of origin	Versailles Ref.#	Core-Collection	At1g71920		At5g10330		Test-cross x Col	Test-cross x Cvi
				Expr <sup>o</sup>	STOP	Expr <sup>o</sup>	Deletions		
Col-0	Poland	186AV	Reference	∅		+			GI (RIL, F2)
Stw-0	Russia	92AV	CC24	∅		+			
Ler	Poland	213AV		+	Ler	+		OK (RIL)	GI (RIL)
Shahdara	Tajikistan	236AV	CC8	+	Ler	+		OK (RIL)	GI (F2)
Bay-0	Germany	41AV		+	Ler	+			GI (F2)
Oy-0	Norway	224AV	CC8	+	Ler	+		OK (RIL)	GI (F2)
Ge-0	Switzerland	101AV	CC16	+	Ler	+		OK (RIL)	
Mh-1	Poland	215AV	CC16	+	Ler	+			
Kn-0	Lithuania	70AV	CC24	+	Ler	+			
Edi-0	United Kingdom	83AV	CC24	+	Ler	+			
Gre-0	USA	200AV	CC24	+	Ler	+			
Akita	Japan	252AV	CC24	+	Ler	+			
Tsu-0	Japan	91AV	CC24	∅	Ler	+		OK (RIL)	GI (F2)
N13	Russia	266AV	CC16	∅	Ler	+			
Bur-0	Ireland	172AV	CC8	+		+		OK (RIL)	OK (F2)
Bih-1	Czechoslovakia	180AV	CC8	+		+		OK (RIL)	OK (F2)
Ct-1	Italy	162AV	CC8	+		+		OK (RIL)	OK (F2)
St-0	Sweden	62AV	CC16	+		+			
Bl-1	Italy	42AV	CC24	+		+		OK (F2)	
Sakata	Japan	257AV	CC24	+		+		OK (F2)	
Aic-0	Spain	178AV	CC16	+		+		OK (F2)	
Pyl-1	France	8AV	CC16	+		+			
Cvi-0	Cape Verde Isl.	166AV	CC8	+		[∅]	I	GI (RIL, F2)	
Ita-0	Morocco	157AV	CC8	+		[∅]	II	GI (RIL, F2)	
Lip-0	Poland	63AV	CC48	+		[∅]	III	GI (F2)	
Ber	Denmark	299AV		+		[∅]	III	Cond. GI (F2)	
Can-0	Spain (Canary Isl.)	163AV	CC16	+	heterozyg. ?	[∅]	III	Cond. GI (F2)	
Mt-0	Libya	94AV	CC16	+	heterozyg. ?	[∅]	III	Cond. GI (F2)	
Eden-1	Sweden	368AV		+	heterozyg. ?	[∅]	III		
Jea	France	25AV	CC8	+		∅		Cond. GI (RIL, F2)	OK (F2)

see <http://dbsgap.versailles.inra.fr/vnat/> ∅ = no expression detected

GI = Genetic Incompatibility  
Cond. GI = Conditional Genetic Incompatibility

**Supporting Table S1:** Intra-specific natural variation at *HPA* genes.

For each accession, RT-PCR as well as sequencing results are reported for each *HPA* paralog according to functional polymorphisms known; incompatibilities (or compatibilities) are reported as tested by crosses to Col and/or Cvi in either RIL sets, F2 populations, or both. Note that a few accessions show signs of a potential third polymorphic location for an active copy of *HPA*, on the basis of (1) sequencing results, (2) allele combination at LD1.1 / LD1.5 and (3) conditional incompatibilities detected in some crosses and reported in this table.

Supplementary References:

- S1. M. Simon *et al.*, *Genetics* **178**, 2253 (2008).
- S2. O. Loudet, V. Gaudon, A. Trubuil, F. Daniel-Vedele, *Theor. Appl. Genet.* **110**, 742 (2005).
- S3. I. Tzafrir *et al.*, *Plant Physiol.* **135**, 1206 (2004).
- S4. X. Mo *et al.*, *Plant Physiol.* **141**, 1425 (2006).
- S5. C. Alonso-Blanco *et al.*, *Plant J.* **14**, 259 (1998).
- S6. J. J. Keurentjes *et al.*, *Genetics* **175**, 891 (2007).
- S7. H. I. McKhann *et al.*, *Plant J.* **38**, 193 (2004).