



HAL
open science

Genome analysis of poplar LRR-RLP gene clusters reveals RISP, a defense-related gene coding a candidate endogenous peptide elicitor

Benjamin Petre, Stéphane Hacquard, Sébastien Duplessis, Nicolas Rouhier

► To cite this version:

Benjamin Petre, Stéphane Hacquard, Sébastien Duplessis, Nicolas Rouhier. Genome analysis of poplar LRR-RLP gene clusters reveals RISP, a defense-related gene coding a candidate endogenous peptide elicitor. *Frontiers in Plant Science*, 2014, 5, 10.3389/fpls.2014.00111 . hal-02548502

HAL Id: hal-02548502

<https://hal.science/hal-02548502>

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Genome analysis of poplar LRR-RLP gene clusters reveals RISP, a defense-related gene coding a candidate endogenous peptide elicitor

Benjamin Petre^{1,2†}, Stéphane Hacquard^{1,2†}, Sébastien Duplessis^{1,2*} and Nicolas Rouhier^{1,2*}

¹ INRA, Interactions Arbres/Microorganismes, UMR 1136, Champenoux, France

² Université de Lorraine, Interactions Arbres/Microorganismes, UMR 1136, Vandoeuvre-lès-Nancy, France

Edited by:

Adi Avni, Tel Aviv University, Israel

Reviewed by:

Firoz Ahmed, New York University, USA

Jung-Gun Kim, Stanford University, USA

*Correspondence:

Sébastien Duplessis, INRA, Interactions Arbres/Microorganismes, UMR 1136 INRA Université de Lorraine, Centre INRA de Nancy, 54280 Champenoux, France
e-mail: duplessi@nancy.inra.fr;
Nicolas Rouhier, Faculté des Sciences, Université de Lorraine, Interactions Arbres/Microorganismes, UMR1136, Bd des Aiguillettes, 54500 Vandoeuvre-lès-Nancy, France
e-mail: nicolas.rouhier@univ-lorraine.fr

† Present address:

Benjamin Petre, The Sainsbury Laboratory, Norwich Research Park, Norwich, UK;
Stéphane Hacquard, Max Planck Institute for Plant Breeding Research, Cologne, Germany

In plants, cell-surface receptors control immunity and development through the recognition of extracellular ligands. Leucine-rich repeat receptor-like proteins (LRR-RLPs) constitute a large multigene family of cell-surface receptors. Although this family has been intensively studied, a limited number of ligands has been identified so far, mostly because methods used for their identification and characterization are complex and fastidious. In this study, we combined genome and transcriptome analyses to describe the *LRR-RLP* gene family in the model tree poplar (*Populus trichocarpa*). In total, 82 *LRR-RLP* genes have been identified in *P. trichocarpa* genome, among which 66 are organized in clusters of up to seven members. In these clusters, *LRR-RLP* genes are interspersed by orphan, poplar-specific genes encoding small proteins of unknown function (SPUFs). In particular, the nine largest clusters of *LRR-RLP* genes (47 *LRR-RLPs*) include 71 *SPUF* genes that account for 59% of the non-*LRR-RLP* gene content within these clusters. Forty-four *LRR-RLP* and 55 *SPUF* genes are expressed in poplar leaves, mostly at low levels, except for members of some clusters that show higher and sometimes coordinated expression levels. Notably, wounding of poplar leaves strongly induced the expression of a defense *SPUF* gene named *Rust-Induced Secreted protein (RISP)* that has been previously reported as a marker of poplar defense responses. Interestingly, we show that the *RISP*-associated *LRR-RLP* gene is highly expressed in poplar leaves and slightly induced by wounding. Both gene promoters share a highly conserved region of ~300 nucleotides. This led us to hypothesize that the corresponding pair of proteins could be involved in poplar immunity, possibly as a ligand/receptor couple. In conclusion, we speculate that some poplar *SPUFs*, such as *RISP*, represent candidate endogenous peptide ligands of the associated *LRR-RLPs* and we discuss how to investigate further this hypothesis.

Keywords: gene clustering, immunity, ligands, receptors, poplar, wounding

INTRODUCTION

Plants possess cell-surface receptors able to recognize extracellular molecules for subsequent intracellular signaling (Monaghan and Zipfel, 2012). Cell-surface receptors are organized into distinct families, which present a common architecture with a cytoplasmic part, a transmembrane domain and an extracellular part. Cytoplasmic domains allow distinguishing the two major classes of receptors: the receptor-like kinases (RLKs, with a cytoplasmic kinase domain) and the receptor-like proteins (RLPs, with a short cytoplasmic tail). In order to trigger signaling, RLKs use their own kinase domain, whereas RLPs likely require association with RLKs at the plasma membrane (Liebrand et al., 2013). A variety of extracellular domains, which directly interact with apoplast ligands, have been described; the most common being composed of leucine-rich repeats (LRRs) (Monaghan and Zipfel, 2012; Sun et al., 2013). RLKs and RLPs with extracellular LRRs are abbreviated LRR-RLKs and LRR-RLPs, respectively.

LRR-RLPs evolved as a multigene family in higher plants. For instance, more than 50 *LRR-RLP* genes have been identified

in the model plant *Arabidopsis thaliana* (Fritz-Laylin et al., 2005). Despite extensive genetic studies, the function of nearly all *A. thaliana* LRR-RLPs remains unknown (Wang and Fiers, 2010). The few plant LRR-RLPs characterized so far are able to perceive peptide ligands and trigger signaling cascades that control immunity or development. For example, tomato Cf-9 and Ve1 LRR-RLPs trigger plant immunity after the recognition of the fungal proteins Avr9 and Ave1, respectively (Jones et al., 1994; de Jonge et al., 2012). In contrast, *A. thaliana* CLV2 and TMM modulate plant development, respectively meristematic cell differentiation and leaf stomatal patterning (Jeong et al., 1999; Nadeau and Sack, 2002). CLV2 achieves its function through the recognition of endogenous peptides from the CLE family (Murphy et al., 2012).

Besides CLE peptides, several families of cell-surface receptor ligands are small endogenous peptides (~10–20 amino acids), which are often cleaved from larger precursor proteins of unknown function (~100–200 amino acids) and released into the apoplast where they function as messengers (Butenko et al.,

2009). Among these endogenous peptides, those that are able to induce immune responses are commonly referred to as endogenous peptide elicitors (Ryan and Pearce, 2003; Yamaguchi and Huffaker, 2011). For instance, the *A. thaliana* plant elicitor peptide (AtPep) family is constituted of small peptides (~23 amino acids) encoded by larger pro-AtPep precursor proteins of ~100 amino acids (Huffaker et al., 2006). Members of the AtPep family bind to PEPR cell-surface receptors from the LRR-RLK family to trigger immune responses (Krol et al., 2010).

Here, we report the genome-wide analysis of the *LRR-RLP* gene family in the model tree species poplar. We observed that *LRR-RLP* genes are frequently associated with genes coding orphan, species-specific small proteins of unknown function (SPUFs). Moreover, transcriptomic data highlighted a coordinated expression for a pair of physically associated *LRR-RLP* and *SPUF* genes in response to wounding, suggesting that the gene products might be functionally linked.

RESULTS

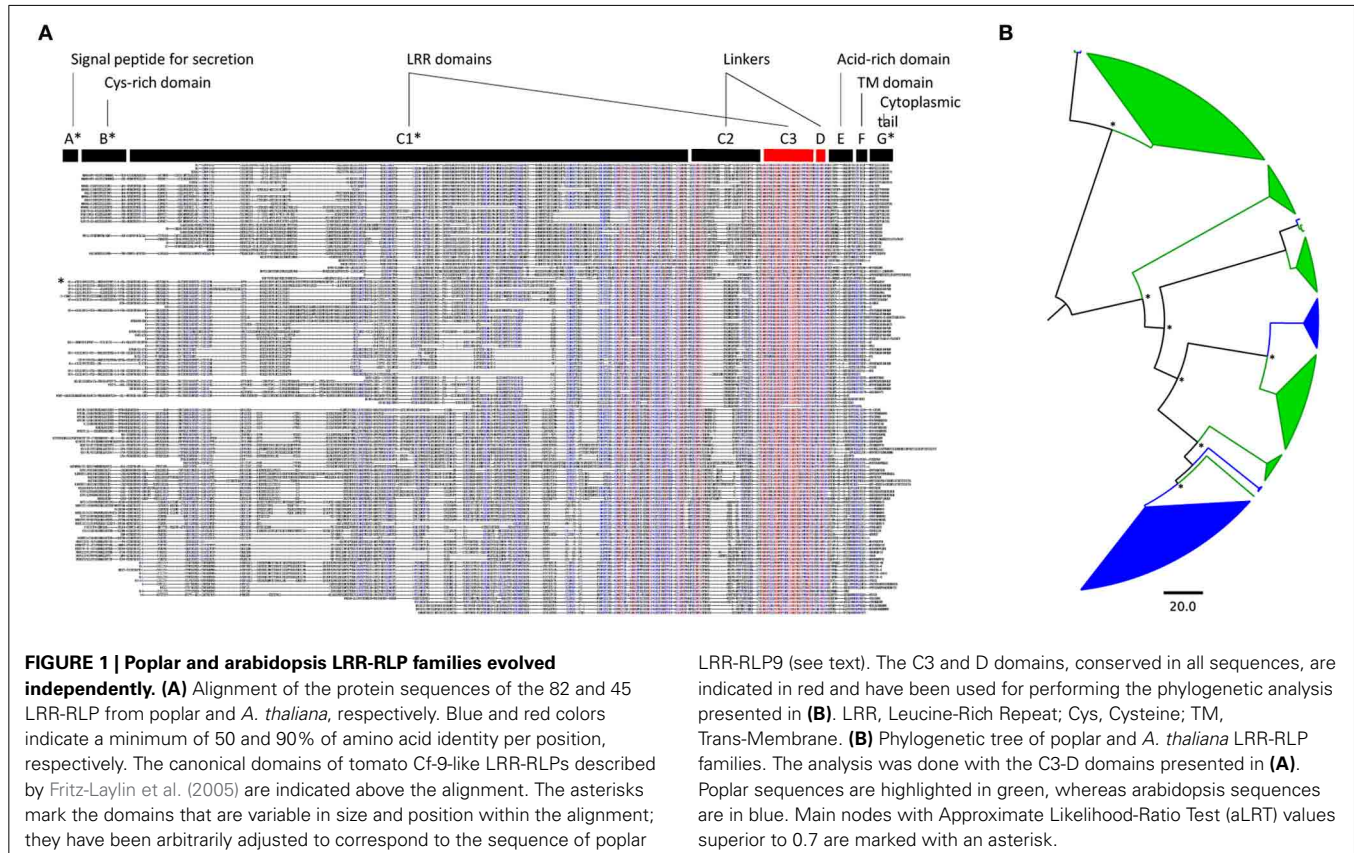
THE POPLAR *LRR-RLP* GENE FAMILY EVOLVED RECENTLY BY TANDEM DUPLICATION

A total of 82 *LRR-RLP* genes were identified in the poplar genome (Table 1; Table S1A). This value is in the range of those reported in *A. thaliana* and rice, 57 and 90, respectively (Fritz-Laylin et al., 2005; Wang et al., 2008). *LRR-RLPs* are composed of several domains numbered from A to G (Fritz-Laylin et al., 2005). Domains such as the LRR domain C1 or the linker domain C2

have a variable number of repeats. In contrast, the LRR domain C3 and the linker domain D are conserved and are thus suitable for sequence comparison and phylogenetic analyses (Figure 1A). A phylogenetic tree was inferred with the C3 and D domains of poplar and *A. thaliana* *LRR-RLPs* (Figure 1B). Sequences from

Table 1 | LRR-RLP and associated SPUF genes in *P. trichocarpa* and *A. thaliana*.

	<i>P. trichocarpa</i>	<i>A. thaliana</i>
<i>LRR-RLP</i> genes	82	45
<i>SPUF</i> genes	87	38
<i>LRR-RLP</i> -associated genes [avg. per <i>LRR-RLP</i> gene]	375 [4.75 ± 2.1]	597 [13.3 ± 3.7]
% of <i>SPUF</i> genes compared with the total no. of <i>LRR-RLP</i> -associated genes	23.2	6.4
% of <i>SPUF</i> genes in the whole predicted proteome	~20	~13
<i>LRR-RLP</i> genes clustered [no. of clusters]	66 [23]	29 [11]
SPUFs predicted secreted	10	10
SPUFs with no paralog	31	21
SPUFs with 1 or 2 paralogs	20	11
SPUFs with no homolog in other plants	34	29

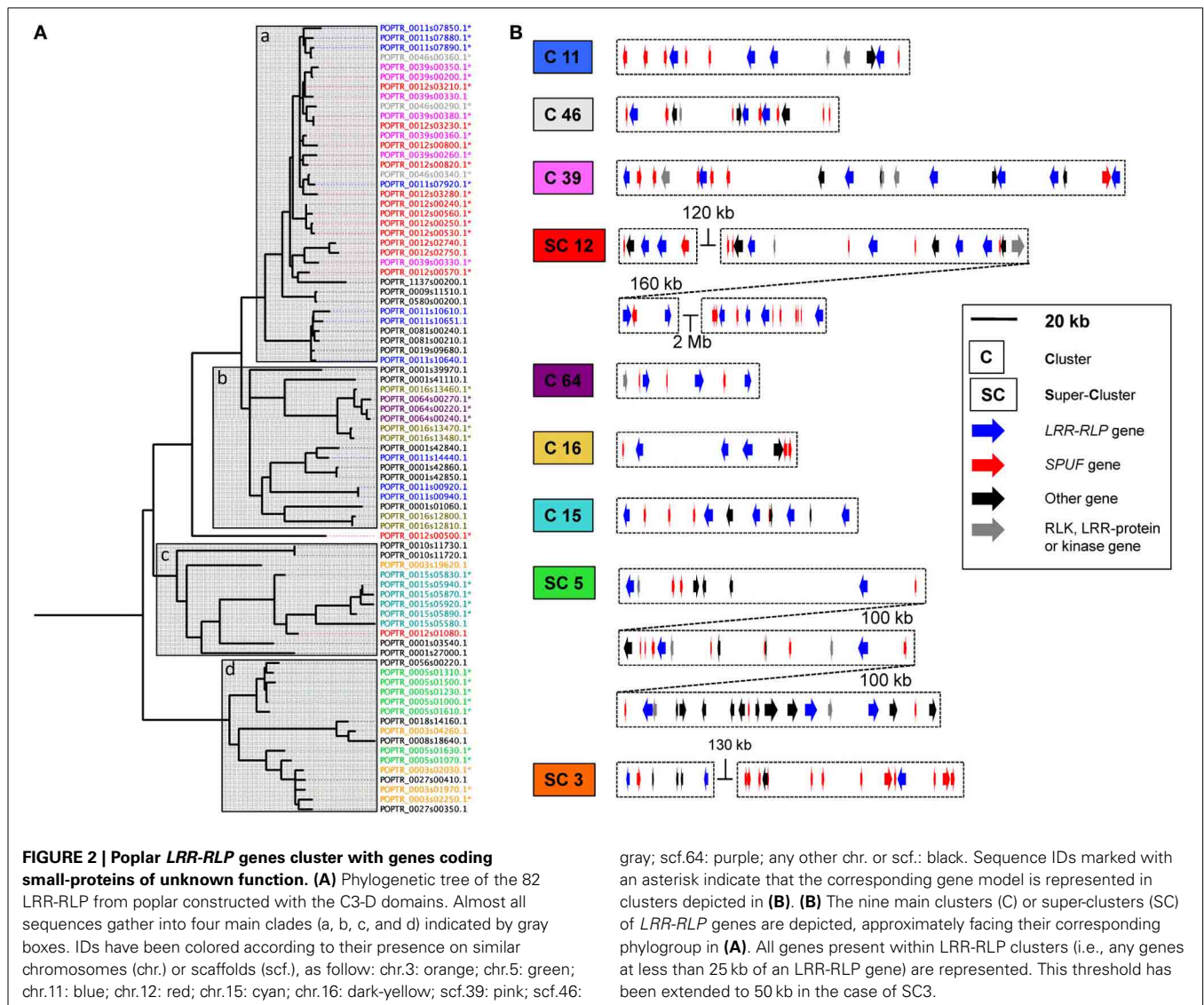


each species gather within a few well-separated clades, highlighting the strong divergence of the two families and their probable independent evolution in poplar and *A. thaliana*. A phylogenetic tree focusing only on poplar LRR-RLPs revealed that 81 of the 82 sequences group into four distinct clades termed a, b, c, and d (Figure 2A). LRR-RLP genomic organization correlates with phylogeny, since genes that cluster in the genome sequence gather in the same phylogenetic clade. For instance, the six LRR-RLPs from chromosome 15 are all grouped into the clade c, whereas the seven LRR-RLPs from chromosome 5 are found in the clade d. A more accurate analysis of gene positions revealed that 66 of the 82 LRR-RLPs are organized in 23 clusters or super-clusters (Table 1). A cluster is constituted by at least two genes within a 50 kb stretch, whereas a super-cluster refers to a group of at least two clusters separated by less than 2 Mb. The nine largest clusters or super-clusters gathering 47 LRR-RLPs (57% of the family) are depicted in Figure 2B.

POPLAR LRR-RLP GENES ARE CLOSELY ASSOCIATED WITH GENES ENCODING ORPHAN, SPECIES-SPECIFIC SMALL PROTEINS OF UNKNOWN FUNCTION (SPUFs)

We investigated carefully all LRR-RLP-associated genes (i.e., genes present within the 25 kb upstream or downstream of a given LRR-RLP gene) in the poplar genome sequence, and found a total of 375 LRR-RLP-associated genes (Table 1). Among these 375 genes, 87 (23% of all LRR-RLP-associated genes) encode small proteins (less than 200 amino acids) of unknown function that we abbreviated SPUF (Table 1). This value is close to the average percentage of SPUF (20%) found within the predicted proteome of *P. trichocarpa* (Table 1). However, within the nine biggest clusters of LRR-RLPs depicted in Figure 2, and without considering the LRR-RLPs, SPUF genes account for 59% of the gene content. Such an abundance of SPUF genes within poplar LRR-RLP clusters prompted us to further investigate them.

Using SPUF sequences as queries, we performed homology searches against the non-redundant protein database at



NCBI. This analysis revealed that 53 SPUFs are poplar-specific whereas the remaining 34 have low homology with proteins from other plant species (Table S1B). Due to the presence of partially conserved or truncated domains (e.g., NB-ARC, glycosyl hydrolase, . . .), a dozen of SPUFs have numerous paralogs and homologs in databases (Figure 3). To estimate SPUF diversity within the poplar genome, we also performed homology searches against the predicted proteome of *P. trichocarpa* and found that 30 of the 87 SPUFs have no paralog and 21 have only one or two paralogs (Figure 3; Table 1). Altogether, these results show that many SPUFs are orphan and poplar-specific genes.

Interestingly, some SPUF paralogs are also physically associated within LRR-RLP clusters, strongly suggesting that they could have been duplicated conjointly with LRR-RLP genes. For example, a family of four related SPUFs are all grouped within the super-cluster 5 (Figure 4). Another family of three closely related SPUFs is dispatched between cluster 15 (two members) and chromosome 12 (one member, nearby a non-clustered LRR-RLP). Both the SPUF and the LRR-RLP from chromosome 12 are the closest paralogs of SPUFs and LRR-RLPs from the cluster 15 (Figures 2, 4). In this case, the LRR-RLP/SPUF pairs from chromosomes 15 and 12 likely result from a recent event of duplication and transposition. This point is supported by the fact that chromosomes 12 and 15 are “twin chromosomes,” resulting from the recent whole genome duplication in poplar (Tuskan et al., 2006).

Noteworthy, 10 SPUFs possess an N-terminal signal peptide predicted to target the proteins to the secretory pathway (Table 1). The gene coding one of these proteins has been previously reported as the most induced gene in poplar leaves during resistance to the leaf rust fungus *Melampsora larici-populina* (Rinaldi et al., 2007). The corresponding protein, composed of 60 amino acids in its predicted mature form, was named Rust-Induced Secreted Protein (RISP). RISP has no known function and no homolog in other plants, and has been speculated to be a novel component of the poplar immune system.

Besides SPUF genes, 30 genes coding proteins with a restricted number of LRR-domains have been found in the vicinity of poplar LRR-RLP genes (in gray in Figure 2B; Table S1A). It is

tempting to consider them as reservoirs of diversity or relics of LRR-RLP gene evolution, although we cannot completely exclude that these are mis-annotated or pseudo-genes. Among other notable LRR-RLP-associated genes identified within RLP-LRR clusters are RLKs and serine-threonine or lectin kinases (Table S1A), which may eventually be part of the same signaling cascades.

SOME CLUSTERS OF LRR-RLP AND SPUF GENES ARE EXPRESSED IN POPLAR LEAVES

The expression of the 82 LRR-RLP genes and of 82 of the 87 SPUF genes was examined in mature poplar leaves using whole-genome poplar oligoarrays (Table S1C). Overall, 44 LRR-RLPs were expressed above the background threshold, but at low to medium levels (<4000) (Figure 5A). POPTR_0009s11510, the only LRR-RLP present on chromosome 9 (hereafter termed LRR-RLP9), is the only one that presented a slightly higher expression in poplar leaves (>4000) (Figure 5A). Among the SPUFs, 55 presented a detectable expression, and five of them showed a very high expression level (>10,000) (Figure 5B). Interestingly, some specific clusters gathering several pairs of LRR-RLP and SPUF genes showed medium to high expression levels in leaves. For example, the four LRR-RLPs and two SPUFs from cluster 11 are concomitantly expressed in leaves (Figure 5C). Similarly, two pairs of LRR-RLP/SPUF from the super-cluster 3 are co-expressed. Interestingly, the three SPUFs dispatched between chromosome 15 and 12 (highlighted previously, see Figure 4) were among the most highly expressed SPUF genes in leaves, along with the nearby LRR-RLP from chromosome 12 (Figure 5C).

To extend our survey of LRR-RLP and SPUF gene expression, we explored the PopGenIE portal (<http://popgenie.org/>), which compiles expression data from rust-infected leaves, mature leaves, leaf disks, seeds, stems and suspension cells from poplar. Expression data were available for 47 LRR-RLP and 36 SPUF genes. A large majority (42 LRR-RLPs and 34 SPUFs) was expressed at low levels and did not exhibit any particular expression profile or obvious regulation in the analyzed conditions (Figure S1). Nevertheless, five LRR-RLPs and two SPUFs showed interesting profiles. Indeed, the POPTR_0005s01490 and

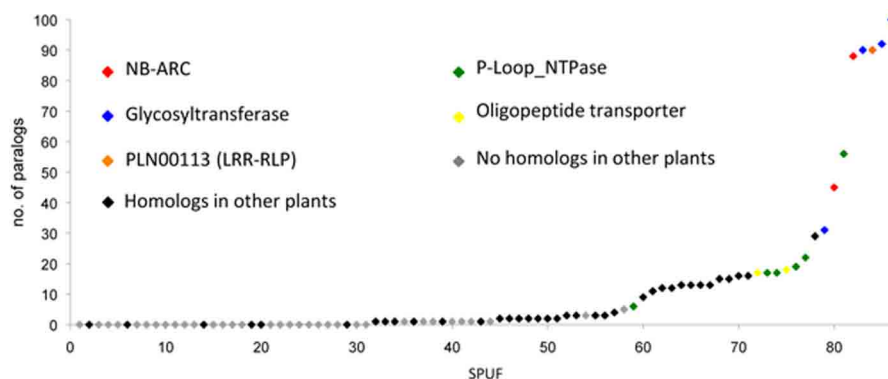
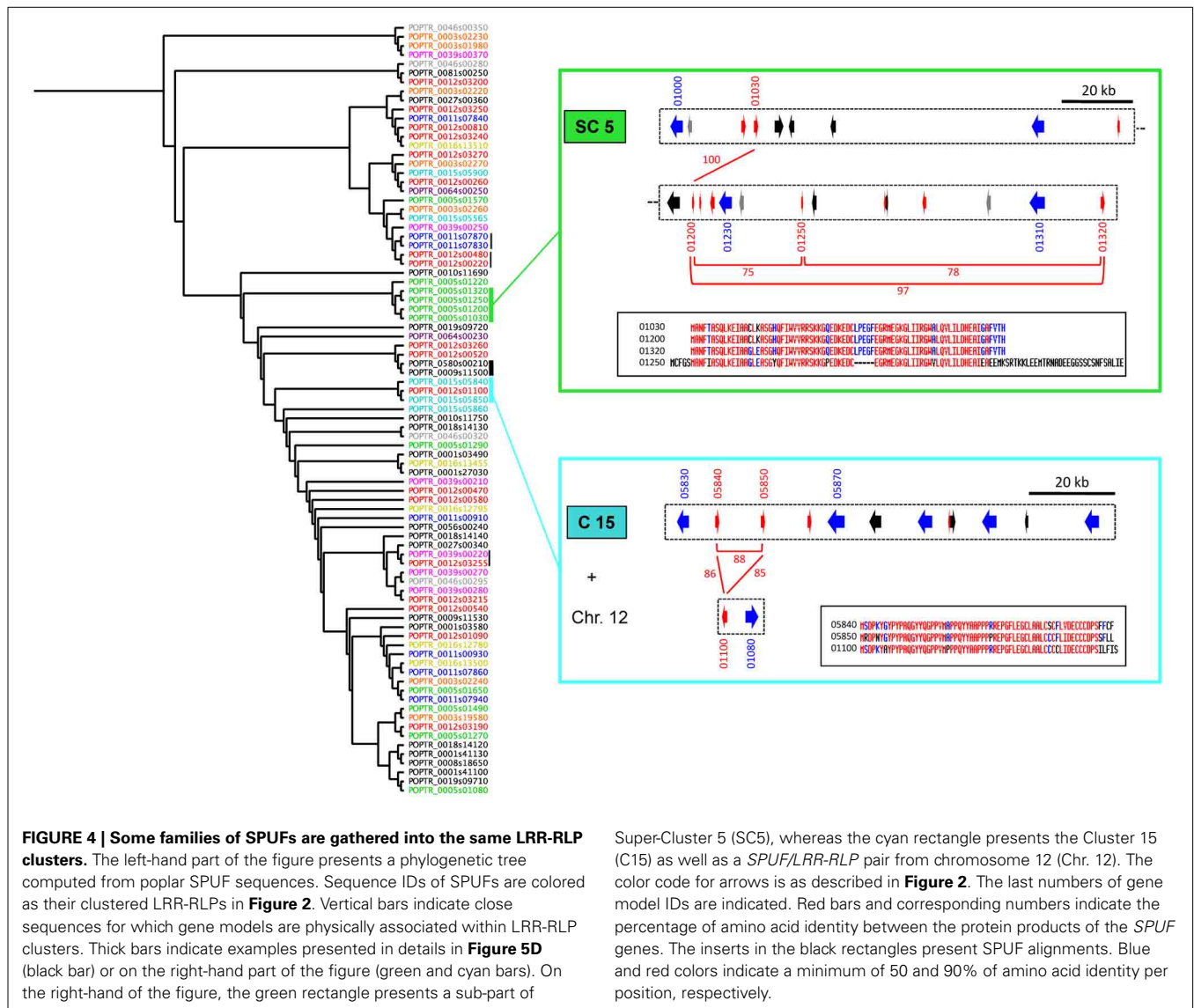


FIGURE 3 | Poplar SPUFs have limited paralogs and homologs in other plants. Poplar SPUF sequences were used for homology searches against the predicted poplar proteome on the Phytozome portal, as well as the

non-redundant protein database at the NCBI website. The dataset used, homology search details and homologs identified in other plants are detailed in Table S1B.

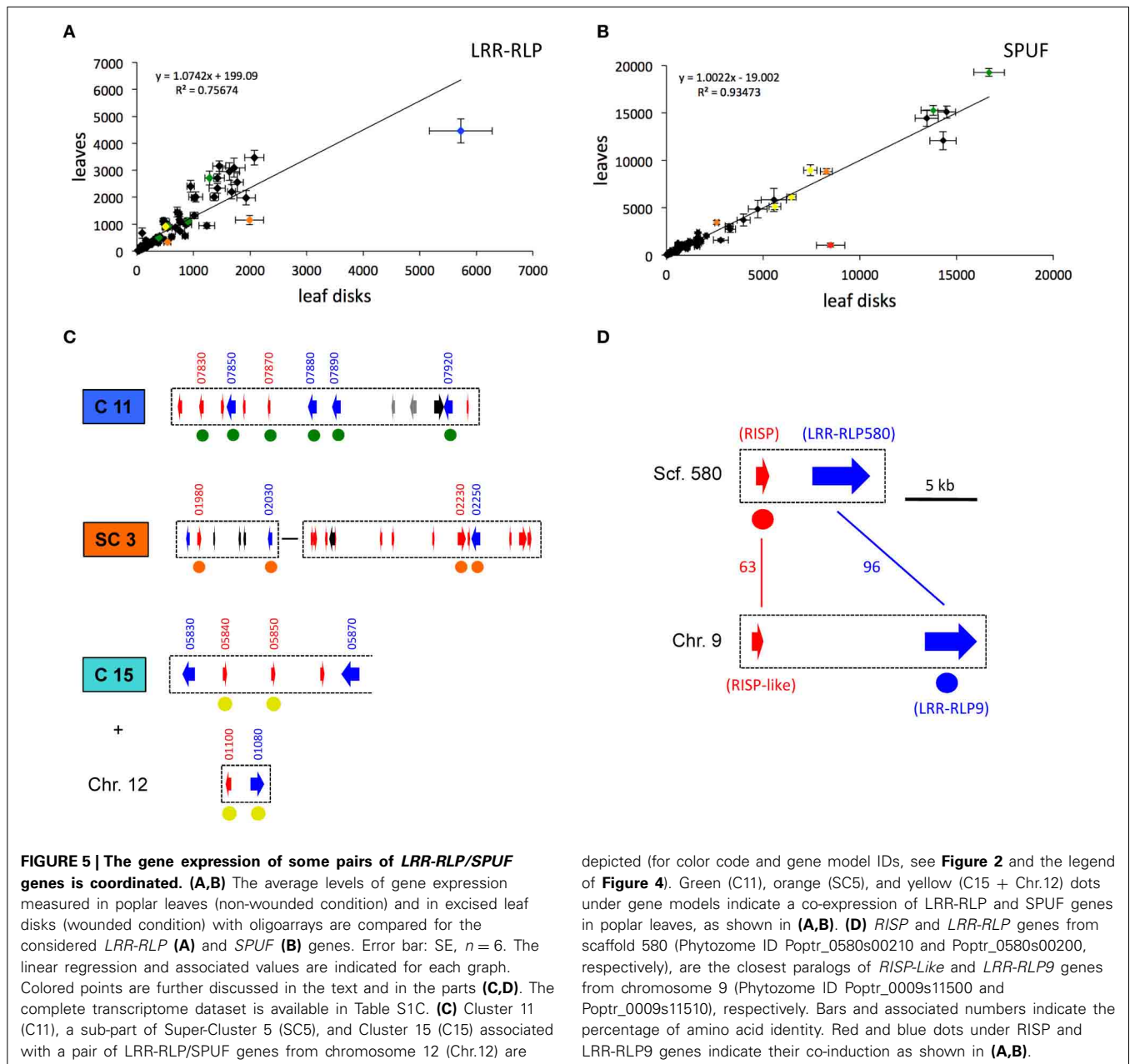


POPTR_0011s07870 *SPUF* genes presented a high expression in stems and rust-infected leaves, respectively, whereas the five *LRR-RLPs*, including the above-mentioned LRR-RLP9, presented a high expression during stress treatments (rust-infected leaves and/or leaf discs).

RISP IS STRONGLY INDUCED IN POPLAR LEAVES UPON WOUNDING

Most genes involved in immune responses (including cell-surface receptors and precursors of endogenous peptide elicitors) have an inducible expression in response to stresses like wounding or pathogen attack (Kemmerling et al., 2011; Yamaguchi and Huffaker, 2011). In order to identify poplar *LRR-RLP* and *SPUF* genes that likely participate to defense mechanisms, we compared transcript expression profiles between excised poplar leaf discs and entire leaves (i.e., wounded vs. non-wounded leaves) by using whole-genome poplar oligoarrays. As expected, the expression of genes classically associated with wounding such as wound-response proteins, chitinases, or protease inhibitors were induced

in leaf discs (Figure S2; Table S1C). The expression levels of all *SPUFs* remained very stable after wounding, except one, the *RISP* gene, which was strongly induced (~10 fold) (**Figure 5B**). This observation supports the above-mentioned statement that *RISP* could be involved in poplar defense responses (Rinaldi et al., 2007). Similarly, almost all *LRR-RLPs* remained stably expressed in both conditions, except two (POPTR_0003s02030 and LRR-RLP9) that were slightly (<2 fold) induced by wounding (**Figure 5A**). Interestingly, the predicted LRR-RLP9 gene is a very close paralog (>96% identity at the amino acid level) of POPTR_0580s00200 (hereafter LRR-RLP580), the *LRR-RLP* gene associated with the *RISP* gene. The co-regulation of both *RISP* and LRR_RLP9 genes observed by these oligoarray data has been further confirmed by a RNA-Seq analysis performed over a time-course infection of poplar leaves with virulent and avirulent strains of *M. larici-populina*. First, we observed a clear co-induction at 6 hpi, which likely correspond to a wounding response due to the leaf detaching (Figure S3). Second, we



confirmed the strong and specific induction of RISP during poplar defense responses against an avirulent strain of *M. larici-populina* as previously observed (Rinaldi et al., 2007). In this case, the expression of LRR-RLP9 was only slightly induced (Figure S3).

Next, the promoter regions of both *RISP* and *LRR-RLP9* genes have been analyzed to identify putative common regulatory elements explaining the observed co-expression. Whereas we did not detect particular signature or conservation in the 0 to -700 nucleotide region, we identified a stretch of ~300 bp with almost 90% nucleotide identity in the -700 to -1000 region (Figure S4). This striking conservation between the promoter regions of two unrelated genes may well explain the observed co-expression and will deserve future analysis.

RISP GENE IS PHYSICALLY ASSOCIATED WITH A CHIMERA OF LRR-RLP9 AND LRR-RLP580

The only paralog of RISP is POPTR_0009s11500 (hereafter termed RISP-like). RISP and RISP-like proteins show 63% identity and both carry a conserved predicted signal peptide (Figure 5D). In the current genome assembly, the *RISP-like* gene is adjacent to *LRR-RLP9* gene onto chromosome 9. Hence, two pairs of paralogous genes (*RISP/LRR-RLP580* and *RISP-like/LRR-RLP9*) seem to be present at different positions in the poplar genome (Figure 5D). Considering that the *RISP* and *LRR-RLP580* genes are located onto the small (~13 kb), unresolved scaffold 580, a poorly assembled part of poplar genome, we sought to investigate further the physical association between *RISP* and *LRR-RLP580* by using a combination of PCR amplification and

sequencing approaches (Figures 5D, 6A). Using *P. trichocarpa* genomic DNA as a template, we successfully amplified a major 8 kb fragment using a forward primer specific to the *RISP* gene (primer 1) and a reverse primer that cannot discriminate between the regions downstream the *LRR-RLP9* and *LRR-RLP580* genes (primer 4), both sequences being too similar (Figure 6B). By using additional primers targeting the coding sequence of both *LRR-RLPs* (primers 3 and 4), we amplified, from the 8 kb PCR product as a template, two fragments of ~5 and 3 kb, respectively, (Figure 6B). Sequencing of these fragments unambiguously confirmed the presence of the *RISP* gene, whereas the sequenced portion of the associated *LRR-RLP* was a chimera of *LRR-RLP9* and *LRR-RLP580*, likely indicating that the current genome assembly is not correct at this particular locus. This chimeric organization was further confirmed by amplifying the sequence coding the extracellular part of both *LRR-RLP* genes from cDNA, using primers 5 and 6. The sequencing of the cloned product revealed that the expressed sequence was indeed a combination of both predicted genes (Figure S5). Altogether, these observations indicated that *LRR-RLP9* and *LRR-RLP580* genes are mis-predicted and that they likely constitute a single gene that is physically associated and co-regulated with the *RISP* gene in response to wounding.

DISCUSSION

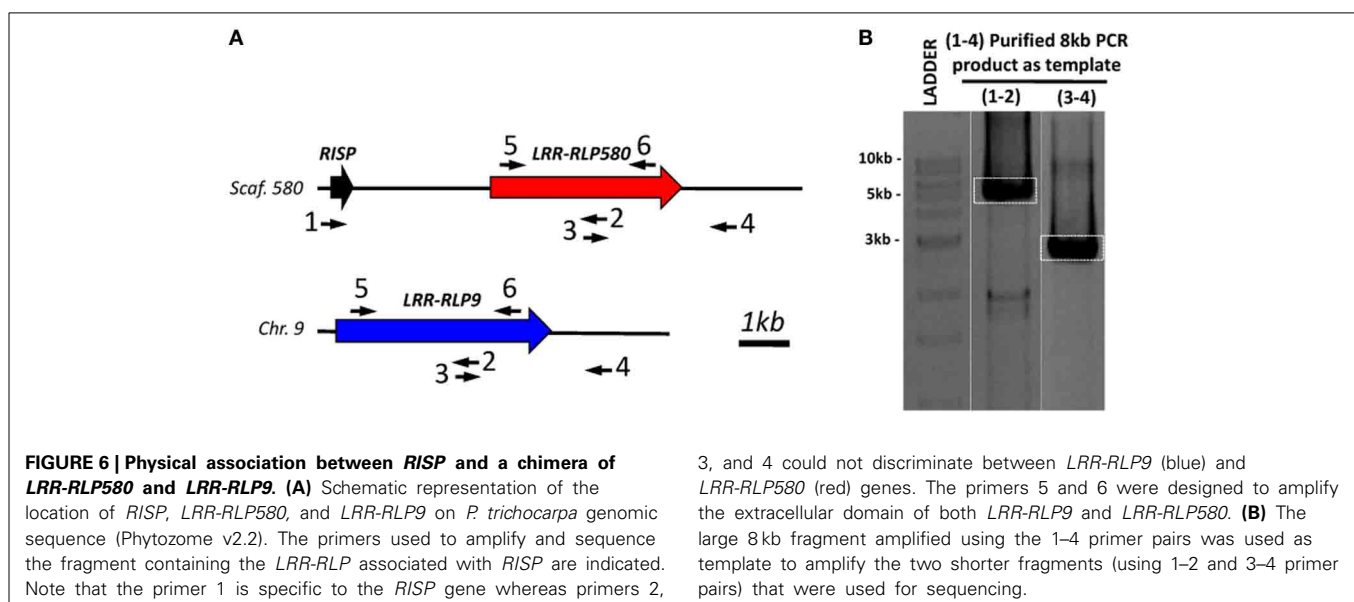
In this study, by analysing the LRR-RLP gene family in poplar, we observed that *LRR-RLP* genes are arranged into large clusters which are interspaced by genes coding SPUFs. Based on the observed co-regulation upon wounding of *RISP* and its physically associated LRR-RLP and on the presence of similar regulatory elements in their promoter regions, we hypothesize that both gene products may be functionally linked and by extension we suggest that some SPUFs could represent candidate endogenous ligands of their genome associated LRR-RLPs. Below, we did not discuss the genomic organization of LRR-RLP in poplar as this is not different from other plant species, but we rather discuss how to

validate LRR-RLP/SPUF interactions and whether such genomic analysis can be extended to other species.

DO SPUF GENES ENCODE CANDIDATE ENDOGENOUS LIGANDS OF LRR-RLPs?

In prokaryotes, genes that are functionally related are often gathered in operons and are co-regulated. Although such organization is less prominent in eukaryotes, a recent review illustrated the presence of operon-like clusters of secondary metabolism genes in plants, suggesting that physical association of genes can indeed reflect functional links between gene products (Kliebenstein and Osbourn, 2012). Besides, in tomato, the *Prf* gene clusters with several *Pto* resistance genes and it has been shown that the proteins physically interact (Rosebrock et al., 2007). Hence, operon-like clusters of genes coding interacting proteins involved in the immune response exist in plants. Nevertheless, physical associations of genes encoding a cell-surface receptor and its endogenous ligand have not been reported in plants so far. This may have been hampered by the limited number of identified receptor/ligand pairs or by the fact that the genes are not systematically associated in all plant genomes. Concerning the first point, genes coding characterized receptor/ligand pairs such as PEPRs/AtPeps or CLV2/CLE are not physically associated into the Arabidopsis genome. For the latter point, the existence of such physical association in one or in a few species only could have allowed in principle identifying functionally related receptor/ligand couples. However, the impossibility for example to unambiguously identify orthologs of AtPep, systemin, or CLV3 peptides in poplar genome suggests that such general inter-species comparisons could be precluded.

We showed that in poplar, *LRR-RLPs* cluster primarily with genes coding SPUFs. This physical association at the genomic level and the coordinated expression pattern of some *LRR-RLP/SPUF* pairs in some stress conditions suggest that they might function together, possibly as ligand/receptor couples. For instance, the physical association of *RISP* gene with an *LRR-RLP*



gene is corroborated by the fact that both genes are induced, although slightly for the *LRR-RLP*, in response to wounding. Previous studies reported *RISP* induction in leaves during effective immune responses to an avirulent strain of the leaf rust pathogen *M. larici-populina* (Rinaldi et al., 2007). Also, comparison of *RISP* sequences between different poplar cultivars suggests an evolution under positive selection, a characteristic of genes involved in biotic interactions (Petre et al., 2012). Taken together, these data strongly suggest a role for *RISP* in poplar immunity. Preliminary studies with a purified recombinant *RISP* have shown its ability to trigger poplar cell-culture alkalization, a hallmark of immune responses activation (Petre et al., unpublished data). Further studies should investigate if a direct interaction exists between selected SPUF/*LRR-RLP* pairs, but this remains technically challenging even in amenable plant species. Although recent studies showed that some cell-surface receptors are still functional when ectopically expressed in plants from different botanical families, such approach failed in most of the cases (Lacombe et al., 2010; de Jonge et al., 2012). Nevertheless, the recent finding that several *LRR-RLPs* require the *LRR-RLK* *SOBIR1/EVR* as a co-receptor to function may explain why transfer between different species did not work and may help overcoming this obstacle (Liebrand et al., 2013). We identified two very close *SOBIR1/EVR* homologs in poplar genome (POPTR_0015s09830 and POPTR_0012s09230, both 66% identity with *SOBIR1/EVR*), which are highly expressed in poplar leaves (Petre et al., 2012). However, although they are positioned on chromosomes 12 and 15, the two chromosomes with the highest *LRR-RLP* content, they are not included into the identified *LRR-RLP* clusters. Thus, possible future experiments could consist in the co-expression of poplar *LRR-RLPs* and poplar *SOBIR1/EVR* homologs in other amenable plants, followed by a treatment with the genome-associated purified SPUF and phenotyping.

DOES THE ANALYSIS OF GENOMIC ORGANIZATION COULD HELP IDENTIFYING *LRR-RLP* LIGANDS IN OTHER PLANTS?

Considering the obstacles encountered by past genetic approaches, our analysis could help accelerating the discovery of *LRR-RLP* functions in poplar and possibly in other model plant systems. In the *A. thaliana* genome, we identified 38 *LRR-RLP*-associated *SPUFs*, which present features similar to poplar *SPUFs* (i.e., orphan genes, species-specific). However, the density of *SPUF* genes in *A. thaliana* *LRR-RLP* clusters is lower compared with the average number of *SPUFs* in the predicted proteome (6 vs. 13%) (Table 1). Besides *SPUFs*, we identified some *LRR-RLP*-associated genes coding known peptide ligands in *A. thaliana* genome as rapid-alkalinization factors (*RALF*) and known defense-related secreted proteins (i.e., defensins) (Table S1D), which would be of interest for further investigations.

MATERIALS AND METHODS

IDENTIFICATION OF *LRR-RLP* GENES IN THE POPLAR GENOME, PHYLOGENY, AND SEQUENCE ANALYSES

LRR-RLP genes have been identified in the *P. trichocarpa* “Nisqually-1” genome sequence, version 2.2, hosted on the Phytozome portal (<http://www.phytozome.net/poplar>) by amino acid sequence homology searches using tomato Cf-9 protein

sequence as a query, then complemented with recursive searches with poplar *LRR-RLP* sequences representative from each clade (A–D). We considered only the *LRR-RLP* genes presenting the Cf-9 canonical domains and more than 15 *LRRs* (Fritz-Laylin et al., 2005). Among the 82 *LRR-RLP* genes retained for the analysis, only two required a manual correction due to intron mis-prediction (Table S1A, Text S1). For the phylogenetic analyses, we have retained the 45 strict Cf-9 homologs among the 57 *LRR-RLP* genes identified previously in *A. thaliana* (Wang et al., 2008). Only the conserved C3 and D domains (Fritz-Laylin et al., 2005) of *LRR-RLP* sequences and the whole *SPUF* sequences were used to build phylogenetic trees on the Phylogeny website (<http://www.phylogeny.fr/>), by using default parameters (Muscle alignment, Gblocks curation, PhyML phylogenetic tree, and SH-like approximate likelihood-ratio test for branch support) (Dereeper et al., 2008). The trees were observed and edited in the FigTree v1.2.3 software (<http://tree.bio.ed.ac.uk/software/figtree/>). All protein-coding genes present in the 50 kb vicinity of poplar and *A. thaliana* *LRR-RLP* genes were inspected manually on the Phytozome portal. This limit was determined to maximize the genomic area investigated and to have a single *LRR-RLP* in the genome portion analyzed. All genes coding for a small protein of unknown function, defined as proteins with a size below 200 amino acids and with no annotation and no pfam available, were retained and designated as *SPUF* genes. Only 12 needed a manual annotation for start and stop codons (Table S1A). Signal peptides for secretion were predicted with SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP-3.0/>) and SecretomeP 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>) servers and amino acid sequence alignments were executed on the Multalin website (<http://multalin.toulouse.inra.fr/multalin/>). Poplar and *A. thaliana* *SPUF* paralogs have been searched in respective predicted proteomes by amino acid homology searches, whereas homologs in other plants as well as conserved domains have been searched in the non-redundant protein database at NCBI (Blastp, *E*-value < 10⁻¹⁰).

POPLAR TRANSCRIPT EXPRESSION ANALYSIS

Poplar genome oligoarray data obtained from hybrid poplar “Beaupré” (*P. trichocarpa* × *P. deltoides*) leaf disks (i.e., wound treatment) collected at six different time-points from 18 to 48 h after inoculation with *M. larici-populina* (GSE39727 at NCBI GEO) were compared with strictly identical experiments performed on entire leaves (i.e., non-wounded treatment; GSE34802 at NCBI GEO). Poplar cultivation and inoculation procedures have been carried out as previously reported (Rinaldi et al., 2007). Isolation of total RNA, cDNA synthesis and oligoarrays were performed as previously described (Petre et al., 2012). The average of the six conditions in each dataset has been considered for quantitative comparison of gene expression with Microsoft Excel and R. Poplar oligoarray data are presented in Table S1C.

RNA-seq data of a time-course infection of poplar leaves with virulent and avirulent strains of *M. larici-populina* (same inoculum as in Rinaldi et al., 2007) used here to confirm *RISP/LRR-RLP9* expression profiles (Figure S3) were produced in the framework of a different research project and will be published elsewhere.

GENOMIC DNA ISOLATION, PCR AMPLIFICATION, AND SEQUENCING

Genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France) from 100 mg of *P. trichocarpa* “Nisqually-1” leaves. To amplify the genomic DNA fragment containing the *RISP* gene and the associated *LRR-RLP*, specific primers were designed in the coding sequence of the *RISP* gene (primer 1: 5'AGTAGCAAACAAAGTTGCCACCCCAGTC3') and in the downstream conserved sequence of both *LRR-RLP9* and *LRR-RLP580* (primer 4: 5'GAGATGCTAATGGGATGAGGTTT3'). The LongRange PCR Kit (Qiagen) was used to amplify an 8 kb fragment using 25 ng of *P. trichocarpa* genomic DNA as a template. Two additional primers (primer 2: 5'CATGCAAGTGGTTATTGCTCA3' and primer 3: 5'TGAGCAATAACCACTTGCATG3') were designed in the conserved coding sequence of both *LRR-RLPs* and two fragments (using the 1–2 and the 3–4 primer pairs) were amplified from 1 ng of the purified 8 kb PCR fragment using the LongRange PCR Kit. The amplified fragments were purified with the QIAquick PCR purification kit (Qiagen) and directly sequenced according to GenomeLab Dye terminator cycle sequencing with Quick Start kit (Beckman Coulter, Villepinte, France) on a Genetic Analysis System CEQ-8000 (Beckman Coulter). The sequence coding the extracellular domain of the chimera between LRR-RLP 580 and LRR-RLP 9 has been amplified from *P. trichocarpa* cDNAs using primer5: 5'ggggggCATATGTTGTCTCAAATTTCTCCTCT3'; and primer6: 5'ggggggGGATCCTTATGCTTTCCATCCAAATCCATC3' and cloned into the NdeI and BamHI restriction sites (underlined in the primers) of a pET28a vector before sequencing.

AUTHOR'S CONTRIBUTION

Benjamin Petre carried out the *in silico* analysis, transcriptome analyses, PCR cloning of the LRR-RLP extracellular part and drafted the manuscript. Stéphane Hacquard performed the PCR amplification/sequencing of the genomic fragment containing *RISP* and the LRR-RLP genes. Sébastien Duplessis performed the promoter analysis and contributed the unpublished RNA-Seq data. All authors contributed to the conceptual design of the experiments, participated in the writing of the manuscript, have read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors thank Cyril Zipfel and Sophien Kamoun at The Sainsbury Laboratory (UK), Hugo Germain at the Université du Québec à Trois-Rivières (Canada) and Armand Séguin at Natural Resources Canada (Canada) for the critical reading of an early version of the manuscript, Francis Martin and Annegret Kohler at INRA Nancy (France) for access to poplar oligoarrays and help with their normalization, Emmanuelle Morin and Emilie Tisserant at INRA Nancy (France) for their kind support with bioinformatics, Pascal Frey, Bénédicte Fabre, and Patrice Vion at INRA Nancy (France) for advices on poplar culture and technical help in green-houses. This work was founded by INRA and Région Lorraine grants to Sébastien Duplessis and Nicolas Rouhier. Benjamin Petre was supported by a young scientist grant (CJS) from INRA. This work was also supported by a grant overseen by the French National Research Agency (ANR) as part of

the “Investissements d'Avenir” program (ANR-11-LABX-0002-01, Lab of Excellence ARBRE).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00111/abstract>

REFERENCES

- Butenko, M. A., Kjersti Vie, A., Brembu, T., Aalen, R. B., and Bones, A. M. (2009). Plant peptides in signaling: looking for new partners. *Trends Plant Sci.* 5, 255–263. doi: 10.1016/j.tplants.2009.02.002
- de Jonge, R., van Esse, H. P., Maruthachalam, K., Bolton, M. D., Santhanam, P., Keykha Saber, M., et al. (2012). Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5110–5115. doi: 10.1073/pnas.1119623109
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, 465–469. doi: 10.1093/nar/gkn180
- Fritz-Laylin, L. K., Krishnamurthy, N., Tör, M., Sjölander, K. V., and Jones, J. D. G. (2005). Phylogenomic analysis of the receptor-like proteins of rice and *Arabidopsis*. *Plant Physiol.* 138, 611–623. doi: 10.1104/pp.104.054452
- Huffaker, A., Pearce, G., and Ryan, C. A. (2006). An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10098–10103. doi: 10.1073/pnas.0603727103
- Jeong, S., Trotochaud, A. E., and Clark, S. E. (1999). The *Arabidopsis* CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* 11, 1925–1934. doi: 10.2307/3871087
- Jones, D. A., Thomas, C. M., Hamond-Kosack, K. E., Balint-Kurti, P. J., and Jones, J. D. G. (1994). Isolation of the tomato Cf-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266, 789–793. doi: 10.1126/science.7973631
- Kemmerling, B., Halter, T., Mazzotta, S., Mosher, S., and Nürnberger, T. (2011). A genome-wide survey for *Arabidopsis* leucine-rich repeat receptor kinases implicated in plant immunity. *Front. Plant Sci.* 2:88. doi: 10.3389/fpls.2011.00088
- Kliebenstein, D. J., and Osbourn, A. (2012). Making new molecules—evolution of pathways for novel metabolites in plants. *Curr. Opin. Plant Biol.* 15, 1–9. doi: 10.1016/j.pbi.2012.05.005
- Krol, E., Mentzel, T., Chinchilla, D., Boller, T., Felix, G., Kemmerling, B., et al. (2010). Perception of the *Arabidopsis* danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. *J. Biol. Chem.* 285, 13471–13479. doi: 10.1074/jbc.M109.097394
- Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., van Esse, H. P., et al. (2010). Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nat. Biotechnol.* 28, 365–369. doi: 10.1038/nbt.1613
- Liebrand, T. W., van den Berg, G. C., Zhang, Z., Smit, P., Cordewener, J. H., America, A. H., et al. (2013). Receptor-like kinase SOBIR1/EVR interacts with receptor-like proteins in plant immunity against fungal infection. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10010–10015. doi: 10.1073/pnas.1220015110
- Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptors complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357. doi: 10.1016/j.pbi.2012.05.006
- Murphy, E., Smith, S., and De Smet, I. (2012). Small signalling peptides in *Arabidopsis* development: how cells communicate over a short distance. *Plant Cell* 24, 3198–3217. doi: 10.1105/tpc.112.099010
- Nadeau, J. A., and Sack, F. D. (2002). Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* 296, 1697–1700. doi: 10.1126/science.1069596
- Petre, B., Morin, E., Tisserant, E., Hacquard, S., Da Silva, C., Poulain, J., et al. (2012). RNA-Seq of early-infected poplar leaves by the rust pathogen *Melampsora larici-populina* uncovers *PtSultr3;5*, a fungal-induced host sulfate transporter. *PLoS ONE* 7:e44408. doi: 10.1371/journal.pone.0044408
- Rinaldi, C., Kohler, A., Frey, P., Duchaussoy, F., Ningre, N., Couloux, A., et al. (2007). Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol.* 144, 347–366. doi: 10.1104/pp.106.094987

- Rosebrock, T. R., Zeng, L., Brady, J. J., Abramovitch, R. B., Xiao, F., and Martin, G. B. (2007). A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature* 448, 370–374. doi: 10.1038/nature05966
- Ryan, C. A., and Pearce, G. (2003). Systemins: a functionally defined family of peptide signals that regulate defensive genes in Solanaceae species. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14577–14580. doi: 10.1073/pnas.1934788100
- Sun, Y., Li, L., Macho, A. P., Han, Z., Hu, Z., Zipfel, C., et al. (2013). Structural basis for flg22-induced activation of the *Arabidopsis* FLS2-BAK1 immune complex. *Science* 342, 624–628. doi: 10.1126/science.1243825
- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691
- Wang, G., Ellendorff, U., Kemp, B., Mansfield, J. W., Forsyth, A., Mitchell, K., et al. (2008). A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. *Plant Physiol.* 147, 503–517. doi: 10.1104/pp.108.119487
- Wang, G., and Fiers, M. (2010). Receptor-like proteins. Searching for functions. *Plant Signal. Behav.* 5, 540–542. doi: 10.4161/psb.11030
- Yamaguchi, Y., and Huffaker, A. (2011). Endogenous peptide elicitors in higher plants. *Curr. Opin. Plant Biol.* 14, 351–357. doi: 10.1016/j.pbi.2011.05.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 August 2013; accepted: 09 March 2014; published online: 28 March 2014.

Citation: Petre B, Hacquard S, Duplessis S and Rouhier N (2014) Genome analysis of poplar LRR-RLP gene clusters reveals RISP, a defense-related gene coding a candidate endogenous peptide elicitor. *Front. Plant Sci.* 5:111. doi: 10.3389/fpls.2014.00111

This article was submitted to *Plant-Microbe Interaction*, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Petre, Hacquard, Duplessis and Rouhier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.