Functional assignment to positively selected sites in the core type III effector RipG7 from solanacearum: RipG7 evolution, structure/function

Keke Wang, Philippe Remigi, Maria Anisimova, Fabien Lonjon, Ilona Kars, Andrey Kajava, Chien-Hui Li, Chiu-Ping Cheng, Fabienne Vailleau, Stéphane Genin, et al.

Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2016) 17(4), 553–564 DOI: 10.1111/mpp.12302

Functional assignment to positively selected sites in the core type III effector RipG7 from *Ralstonia solanacearum*

KEKE WANG[1,2], PHILIPPE REMIGI[1,2,†], MARIA ANISIMOVA[3], FABIEN LONJON[1,2], ILONA KARS[1,2,‡], ANDREY KAJAVA[4], CHIEN-HUI LI[5], CHIU-PING CHENG[5], FABIENNE VAILLEAU[1,2,6], STÉPHANE GENIN[1,2] AND NEMO PEETERS[1,2,*]

[1]*INRA, Laboratoire des Interactions Plantes Micro-organismes (LIPM), UMR441, CS52627 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France*

[2]*CNRS, Laboratoire des Interactions Plantes Micro-organismes (LIPM), UMR2594, CS52627 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France*

[3]*Institute of Applied Simulations, School of Life Sciences and Facility Management, Zürich University of Applied Sciences, Gruenalstrasse 14, 8820 Wädesnwil, Switzerland*

[4]*Centre de Recherche de Biochimie Macromoléculaire, CNRS, UMR5237, 1919 Route de Mende, 34000 Montpellier, France*

[5]*Institute of Plant Biology, National Taiwan University, Taipei 11529, Taiwan, R.O.C*

[6]*Université de Toulouse, INP, ENSAT, 18 Chemin de Borde Rouge, Castanet-Tolosan 31326, France*

SUMMARY

The soil-borne pathogen *Ralstonia solanacearum* causes bacterial wilt in a broad range of plants. The main virulence determinants of *R. solanacearum* are the type III secretion system (T3SS) and its associated type III effectors (T3Es), translocated into the host cells. Of the conserved T3Es among *R. solanacearum* strains, the Fbox protein RipG7 is required for *R. solanacearum* pathogenesis on *Medicago truncatula*. In this work, we describe the natural *ripG7* variability existing in the *R. solanacearum* species complex. We show that eight representative *ripG7* orthologues have different contributions to pathogenicity on *M. truncatula*: only *ripG7* from Asian or African strains can complement the absence of *ripG7* in GMI1000 (Asian reference strain). Nonetheless, RipG7 proteins from American and Indonesian strains can still interact with *M. truncatula* SKP1-like/MSKa protein, essential for the function of RipG7 in virulence. This indicates that the absence of complementation is most likely a result of the variability in the leucine-rich repeat (LRR) domain of RipG7. We identified 11 sites under positive selection in the LRR domains of RipG7. By studying the functional impact of these 11 sites, we show the contribution of five positively selected sites for the function of $RipG7_{CMR15}$ in *M. truncatula* colonization. This work reveals the genetic and functional variation of the essential core T3E RipG7 from *R. solanacearum*. This analysis is the first of its kind on an essential disease-controlling T3E, and sheds light on the co-evolutionary arms race between the bacterium and its hosts.

Keywords: LRR, *Medicago truncatula*, positive selection, *Ralstonia solanacearum*, type III effector, virulence function.

**Correspondence*: Email: nemo.peeters@toulouse.inra.fr

†*Present address*: New Zealand Institute for Advanced Study, Massey University, Auckland, New Zealand.

‡*Present address*: Bejo Zaden, B.V. Trambaan 1, 1749CZ Warmenhuizen, The Netherlands.

INTRODUCTION

During the co-evolution of host–pathogen interactions, plant hosts have evolved complex immune systems to prevent infections caused by pathogens. The latter have evolved infection strategies to suppress these host defence responses and colonize hosts (Dou and Zhou, 2012; Win *et al*., 2012). Among these strategies are the bacterial type III effectors (T3Es). They represent one of the most effective means used by many Gram-negative pathogenic bacteria of both plants and animals (Galán and Collmer, 1999; Raymond *et al*., 2013). Bacterial pathogens, such as the plant pathogens *Pseudomonas syringae*, *Xanthomonas* spp. and *Ralstonia solanacearum*, or the animal pathogens *Salmonella* spp., *Shigella* spp. and *Yersinia* spp., have repertoires of T3Es that they can inject into host cells to manipulate cellular pathways (Galán, 2009; Win *et al*., 2012).

Various isolates of these bacterial pathogens possess a set of T3Es consisting of a combination of conserved (or 'core') T3Es and specific T3Es. This has been described for the plant pathogens *P. syringae* (Baltrus *et al*., 2011) and *R. solanacearum* (Peeters *et al*., 2013a). Each bacterial T3E repertoire is shaped during the co-evolutionary arms race with the host plant immune system (Jones and Dangl, 2006). The persistence and further evolution of a given T3E is determined by a tradeoff between its capacity to participate in coaxing the host into accommodating the pathogen and its recognition by the plant specific (*R* gene-based) immunity.

Ralstonia solanacearum is a large species complex, present worldwide and divided into four phylotypes: Asian strains (phylotype I), American strains (phylotype II), African strains (phylotype III) and Indonesian strains (phylotype IV). Considering the sequence variability of some of the sequenced genomes, it has been debated whether this species complex should be classified into three subspecies (grouping together the Asian and African strains) (Peeters *et al*., 2013b; Safni *et al*., 2014). This species complex is remarkable for having a particularly large host range (Peeters *et al*., 2013a), as well as a large set of T3Es, with, for

instance, up to 75 T3Es in the strain Po82 (Peeters *et al*., 2013b; Xu *et al*., 2011). The analysis of sequenced genomes from this species complex enabled us to define a large set of 32 conserved or 'core' T3Es (Peeters *et al*., 2013b).

So far, the relationship between host specificity and T3E repertoire has not been clearly demonstrated (Baltrus *et al*., 2011; Lindeberg *et al*., 2009), probably owing to the fact that: (i) T3Es can have overlapping functions; and (ii) T3Es are not the only virulence determinants. One notable exception is the T3E RipG7 which has been shown to be required for the compatible interaction between strain GMI1000 of *R. solanacearum* and the host plant *Medicago truncatula* ecotype A17 (Angot *et al*., 2006). Interestingly, RipG7 is conserved in all strains sequenced so far and, as such, is part of the core T3E of *R. solanacearum* (Peeters *et al*., 2013b). These core T3Es are, by definition, conserved in the diverse strains of the *R. solanacearum* species complex, indicating a likely benefit to bacterial fitness. This is indeed the case, as it has been shown that, in addition to the requirement of RipG7 on *M. truncatula* (Angot *et al*., 2006), RipG7 is also necessary for the full aggressiveness of the strain GMI1000 on *Arabidopsis thaliana* and tomato plants (Remigi *et al*., 2011).

We have also shown that the Fbox motif is essential for RipG7 virulence function on *M. truncatula* (Angot *et al*., 2006). In eukaryotes, the Fbox-containing proteins have been shown to be essential for many cellular processes (Lechner *et al*., 2006). They form an SCF ('SKP1-CUL1-Fbox') protein complex that has an E3 ubiquitin ligase activity, directing specific proteins for ubiquitination and proteasome degradation. The Fbox domain is crucial as it enables the direct interaction of the Fbox protein with the SKP1-like protein (Zheng *et al*., 2002).

Other bacterial pathogens have been reported to contain Fbox proteins among their virulence arsenal (Kubori *et al*., 2010; Magori and Citovsky, 2011; Zaltsman *et al*., 2010). The C-terminal domain of Fbox proteins typically contains protein–protein interaction domains, allowing for the recruitment of host proteins prior to their ubiquitination and further proteasome degradation. Notably, the latter domains are very different, defining several families, with probably different target recognition, in a given organism (Xu *et al*., 2011). Among the Fbox proteins with a leucine-rich repeat (LRR) as C-terminal domain, there is evidence of positive selection on the LRR, suggesting diversification of target recognition (Wang *et al*., 2014). Other plant LRR-containing proteins, related to immunity functions, are also subject to strong evolutionary constraints. Positive selection has been detected in the LRR domains of the polygalacturonase-inhibiting proteins (PGIPs) (Casasoli *et al*., 2009; Kalunke *et al*., 2015). Similarly, it has been shown that the key amino acids of the resistance protein Cf-9 are hypervariable and solvent-exposed residues of the LRR (Wulff *et al*., 2009). The plant surface receptors EF-Tu receptor (EFR) and flagellin-sensing 2 (FLS2) are both mostly under purifying selection (Helft *et al*., 2011), consistent with the fact that they enable the plant cell to detect conserved pathogen peptides (Zipfel, 2009). Mutagenesis was instrumental in all of these studies to show the essential contribution of specific LRR amino acids to their overall or specific functions (Benedetti *et al*., 2013; Cao *et al*., 2013; Koller and Bent, 2014; Sun *et al*., 2012; Wulff *et al*., 2009).

In this work, a similar approach was carried out on the RipG7 core T3E of *R. solanacearum*. This enabled us to highlight the contribution to the function of RipG7 of positively selected residues situated in the LRR domain. Indeed, we took advantage of the fact that RipG7 is required for disease on *M. truncatula* to first test the functionality of different *ripG7* orthologues present in phylogenetically diverse strains of *R. solanacearum*. We then refined a previous analysis that indicated the presence of positive selection (Remigi *et al*., 2011), and identified several residues under high probability of positive selection. We then showed that some of these residues are important for the function of RipG7 on *M. truncatula*, demonstrating that the diversifying selection likely to have been imposed during the co-evolutionary arms race has a direct impact on the function of this virulence determinant.

RESULTS

ripG7 is sufficient, among the seven paralogous *ripG1* to *ripG7*, for *R. solanacearum* strain GMI1000 to colonize and cause wilting in *M. truncatula*

Previous virulence tests of single mutants generated in each of the seven paralogous *ripG1* to *ripG7* genes have shown that only the mutation in *ripG7* affects the pathogenicity of *R. solanacearum* on *M. truncatula* (Angot *et al*., 2006). In this work, we generated GRS138 (*ripG7* mutant, Table S1, see Supporting Information), GRS460 (*ripG2 ripG3 ripG6 ripG7* quadruple mutant, Table S1) and GRS447 (septuple *ripG1–ripG7* mutant, Table S1) expressing a chimeric $ripG7_{GMI1000}$ from an ectopic genomic integration, and tested the virulence of these strains on *M. truncatula*. Figure 1 represents the $\log_{10}$(hazard ratio), the hazard ratio being the wilting rate for plants root inoculated with GMI1000 divided by the wilting rate of plants inoculated with mutant or complemented strains (see Experimental details). No or little wilting occurs for plants inoculated with the quadruple (*ripG2 ripG3 ripG6 ripG7*, GRS460) or septuple (*ripG1–7*, GRS447) mutant, whereas the single *ripG7* mutant (GRS138) displays some occasional symptomatic plants (Fig. 1, strain GRS138 is significantly different from strains GRS460 or GRS447 with both $P < 0.001$, unpaired *t*-test). The pathogenicity of the three mutants was fully restored by $RipG7_{GMI1000}$ (Fig. 1). We then performed a quantification of bacterial growth in a gnotobiotic *M. truncatula in vitro* infection system (Fig. 1). The wilting phenotype is associated with a strong *in planta* bacterial growth; indeed, GMI1000 and complemented strains all show a median colonization higher than 10^9 colony-forming units/g fresh weight (cfu/gFW), whereas the three

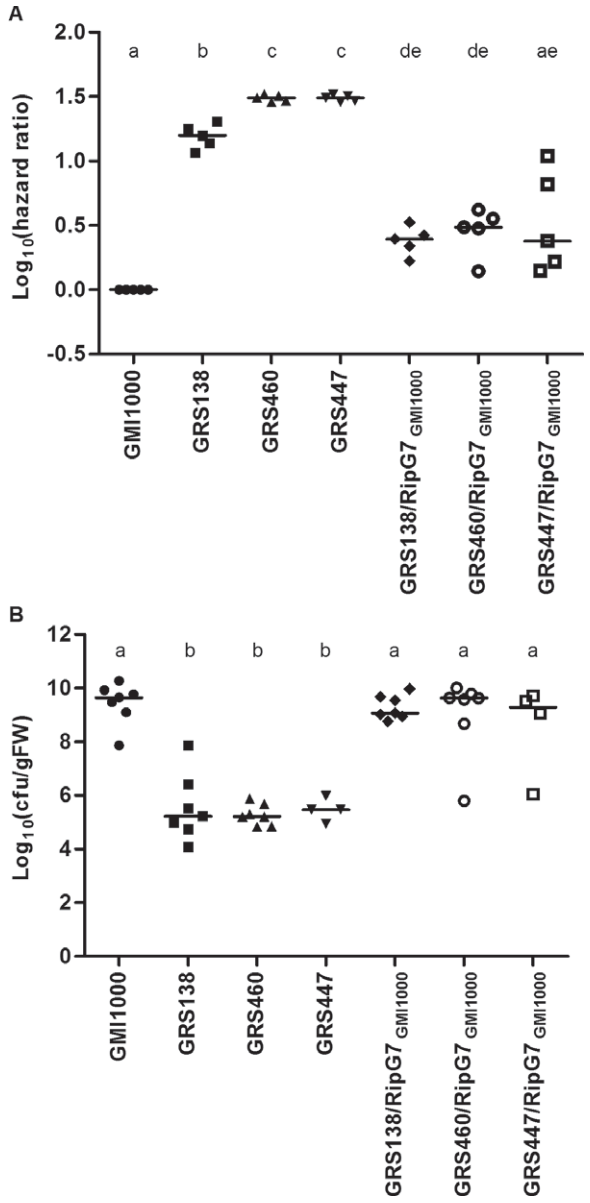


Fig. 1 Pathogenicity and *in planta* growth of *ripG7* mutants and complemented strains vs. wild-type GMI1000 on *Medicago truncatula*. (A) The data are represented as the hazard ratio between the tested strain and GMI1000. The death rates of plants inoculated with GRS138, GRS460 and GRS447 are significantly lower than the death rates of those inoculated with GMI1000. Moreover, GRS460 and GRS447 fall into the same group, according to unpaired *t*-test. GRS138/RipG7$_{GMI1000}$, GRS460/RipG7$_{GMI1000}$ and GRS447/RipG7$_{GMI1000}$ have a similar level of pathogenicity to GMI1000. Letters are used to represent groups after pairwise *t*-test ($P < 0.05$). (B) High levels of bacterial growth (colony-forming units per gram fresh weight, cfu/gFW) in *M. truncatula* are observed after inoculation with wild-type GMI1000 and complemented strains (GRS138/RipG7$_{GMI1000}$, GRS460/RipG7$_{GMI1000}$ and GRS447/RipG7$_{GMI1000}$). GRS138, GRS460 and GRS447 mutants all display significantly lower colonization than GMI1000. Letters are used to represent groups after pairwise Mann–Whitney tests ($P < 0.05$).

mutants (GRS138, GRS460 and GRS447) systematically display a much lower colonization (median < 10^6 cfu/gFW). This study aims to test the functionality of different RipG7 natural orthologues and mutagenized RipG7 sequences. In order to maximize the observed disease symptoms and bacterial colonization, we decided to use the quadruple GRS460 mutant strain as the tester recipient strain.

Only ripG7 orthologues from phylotype I and phylotype III strains can complement the loss of virulence of the GMI1000 *ripG7* mutant on *M. truncatula*

In order to obtain a broader knowledge of the existing genetic diversity, we cloned 16 *ripG7* orthologues (Fig. S1, Table S1 and Dataset S1, see Supporting Information) from a set of strains spanning the whole diversity of *R. solanacearum*. This dataset contains seven RipG7 orthologues from phylotype I strains and three from each of phylotypes II, III and IV.

To investigate the genetic and functional variation of *ripG7*, we selected eight *ripG7* orthologues among the 16 cloned *R. solanacearum* strains to evenly represent the four phylotypes (phylotype I: GMI1000, Pss190; phylotype II: Molk2, IPO1609; phylotype III: CMR15, CMR66; phylotype IV: Psi07, MAFF301558). Among these strains, only GMI1000, CMR15 and CMR66 are successful in causing wilting of *M. truncatula* plants (Fig. S2, see Supporting Information). We then tested whether the defective mutant GRS460 could be complemented for its virulence phenotype on *M. truncatula* by the different *ripG7* orthologues, fused to a triple haemagglutinin (HA) epitope tag, under the control of the GMI1000 *RipG7* promoter (see Experimental procedures). Figure 2 shows that the $\log_{10}$(hazard ratio) for plants inoculated with the GRS460 mutant is significantly higher than zero ($P < 0.001$, one-sample *t*-test), indicating that the death rate of plants inoculated with GRS460 is significantly lower than the death rate of those inoculated with the wild-type strain GMI1000 (as evidenced in Fig. 1). The $\log_{10}$(hazard ratio) of the GRS460 strain expressing a mutant version with a deletion of the RipG7 Fbox domain (GRS460/ripG7ΔF$_{GMI1000}$), known to be required for *R. solanacearum* virulence on *M. truncatula* (Angot *et al*., 2006), is also significantly higher than zero ($P = 0.003$, one-sample *t*-test). The $\log_{10}$(hazard ratios) of plants inoculated with GRS460/RipG7$_{GMI1000}$ or GRS460/RipG7$_{Pss190}$ are not significantly different from zero ($P = 0.099$ and $P = 0.276$, respectively, one-sample *t*-test), which shows that RipG7$_{Pss190}$ and RipG7$_{GMI1000}$ (see also Figs 1 and 2) can complement the GRS460 mutant and restore its virulence on *M. truncatula*. The strain GRS460/RipG7$_{CMR15}$ shows an intermediate phenotype, significantly different from both GMI1000 ($P < 0.001$, one-sample *t*-test) and GRS460 ($P < 0.001$, unpaired *t*-test). The complementation of GRS460 did not occur with the other phylotype III variant RipG7$_{CMR66}$ nor with orthologues from phylotype II (RipG7$_{Molk2}$ and RipG7$_{IPO1609}$) or

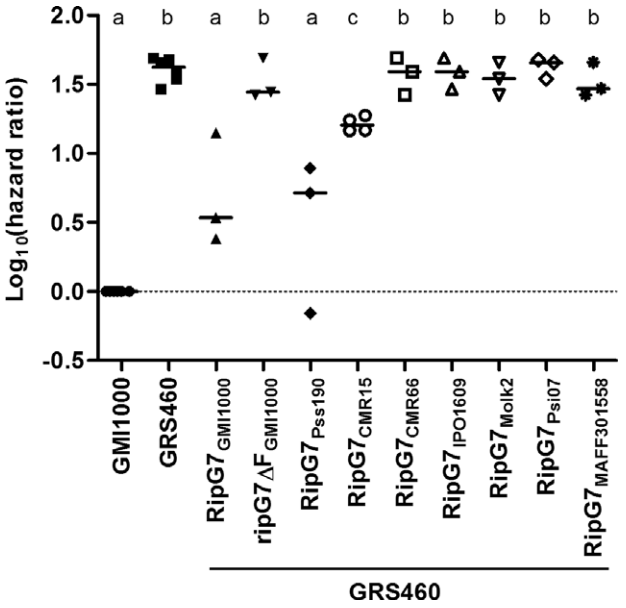


Fig. 2 Pathogenicity of GRS460 complemented with natural *ripG7* variants vs. wild-type GMI1000 on *Medicago truncatula*. *ripG7* orthologues can be separated into three groups: (i) $RipG7_{GMI1000}$ and $RipG7_{Pss190}$ fully restore the pathogenicity of GRS460 on *M. truncatula* to the wild-type level; (ii) $RipG7_{CMR15}$ partially restores the pathogenicity of GRS460; and (iii) $RipG7_{CMR66}$, $RipG7_{IPO1609}$, $RipG7_{Molk2}$, $RipG7_{Psi07}$ and $RipG7_{MAFF301558}$ fail to restore the pathogenicity of GRS460, like the negative control RipG7ΔF. Letters are used to represent groups after pairwise *t*-tests ($P < 0.05$).

phylotype IV ($RipG7_{MAFF301558}$ and $RipG7_{Psi07}$); the $\log_{10}$(hazard ratios) for the latter strains were significantly greater than zero ($P = 0.003$, $P = 0.002$, $P = 0.002$, $P = 0.001$ and $P = 0.002$, respectively; one-sample *t*-test) and similar to that of the GRS460 mutant strain ($P = 0.651$, $P = 0.398$, $P = 0.766$, $P = 0.265$ and $P = 0.735$, respectively; unpaired *t*-test).

To characterize the partial complementation of GRS460 by $RipG7_{CMR15}$, we quantified the *in planta* bacterial growth of GRS138 and GRS460 expressing $RipG7_{CMR15}$ (Fig. S3, see Supporting Information). There seems to be a higher variability in the measured bacterial load of GRS138/$RipG7_{CMR15}$ and GRS460/$RipG7_{CMR15}$, with instances of strong colonization, when compared with the recipient GRS138 and GRS460 mutants. Nonetheless, the bacterial growth of neither GRS138/$RipG7_{CMR15}$ nor GRS460/$RipG7_{CMR15}$ can be distinguished from that of GRS138 or GRS460 ($P = 0.063$ and $P = 0.297$, respectively; Mann–Whitney test; Fig. S3). This suggests that the low bacterial colonization of GRS460/$RipG7_{CMR15}$ could be the cause of the delayed symptoms observed for the *M. truncatula* plants inoculated with GRS460/$RipG7_{CMR15}$.

In order to verify that the ectopically expressed RipG7 orthologues are indeed substrates of the type III secretion system (T3SS), we performed *in vitro* secretion assays on the different strains. We showed that the eight RipG7 allelic versions tested could all be produced and secreted by *R. solanacearum*. As a control, a T3SS-defective mutant (*hrcV* mutant) failed to secrete $RipG7_{GMI1000}$ (Fig. S4, see Supporting Information).

Among the eight ripG7 orthologues tested, only $RipG7_{IPO1609}$ does not interact with MSKa, the *M. truncatula* SKP1-like SCF core component

A functional Fbox is required for the complementation of GRS138 by $RipG7_{GMI1000}$ (Angot *et al*., 2006). This is most probably a result of the fact that, for its function, RipG7 needs to interact with a plant SKP1-like protein (Angot *et al*., 2006), similar to a bona fide Fbox protein (Zheng *et al*., 2002). Considering that five of the eight diverse *ripG7* orthologues tested were not able to complement the GRS460 mutant strain for bacterial wilting (Fig. 2), we wanted to test whether the absence of complementation could be explained by an absence of interaction with MSKa, one of the main *M. truncatula* SKP1-like proteins (Angot *et al*., 2006). Co-immunoprecipitation experiments showed that $RipG7_{GMI1000}$-green fluorescent protein ($RipG7_{GMI1000}$-GFP), but not the Fbox-deleted version $RipG7\Delta F_{GMI1000}$-GFP, interacts with MSKa-cMyc (Fig. 3A). An interaction with MSKa could be detected for $RipG7_{Pss190}$ and $RipG7_{CMR15}$, the two orthologues that complement GRS460, as well as for several other orthologues that do not complement GRS460 ($RipG7_{CMR66}$, $RipG7_{Psi07}$, $RipG7_{MAFF301558}$, $RipG7_{Molk2}$). Only phylotype II $RipG7_{IPO1609}$ could not interact with MSKa-cMyc, neither in co-immunoprecipitation experiments (Fig. 3A), nor in yeast two-hybrid (Y2H) assays (Fig. 3B).

Selection study on RipG7 from 16 diverse strains identifies 11 sites under strong positive selection

As published previously, there is evidence that RipG7 is among the T3Es that seem to be evolving under significant positive selection (Peeters *et al*., 2013b; Remigi *et al*., 2011). The selection analysis was reiterated on the larger dataset of 16 orthologues (Dataset S1; Fig. S1) and allowed us to identify specific amino acids under positive selection. Significant across-site variation of selection on the protein was detected using the likelihood ratio test (LRT) of models M0 vs. M3 ($P < 0.01$) (see Experimental procedures). Accordingly, it was estimated that the majority of RipG7 sites were either strongly conserved with $\omega = 0.06$ (53.9% sites) or evolving almost free of selection pressure with $\omega = 0.83$ (39.3% sites). The remaining 6.8% of sites were estimated to be evolving under recurrent positive selection pressure with $\omega = 3.36$. Furthermore, strong evidence of positive selection on the protein was obtained using the LRTs of nested models M1a vs. M2a and M8a vs. M8 ($P < 0.01$ for both tests). As a result, we identified 11 sites in RipG7 under positive selection with high posterior probability ($P > 0.95$, see Fig. 4A). One site is located in the Fbox domain and the 10 others are located within, or in loops between, predicted LRR domains (Fig. 4A) (Kajava *et al*., 2008).

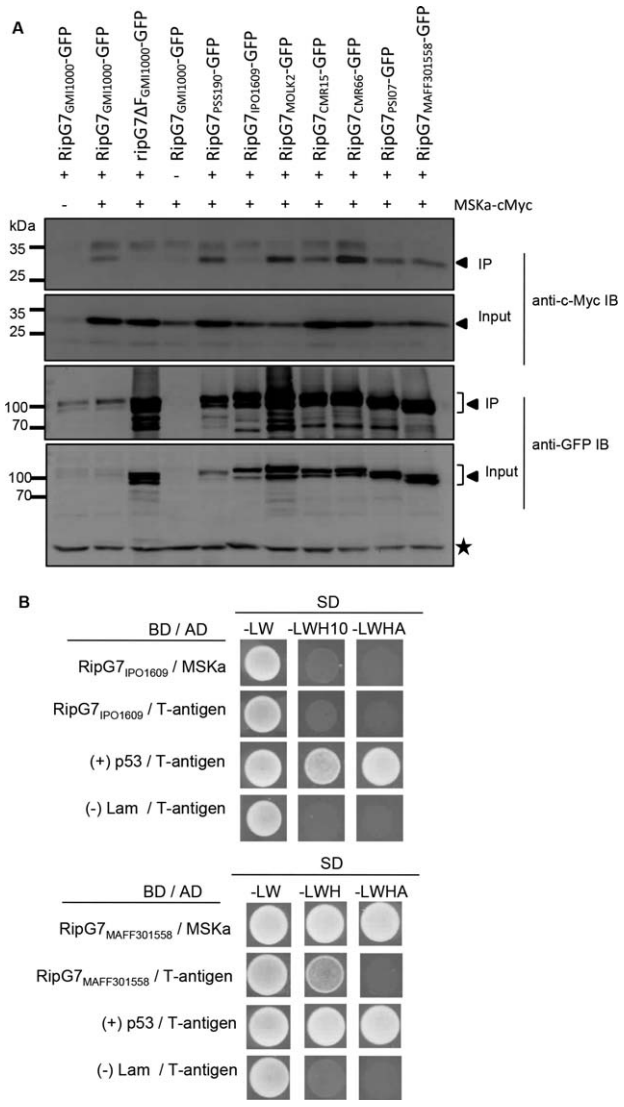


Fig. 3 Interaction between RipG7 natural variants and MSKa. (A) Co-immunoprecipitation of RipG7 with MSKa after transient expression in *Nicotiana benthamiana*. Cell lysates were immunoprecipitated with anti-green fluorescent protein (anti-GFP). The cell lysates (Input) and immunoprecipitates (IP) were analysed by immunoblotting with antibodies anti-c-Myc and anti-GFP. Black triangles point to the specific corresponding protein bands after immunoblotting (IB). A non-specific band shown with a black star confirms equal loading. Co-immunoprecipitation assay indicates that RipG7 natural variants can interact with MSKa in plants, except for $RipG7_{IPO1609}$. (B) Yeast two-hybrid (Y2H) analysis testing the interaction between BD-$RipG7_{IPO1609}$ and AD-MSKa in yeast cells. The two-hybrid interaction was tested with positive control BD-p53/AD-T-antigen and negative control BD-Lam/AD-T-antigen. No interaction between $RipG7_{IPO1609}$ and MSKa was indicated by the absence of growth on selective medium. The interaction between $RipG7_{MAFF301558}$ and MSKa was confirmed by growth on selective medium. SD-LW, medium without leucine (Leu) and tryptophan (Trp); SD-LWH10, medium without Leu, Trp and histidine (His) supplemented with 10 mM of 3-amino-1,2,4-triazole (3-AT); SD-LWHA, medium without Leu, Trp, His and adenine (Ade). AD, activation domain; BD, binding domain.

The amino acid variability at the selected sites is indicated for each of the natural variants of RipG7 in Fig. 4B.

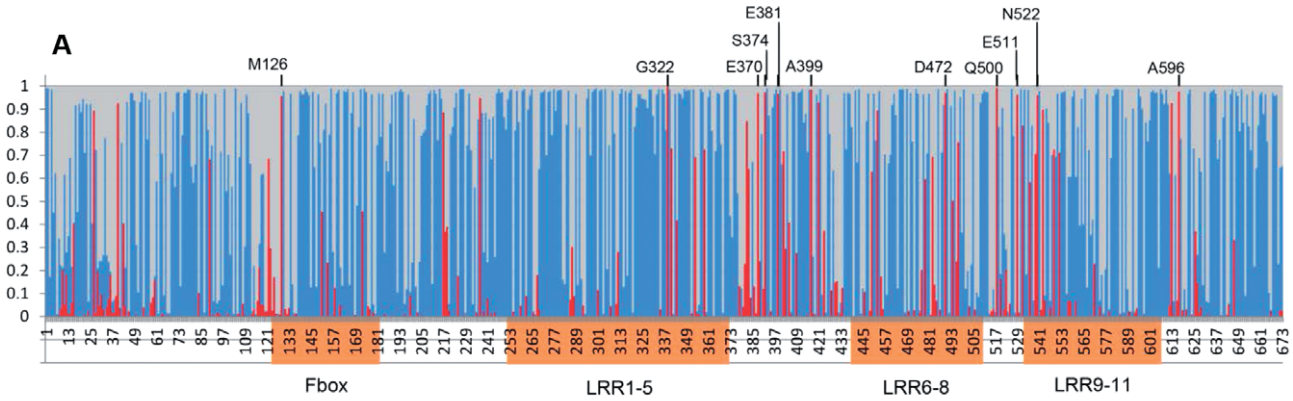
Loss-of-function site-directed mutagenesis reveals positively selected sites L126, G322, E500, D511 and D522 as contributing to the virulence function of $RipG7_{CMR15}$, but not of $RipG7_{GMI1000}$

To assess the functional role of the 11 positively selected sites in RipG7, we conducted site-directed mutagenesis using $RipG7_{GM1000}$ as template and generated mutations in each of the positively selected sites. We chose to replace the selected $RipG7_{GMI1000}$ sites with the corresponding residue of $RipG7_{MAFF301558}$ (phylotype IV, see Fig. 4), this orthologue being unable to complement the GRS460 mutant for disease establishment (Fig. 2), but still being capable of interacting with MSKa (Fig. 3A). The $\log_{10}$(hazard ratio) of the $RipG7_{GMI1000}$(11*) mutant cumulating the 11 positively selected mutated sites was not significantly different from zero ($P = 0.114$, one sample *t*-test; Fig. 5A), indicating that it could not be distinguished from its wild-type progenitor $RipG7_{GMI1000}$.

The same approach was performed between the more closely related phylotype III *ripG7* orthologues $ripG7_{CMR15}$ and $ripG7_{CMR66}$, displaying 92% identical amino acids. Although both strains are compatible on *M. truncatula*, only $RipG7_{CMR15}$ can complement the wilting defect of GRS460 (although partially, Fig. 2). Only five of the 11 positively selected sites are polymorphic between $RipG7_{CMR15}$ and $RipG7_{CMR66}$ (L126, G322, E500, D511 and D522, see Fig. 4B). We generated $RipG7_{CMR15}$(5*), cumulating changes in all five sites, using $RipG7_{CMR66}$ as a template. The GRS460 strain complemented with $RipG7_{CMR15}$(5*) was clearly less aggressive on *M. truncatula* than the strain GRS460 complemented with the wild-type version $RipG7_{CMR15}$ (Fig. 5B; $P = 0.009$, one-sample *t*-test). This result indicates that one or several residues among L126, G322, E500, D511 and D522 contribute significantly to the virulence function of $RipG7_{CMR15}$.

Modification of the positively selected sites is not sufficient to provide a gain of function to $RipG7_{MAFF3010558}$ or $RipG7_{CMR66}$

In a reverse experiment, we modified the 11 positively selected sites of $RipG7_{MAFF3010558}$ so that each site would display the corresponding $RipG7_{GMI1000}$ residue at each of the 11 sites. This mutated RipG7 was not functional on *M. truncatula* (Fig. S5A, see Supporting Information), as the recipient strain GRS460 was not distinguishable from the non-complementing $RipG7_{MAFF3010558}$ wild-type orthologue ($P = 0.939$, unpaired *t*-test). The same approach was performed between the more closely related phylotype III *ripG7* orthologues $RipG7_{CMR15}$ and $RipG7_{CMR66}$. In $RipG7_{CMR66}$, we replaced the five positively selected and polymorphic sites using $RipG7_{CMR15}$ residues, and these five point mutations were also insufficient to



B

Phylotype	Allele	126	322	370	374	381	399	472	500	511	522	596
I	GMI1000	M	G	E	S	E	A	D	Q	E	N	A
I	Pss190	I	R	E	S	E	A	D	Q	E	N	E
I	Pss81	I	R	E	S	E	A	D	Q	E	N	E
I	Pw12	M	G	E	S	E	A	D	Q	E	N	A
I	Pss166	M	G	E	S	E	A	D	Q	E	N	A
I	Pss79	M	G	E	S	E	A	D	Q	E	N	A
I	ss1-3	M	G	E	S	E	A	D	Q	E	N	E
III	CMR15	L*	G*	T	E	L	Y	R	E*	D*	D*	A
III	CMR66	I*	D*	T	E	L	Y	R	G*	G*	E*	A
III	NCPPB332	L	S	T	E	M	Y	R	G	G	E	A
IV	MAFF301558	L	R	A	G	G	S	F	G	G	I	Q
IV	PSI07	I	S	V	D	M	A	G	K	D	N	Q
IV	ACH732	L	Q	A	D	Q	S	S	D	E	G	A
II	CFBP2957	M	K	E	L	E	W	L	G	G	A	A
II	Molk2	M	K	R	S	G	W	C	T	Y	V	E
II	IPO1609	M	K	-	L	E	W	C	G	G	A	Q

Fig. 4 The 11 sites under strong positive selection in RipG7. (A) Posterior probability of positive (red), negative (blue) or relaxed (grey) selection is plotted along the coding sequence of RipG7. The *x*-axis coordinates correspond to the codons along the multiple alignments of all 16 RipG7 orthologues. On the bottom part of the graph are indicated the positions of the Fbox domain ($RipG7_{GMI1000}$ amino acid coordinates, 121–182), and the leucine-rich repeat (LRR) domains LRR1–5 ($RipG7_{GMI1000}$ amino acid coordinates, 236–355), LRR6–8 ($RipG7_{GMI1000}$ amino acid coordinates 421–492) and LRR9–11 ($RipG7_{GMI1000}$ amino acid coordinates, 515–586). On the top part of the graph are indicated the 11 sites (with the corresponding residues from $RipG7_{GMI1000}$) with the highest probability ($P > 0.95$) of positive selection: M126 ($P = 0.957$), G322 ($P = 0.999$), E370 ($P = 0.969$), S374 ($P = 0.973$), E381 ($P = 0.967$), A399 ($P = 0.984$), D472 ($P = 0.97$), Q500 ($P = 0.994$), E511 ($P = 0.963$), N522 ($P = 0.962$) and A596 ($P = 0.976$). (B) Table highlighting the polymorphism at the 11 positively selected sites for each of the 16 natural variants studied. An amino acid in a blue background is hydrophobic, in green is hydrophilic, in pink is negatively charged and in red is positively charged. The star symbols indicate the sites polymorphic between CMR15 and CMR66.

generate a gain-of-function RipG7 variant. Indeed, $RipG7_{CMR66}$(5*) cannot be distinguished from $RipG7_{CMR66,}$ and they are both unable to complement GRS460 to cause wilting of *M. truncatula* (Fig. S5B; $P = 0.223$, unpaired *t*-test).

DISCUSSION

The core T3E RipG7 is required for full virulence on the legume host plant *M. truncatula.* In contrast with that which has been published previously (Angot *et al.*, 2006), in our current and updated controlled inoculation procedures, the single *ripG7* mutant (strain GRS138) is not completely abolished in its virulence capacity (Fig. 1), as is the case for the quadruple mutant (*ripG2 ripG3 ripG6 ripG7* mutant, strain GRS460) or the septuple *ripG* mutant GRS447. This indicates that there seems to be a small contribution to virulence of any or several of RipG2, RipG3 and RipG6. Interestingly, this was not apparent when *in planta* bacterial colonization was assessed for these mutant strains (Fig. 1). The

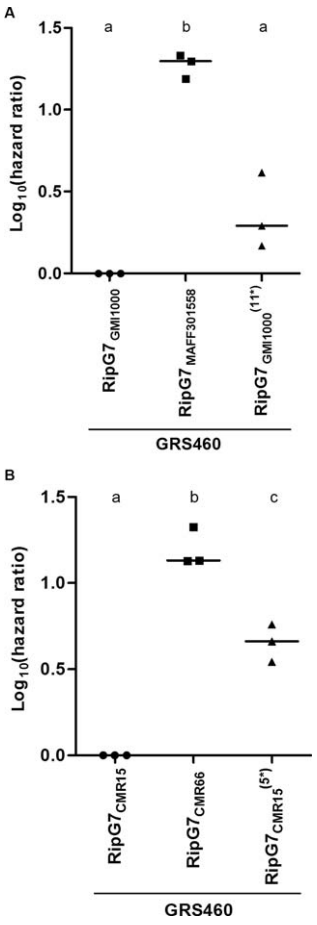


Fig. 5 Pathogenicity of GRS460/RipG7$_{GMI1000}$ and GRS460/RipG7$_{CMR15}$ vs. cumulated mutant strains on *Medicago truncatula*. (A) The death rate of plants inoculated with GRS460 containing RipG7$_{GMI1000}$(11*) could not be distinguished from their wild-type progenitor RipG7$_{GMI1000}$. (B) GRS460 containing RipG7$_{CMR15}$(5*) showed reduced pathogenicity compared with RipG7$_{CMR15}$. Letters are used to represent groups after pairwise *t*-test ($P < 0.05$).

marginal contribution of RipG2, RipG3 or RipG6 was ignored further in this work as the GRS460 quadruple mutant could be fully restored when complemented with the cognate GMI1000 *ripG7*-expressing construct.

None of the strains studied in detail in this work were sampled from legume plants, but rather from tomato (GMI1000, Pss190, CMR15, Psi07), potato (CMR66, MAFF301558, IPO1609) or banana crop (MolK2). Nonetheless, the legume plant *M. truncatula* is a host plant for *R. solanacearum* (Vailleau *et al*., 2007). When each of the eight above-mentioned strains were inoculated on *M. truncatula*, they fell into two categories: those able (GMI1000, CMR15 and CMR66) and those unable (Pss190, IPO1609, MolK2, MAFF301558 and Psi07) to produce wilting symptoms on the host. We know that: (i) RipG7 is an essential T3E of GMI1000 for its virulence on this host (Angot *et al*., 2006); and (ii) numerous other virulence determinants are needed for appropriate infection, colonization and wilting (Genin and Denny, 2012). The straightforward explanation would be that the first group (GMI1000, CMR15 and CMR66) has both a functional RipG7 and other required virulence determinants (e.g. RipAA, Turner *et al*., 2009; other virulence functions, Valls *et al*., 2006), whereas the second group could be deficient in either or both. In this study, we investigated further the sequence variability in relation to the function of these diverse RipG7 proteins. Only RipG7$_{GMI1000}$, RipG7$_{Pss190}$ and RipG7$_{CMR15}$ were able to complement the absence of wilting symptoms of the GRS460 strain, a GMI1000-based mutant. The first conclusion would be that GMI1000 and CMR15 have both functional RipG7s and the other required virulence determinants to cause wilting of *M. truncatula*. Pss190 has a functional RipG7, but not the other required virulence determinants. As the Pss190 T3E repertoire is very similar (presence/absence polymorphism and allelic diversity) to that of GMI1000 (C.-P. Cheng, personal observation, unpublished data), we can hypothesize that allelic differences, rather than presence/absence polymorphism in either the other T3Es or non-T3E virulence determinants, account for the inability of the Pss190 strain to cause wilting in *M. truncatula*. CMR66 does not have a GMI1000-compatible functional RipG7, but is capable of causing wilting in this host. For this latter puzzling case, we cannot be sure that the virulence of CMR66 is indeed associated with endogenous RipG7, nor do we know the exact virulence gene repertoire, as this strain has not yet been sequenced.

The case of CMR15 is compelling. On the one hand, the CMR15 strain is very aggressive on *M. truncatula* (Fig. S2, although the difference from GMI1000 is not significant, we recorded a negative hazard ratio for CMR15). On the other, the RipG7$_{CMR15}$ orthologue only weakly restores the wilting capacity of GRS460. We ruled out the possibility that this could be because of a dependence on any of RipG2, RipG3 or RipG6 for the colonization of gnotobiotic *M. truncatula* plantlets as the single *ripG7* mutant strain GRS138 expressing RipG7$_{CMR15}$ cannot be distinguished from strain GRS460/RipG7$_{CMR15}$ (see Fig. S3). We can speculate that CMR15 has evolved other virulence determinants (T3Es or others) required for full virulence on this host. It is interesting to note that CMR15 is the only *R. solanacearum* strain to have a *ripG8* gene (Remenant *et al*., 2010; Remigi *et al*., 2011). This paralogue is quite distant from RipG7 (Peeters *et al*., 2013b), but only a mutation analysis could inform us about its involvement in the virulence on this host.

RipG7 is a peculiar T3E, interacting directly with plant proteins (SKP1-like proteins) to create a bacterium–plant composite SCF E3-ubiquitin ligase in plant cells (Angot *et al*., 2006, 2007). Co-immunoprecipitation experiments allowed us to rule out the possibility that the absence of complementation is a result of the absence of interaction with SKP1-like (MSKa) for all but the phylotype II strain IPO1609. This result is in agreement with the fact that the LRR domains of Fbox proteins are key for the interaction with substrate proteins (Gagne *et al*., 2002; Wang

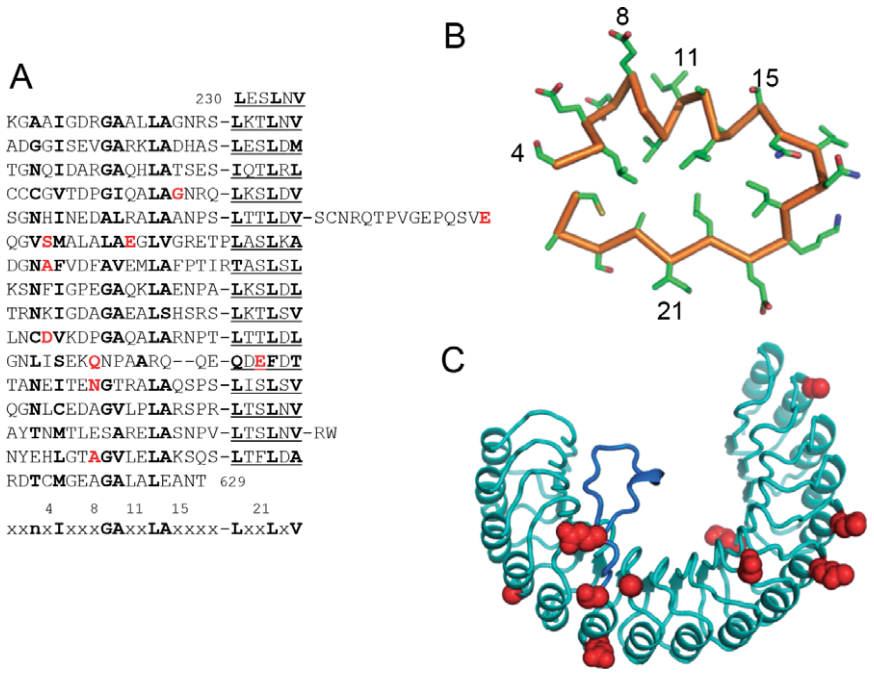


Fig. 6 Mapping of positively selected sites on a structural model of the RipG7$_{GMI1000}$ leucine-rich repeat (LRR) domain. (A) Alignment of the LRRs of RipG7$_{GMI1000}$. Conserved motifs corresponding to β-structural regions are underlined. Residues of LRR that are located inside the structure are indicated in bold. The positively selected sites are indicated in red. (B) A typical structural unit of GALA LRR observed in the crystal structure of GALA protein LegL1 from *Legionella pneumophila* (pdb code 4XA9). Positively selected sites are located in positions 4, 8, 11, 15 and 21 of the LRR, exposed on the surface of the structure. (C) A structural model of the LRR domain of RipG7 with the positively selected sites shown in red. A long loop of one of the LRRs is shown in dark blue.

et al., 2014). Our selection study on a global diversity set of 16 *ripG7* orthologues further emphasizes the importance of this domain, as 10 of the 11 strongest candidate sites for positive selection were located in, or in between, LRR repeats. This supports the hypothesis of an evolutionary arms race between *R. solanacearum* and its hosts, occurring at the interaction interface between RipG7 and its putative plant targets. Figure 6 presents the predicted structure of the LRR domain of RipG7$_{GMI1000}$; all sites but E511 are located in the α-helices of the LRR. As a result, 10 of the 11 selected sites are exposed (Fig. 6A) on the surface of the convex side of the predicted LRR of RipG7. For RipG7$_{CMR15}$, collectively or individually, the sites 126, 322, 500, 511 and 522 are required for full virulence contribution of this orthologue. At this stage, it is hard to speculate whether the concave part of the LRR of RipG7$_{CMR15}$ (with site 511), or the convex site (with sites 322, 500 and 522), is directly involved in target recognition. It should be noted that, for the plant Fbox LRR phytohormone receptor proteins TIR1 and COI1, the surface of interaction with IAA7 and JAZ1 proteins, respectively, has been mapped on the concave side of the LRR structure (Sheard *et al*., 2010; Tan *et al*., 2007).

Site 126 is the only non-LRR-located site and is surprisingly located in the Fbox domain for which the overall selection is, rather, purifying (Wang *et al*., 2014) (Fig. 4A). It is unexpected that there should be a positively selected site in the Fbox domain, as it interacts with a conserved SKP1-like plant protein (Kuroda *et al*., 2012). It should be noted that, although clearly under positive selection, this site is a methionine, leucine or isoleucine, all very similar in their chemical nature, and thus should not affect the interaction with MSKa, but could contribute to modulate this interaction, possibly in a host-dependent fashion.

Our analysis uncovered sites under significant positive selection in the overall dataset of 16 orthologues. We also showed that there is a different contribution to RipG7-mediated virulence of some of these amino acids, depending on the allelic background in which they are tested (GMI1000 or CMR15). This is evidence indicating that the T3Es (and RipG7) of specific *R. solanacearum* strains co-evolve with their specific host targets, for a constant best fit (RipG7 LRR domain interacting with putative plant targets), for a full virulence function. Although GMI1000, a phylotype I strain, and CMR15, a phylotype III strain, were both isolated from tomato (in French Guyana and Cameroon, respectively), they may have co-evolved on additional host plants. Recent work has shown that one single amino acid substitution in a virulence effector of *Phytophthora infestans* was sufficient to mediate a host shift in this species (Dong *et al*., 2014). Furthermore, a single amino acid modification in the *Fusarium verticilloides* endo-polygalacturonase enables it to avoid recognition by the LRR domain of PGIP2 in *Phaseolus vulgaris* (Benedetti *et al*., 2013). The recognition of polygalacturonases has also been shown to occur at the concave part of the PGIP2 LRR (Casasoli *et al*., 2009).

Gain-of-function mutagenesis has been successfully achieved by random modification of ATR1, an effector from the oomycete *Hyaloperonospora arabidopsidis*, for it to gain recognition by the plant LRR-immune receptor protein RPP1 (Steinbrenner *et al*., 2015). In our case, as we are testing for the contribution to aggressiveness, we could not embark on a systematic random mutagenesis, and so we tested the contribution of the 11 positively selected sites. The RipG7$_{MAFF301558}$(11*) cumulative *ripG7* mutant was unable to cause wilting in *M. truncatula* (Fig. S5). We

believed that a gain of function would be more likely to be achieved by modification of $RipG7_{CMR66}$ into a functional orthologue using the positive sites of the closely related $RipG7_{CMR15}$ as a template. Although there were only five positively selected sites among the 50 existing polymorphic amino acids, this gain of function was unsuccessful. It should be remembered here that the CMR66 strain is virulent on *M. truncatula*, despite $RipG7_{CMR66}$ being unable to complement the absence of virulence of GRS460. If the strain CMR66 evolved other virulence determinants, $RipG7_{CMR66}$ could have been under less stringent selection and could have accumulated deleterious mutations. A counter-argument to this is the fact that $RipG7_{CMR66}$ has maintained its ability to interact with MSKa.

Ralstonia solanacearum core T3Es are conserved across phylotypes which are spread over all continents and infect diverse host plants (Peeters *et al*., 2013b). These core T3Es must thus be essential for bacterial virulence and fitness. Nonetheless, each specific strain is adapted to a given set of hosts (or restricted to a specific host for some strains; Remenant *et al*., 2011). We can thus expect that the overall T3E repertoire, including the core T3Es, constantly co-evolves with the host plant genes whose products are targeted by these T3Es. This leads to the apparent contradiction in which a T3E can be ubiquitously represented and required in all *R. solanacearum* species complex strains and, at the same time, possess specific and minute changes enabling an adaptation to the specific and different host targets of the different strains.

EXPERIMENTAL PROCEDURES

New *ripG7* orthologues

Based on the conservation of RipG7 in each phylotype, we designed primers on the known RipG7 from reference strains GMI1000 (phylotype I), CMR15 (phylotype III) and Psi07 (phylotype IV) (Table S2, see Supporting Information). We then amplified, BP cloned (Gateway® cloning system, Life Technologies, Carlsbad, CA, USA) and sequenced the *ripG7* orthologues from the other phylotype I strains: Pss79, Pss81, Pss166, Pss190, Pw12 and ss1-3, all originating from Taiwan. Similarly, we cloned RipG7 from phylotype III strains CMR66 and NCPPB332 and phylotype IV strains MAFF301558 and ACH732. All the strain metadata and *ripG7* sequences are available (Table S2, Dataset S1). All accession numbers are given at the end of the article. All the variant *ripG7* sequences can unequivocally be assigned as being *ripG7* orthologues, as both the number and specific sequences of the LRR are strictly conserved (Peeters *et al*., 2013b; Remigi *et al*., 2011).

Plasmid constructs and *R. solanacearum* mutant strains generated

For *R. solanacearum* virulence tests, *ripG7* orthologues were cloned in a complementation vector pNP329 (Lohou *et al*., 2014; Monteiro *et al*., 2012) by the Gateway® cloning system (Life Technologies). These constructs allow for the expression of the diverse RipG7 clones under the control of the promoter from $RipG7_{GMI1000}$ and fused to a terminal triple HA epitope tag (Table S2). GRS138, GRS460 and GRS447, expressing RipG7 proteins, were generated by natural transformation of *R. solanacearum* (Bertolla *et al*., 1997). The strains used in this study are given in Table S2. All RipG7 constructs for Y2H assays and for the agroinfiltration of *Nicotiana benthamiana* leaves were also cloned using the Gateway® system; details are given in Table S2.

Virulence tests, bacterial internal growth curves (IGCs) and statistical analysis

Ralstonia solanacearum strains were grown overnight in B liquid medium (Plener *et al*., 2012) with appropriate antibiotics. Each soil inoculation consisted of 16 2-week-old *M. truncatula* (ecotype A17) plants grown in Jiffy pots that were inoculated by soil drenching with a 10^8 cfu/mL bacterial suspension after root cutting. Visual scoring of plant symptoms according to a scale ranging from '0' (no symptoms) to '4' (complete wilting) was performed as described by Vailleau *et al*. (2007). This procedure was repeated at least three times for each tested strain. The disease scoring was then transformed into binary data, with a disease index below '2' corresponding to '0' and a disease index equal to or higher than '2' corresponding to '1'. This transformation allowed us to apply the survival analysis statistical protocols particularly suited for distinguishing between two different strains (Machin *et al*., 2006; Remigi *et al*., 2011). The hazard is defined as the slope of the survival curve and is a measure of how rapidly plants are dying; the hazard ratio compares two survival curves. The hazard ratios were calculated by Graphpad Prism 5.0 software. The $\log_{10}$(hazard ratio) transformation was used to analyse the data, as it is assumed to follow a normal distribution (Machin *et al*., 2006). One-sample *t*-test was used to test the null hypothesis of the $\log_{10}$(hazard ratio) as being equal to the specified value of '0'. If $P < 0.05$, the null hypothesis that the tested strain behaves like the reference strain will be rejected. For IGCs, 2-week-old, *in vitro*-grown *M. truncatula* plantlets were inoculated with the bacterial suspension (10^8 cfu/mL). IGCs were performed by grinding six plantlets (surface sterilized) with a mortar and pestle. Bacterial concentrations were determined by plating, with a dedicated easySpiral® (Interscience 78860, Saint Nom, France), different bacterial dilutions on solid B medium with appropriate antibiotics. This gnotobiotic inoculation procedure was repeated at least three times for each tested bacterial strain. The $\log_{10}$(cfu/gFW) transformation was used to analyse the data. As the IGC data are known not to follow a normal distribution, a non-parametric Mann–Whitney test was used to determine whether or not a series of biological repetitions of two bacterial strain *in planta* IGCs can be considered as different.

Selection study

RipG7 sequences were aligned using the ProGraphMSA + TR program (Szalkowski and Anisimova, 2013), which implements the evolution-aware alignment algorithm performing well with indel-rich data and particularly with tandem repeats, as is our case. The phylogenies of all RipG7s and of individual RipG7 genes were reconstructed by maximum likelihood under model LG (Le and Gascuel, 2008) with Γ-rate variation among sites (Yang, 1994), as implemented in PhyMLv3.0 (Guindon *et al*., 2010). Branch

supports were estimated using the aBayes method (Anisimova *et al*., 2011). Using these phylogenies, selective pressures on RipG7 were analysed with codon substitution models (Yang *et al*., 2000) (as implemented in PAML), which describe selection on the protein by a distribution of the ω ratio—the ratio of non-synonymous to synonymous substitution rates. Heterogeneity of selective pressures across codon sites was tested using an LRT of codon models M0 (equal selection pressure over sites) vs. M3 (with three categories of sites). Tests for positive selection on the protein were also performed using LRT of pairs of nested models: M1a vs. M2a, and M8a vs. M8 (Anisimova and Kosiol, 2009). In these LRTs, for each pair of models, the null hypothesis (no positive selection) was represented by a model that constrains $\omega \leq 1$ (models M1a and M8a), whereas the alternative hypothesis (positive selection is allowed) allowed for $\omega > 1$ (models M2a and M8). Bayesian empirical Bayesian prediction (Yang *et al*., 2005) was then used to detect sites under positive selection employing the best-fitting model M8 (according to the Akaike Information Criterion). Dataset S2 presents the overall selection analysis on the RipG7 dataset.

Molecular modelling

The initial template for the $RipG7_{GMI1000}$ LRR was taken from a 24-residue LRR of the known crystal structure of GALA protein LegL1 from *Legionella pneumophila* (pdb code 4XA9; M. L. Urbanus *et al*., unpublished data, Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada). Based on this template, the LRR domain was built using the Insight II program (Dayringer *et al*., 1986). The amino acid sequence of the template was edited in accordance with the $RipG7_{GMI1000}$ sequences using the homology modelling option of the Insight II program. The structure was further refined as described in Kajava *et al*. (2008).

Site-directed mutagenesis

Site mutations were introduced by polymerase chain reaction (PCR) using a QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer's protocol. The mutagenic primers are shown in Table S2. PCRs for point mutations were run for 18 cycles of 50 s at 95 °C and 50 s at 60 °C, followed by 5 min at 68 °C. The resulting mutant plasmids were verified by sequencing.

Y2H analysis

RipG7 protein was fused to the GAL-binding domain (BD) and MSKa was fused to the GAL4-activation domain (AD) by the Gateway® cloning system. AH109 yeast cells were co-transformed with a BD-RipG7 fusion construct and AD-MSKa construct, and grown on SD-LW [medium without leucine (Leu) and tryptophan (Trp)]. Three co-transformants per interaction were spotted onto SD-LWH-10 [medium without Leu, Trp and histidine (His) supplemented with 10 mM of 3-amino-1,2,4-triazole] and SD-LWHA [medium without Leu, Trp, His and adenine (Ade)] selective medium and grown at 28 °C for 2–3 days.

Co-immunoprecipitation and immunoblot analyses

RipG7 orthologues and MSKa were fused to a GFP tag in pB7FWG2 and c-Myc tag in pGWB17 by Gateway® cloning, respectively. First, *N. benthamiana* leaves were co-infiltrated with two GV3101 *Agrobacterium tumefaciens* strains harbouring a RipG7-GFP and MSKa-cMyc-expressing plasmid. Then, ground leaf samples were resuspended in 400 μL of protein extraction buffer [GTEN buffer: 10% glycerol, 25 mM Tris-HCl, pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA), 150 mM NaCl, 10 mM dithiothreitol (DTT), protease inhibitor cocktail, 1 mM phenylmethylsulfonylfluoride (PMSF), 0.15% Nonidet P40]. The supernatant of this lysate and 20 μL of equilibrated GFP-Trap®_M beads (Chromotek, Planegg, Germany) were incubated for 1 h, followed by three washes with 500 μL of buffer (GTEN buffer, protease inhibitor cocktail, 1 mM PMSF). Proteins were eluted from the GFP-trap beads with 50 μL of 2 × SDS-sample buffer [100 mM Tris-HCl, pH 6.8, 4% sodium dodecylsulfate (SDS), 0.2% bromophenol blue, 20% glycerol) for 10 min at 95 °C. The eluted proteins were subjected to immunoblot analyses with anti-c-Myc and anti-GFP antibody.

Secretion assay

The secretion assay was performed as described previously (Guéneron *et al*., 2000; Poueymiro *et al*., 2009). The pAM5 plasmid (Guéneron *et al*., 2000) was introduced into GRS138/RipG7s mutant strains in order to overexpress the T3SS and T3Es. $RipG7_{GMI1000}$-expressing plasmid was introduced into the GMI1694 strain (*hrcV* mutant). All strains were grown overnight in B medium with appropriate antibiotics. B medium culture was washed by centrifugation (2300 *g*, 10 min), and bacterial pellets were resuspended in minimal medium (Plener *et al*., 2010) supplemented with 10 mM glutamate, 10 mM sucrose and 100 μg/mL of Congo Red at an optical density (OD) of 0.2. The minimal medium cultures were harvested after 8 h of growth at 28 °C and normalized to the same bacterial concentration. After centrifugation (2300 *g*, 10 min), bacterial cell pellets and culture supernatants were treated as described by Poueymiro *et al*. (2009). The cell pellets and culture supernatants were then subjected to immunoblot analysis with anti-HA antibody (diluted at 1:5000).

Phylogenetic analysis

All 16 RipG7 orthologues and the outgroup RipG6 from GMI1000 were aligned by the PRANK program (Löytynoja and Goldman, 2008) and subjected to phylogenetic tree reconstruction using PhyML (Guindon *et al*., 2010). Tree rendering was performed with iTOL (Letunic and Bork, 2011).

Accession numbers

The *ripG7* accession numbers are as follows. Phylotype I strains: GMI1000 (CAD15059), Pss79 (KR349453), Pss81 (KR349454), Pss166 (KR349452), Pss190 (KR349445), Pw12 (KR349455) and ss1-3 (KR349456); phylotype II strains: IPO1609 (GCA_001050995.1), MolK2 (GCA_000212635.2), CFBP2957 (CBJ43267); phylotype III strains: CMR15 (CBJ38611), CMR66 (KR349448) and NCPPB332 (KR349451); phylotype IV strains: Psi07 (CBJ51373), MAFF301558 (KR349450) and ACH732 (KR349449).

ACKNOWLEDGEMENTS

This work was supported by ANR-07-JCJC-0133 to NP, and was carried out at the Laboratoire des Interactions Plantes Micro-organismes (LIPM)

funded by LABEX 'TULIP' (ANR-10-LABX-41). This work benefited from an Short-term Scientific Mission grant promoted by the COST Action SUSTAIN FA1208. KW was supported by a grant from the China Scholarship Council. We thank Dr Miles Armstrong and Professor Paul Birch, from the James Hutton Institute, Dundee, UK, for technical advice. The authors declare the absence of any conflicts of interest.

REFERENCES

Angot, A., Peeters, N., Lechner, E., Vailleau, F., Baud, C., Gentzbittel, L., Sartorel, E., Genschik, P., Boucher, C. and Genin, S. (2006) *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. *Proc. Natl. Acad. Sci. USA*, **103**, 14 620–14 625.

Angot, A., Vergunst, A., Genin, S. and Peeters, N. (2007) Exploitation of eukaryotic ubiquitin signaling pathways by effectors translocated by bacterial type III and type IV secretion systems. *PLoS Pathog.* **3**, e3.

Anisimova, M. and Kosiol, C. (2009) Investigating protein-coding sequence evolution with probabilistic codon substitution models. *Mol. Biol. Evol.* **26**, 255–271.

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C. and Gascuel, O. (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**, 685–699.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D. and Dang, J.L. (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* **7**, e1002132.

Benedetti, M., Andreani, F., Leggio, C., Galantini, L., Di Matteo, A., Pavel, N.V., De Lorenzo, G., Cervone, F., Federici, L. and Sicilia, F. (2013) A single amino-acid substitution allows endo-polygalacturonase of *Fusarium verticillioides* to acquire recognition by PGIP2 from *Phaseolus vulgaris*. *PLoS ONE*, **8**, e80610.

Bertolla, F., Van Gijsegem, F., Nesme, X. and Simonet, P. (1997) Conditions for natural transformation of *Ralstonia solanacearum*. *Appl. Environ. Microbiol.* **63**, 4965–4968.

Cao, Y., Aceti, D.J., Sabat, G., Song, J., Makino, S.-I., Fox, B.G. and Bent, A.F. (2013) Mutations in FLS2 Ser-938 dissect signaling activation in FLS2-mediated Arabidopsis immunity. *PLoS Pathog.* **9**, e1003313.

Casasoli, M., Federici, L., Spinelli, F., Di Matteo, A., Vella, N., Scaloni, F., Fernández-Recio, J., Cervone, F. and De Lorenzo, G. (2009) Integration of evolutionary and desolvation energy analysis identifies functional sites in a plant immunity protein. *Proc. Natl. Acad. Sci. USA*, **106**, 7666–7671.

Dayringer, H.E., Tramontano, A., Sprang, S.R. and Fletterick, R.J. (1986) Interactive program for visualization and modelling of proteins, nucleic acids and small molecules. *J. Mol. Graph.* **4**, 82–87.

Dong, S., Stam, R., Cano, L.M., Bozkurt, T.O., Oliva, R., Liu, Z., Tian, M., Win, J., Banfield, M.J., Jones, A.M.E., Van der Hoorn, R.A.L. and Kamoun, S. (2014) Effector specialization in a lineage of the Irish potato famine pathogen. *Science*, **343**, 552–555.

Dou, D. and Zhou, J.-M. (2012) Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host Microbe*, **12**, 484–495.

Gagne, J.M., Downes, B.P., Shiu, S.-H., Durski, A.M. and Vierstra, R.D. (2002) The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **99**, 11 519–11 524.

Galán, J.E. (2009) Common themes in the design and function of bacterial effectors. *Cell Host Microbe*, **5**, 571–579.

Galán, J.E. and Collmer, A. (1999) Type III secretion machines: bacterial devices for protein delivery into host cells. *Science*, **284**, 1322–1328.

Genin, S. and Denny, T.P. (2012) Pathogenomics of the *Ralstonia solanacearum* species complex. *Annu. Rev. Phytopathol.* **50**, 67–89.

Guéneron, M., Timmers, A.C., Boucher, C. and Arlat, M. (2000) Two novel proteins, PopB, which has functional nuclear localization signals, and PopC, which has a large leucine-rich repeat domain, are secreted through the hrp-secretion apparatus of *Ralstonia solanacearum*. *Mol. Microbiol.* **36**, 261–277.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321.

Helft, L., Reddy, V., Chen, X., Koller, T., Federici, L., Fernández-Recio, J., Gupta, R. and Bent, A. (2011) LRR conservation mapping to predict functional sites within protein leucine-rich repeat domains. *PLoS ONE*, **6**, e21614.

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

Kajava, A.V., Anisimova, M. and Peeters, N. (2008) Origin and evolution of GALA-LRR, a new member of the CC-LRR subfamily: from plants to bacteria? *PLoS ONE*, **3**, e1694.

Kalunke, R.M., Tundo, S., Benedetti, M., Cervone, F., De Lorenzo, G. and D'Ovidio, R. (2015) An update on polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein that protects crop plants against pathogens. *Front. Plant Sci.* **6**, 146.

Koller, T. and Bent, A.F. (2014) FLS2-BAK1 extracellular domain interaction sites required for defense signaling activation. *PLoS ONE*, **9**, e111185.

Kubori, T., Shinzawa, N., Kanuka, H. and Nagai, H. (2010) Legionella metaeffector exploits host proteasome to temporally regulate cognate effector. *PLoS Pathog.* **6**, e1001216.

Kuroda, H., Yanagawa, Y., Takahashi, N., Horii, Y. and Matsui, M. (2012) A comprehensive analysis of interaction and localization of Arabidopsis SKP1-like (ASK) and F-box (FBX) proteins. *PLoS ONE*, **7**, e50009.

Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320.

Lechner, E., Achard, P., Vansiri, A., Potuschak, T. and Genschik, P. (2006) F-box proteins everywhere. *Curr. Opin. Plant Biol.* **9**, 631–638.

Letunic, I. and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* **39**, W475–W478.

Lindeberg, M., Cunnac, S. and Collmer, A. (2009) The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. *Mol. Plant Pathol.* **10**, 767–775.

Lohou, D., Turner, M., Lonjon, F., Cazalé, A.C., Peeters, N., Genin, S. and Vailleau, F. (2014) HpaP modulates type III effector secretion in *Ralstonia solanacearum* and harbours a substrate specificity switch domain essential for virulence. *Mol. Plant Pathol.* **15**, 601–614.

Löytynoja, A. and Goldman, N. (2008) Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science*, **320**, 1632–1635.

Machin, D., Cheung, Y.B. and Parmar, M. (2006) *Survival Analysis*. Chichester: John Wiley & Sons.

Magori, S. and Citovsky, V. (2011) Hijacking of the host SCF ubiquitin ligase machinery by plant pathogens. *Front. Plant Sci.* **2**, 87.

Monteiro, F., Solé, M., van Dijk, I. and Valls, M. (2012) A chromosomal insertion toolbox for promoter probing, mutant complementation, and pathogenicity studies in *Ralstonia solanacearum*. *Mol. Plant–Microbe Interact.* **25**, 557–568.

Peeters, N., Guidot, A., Vailleau, F. and Valls, M. (2013a) *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. *Mol. Plant Pathol.* **14**, 651–662.

Peeters, N., Carrere, S., Anisimova, M., Plener, L., Cazalé, A.C. and Genin, S. (2013b) Repertoire, unified nomenclature and evolution of the Type III effector gene set in the *Ralstonia solanacearum* species complex. *BMC Genomics*, **14**, 859.

Plener, L., Manfredi, P., Valls, M. and Genin, S. (2010) PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. *J. Bacteriol.* **192**, 1011–1019.

Plener, L., Boistard, P., González, A., Boucher, C. and Genin, S. (2012) Metabolic adaptation of *Ralstonia solanacearum* during plant infection: a methionine biosynthesis case study. *PLoS ONE*, **7**, e36877.

Poueymiro, M., Cunnac, S., Barberis, P., Deslandes, L., Peeters, N., Cazale-Noel, A.-C., Boucher, C. and Genin, S. (2009) Two type III secretion system effectors from *Ralstonia solanacearum* GMI1000 determine host-range specificity on tobacco. *Mol. Plant–Microbe Interact.* **22**, 538–550.

Raymond, B., Young, J.C., Pallett, M., Endres, R.G., Clements, A. and Frankel, G. (2013) Subversion of trafficking, apoptosis, and innate immunity by type III secretion system effectors. *Trends Microbiol.* **21**, 430–439.

Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., Fegan, M., Pruvost, O., Elbaz, M., Calteau, A., Salvignol, G., Mornico, D., Mangenot, S., Barbe, V., Médigue, C. and Prior, P. (2010) Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. *BMC Genomics*, **11**, 379.

Remenant, B., de Cambiaire, J.-C., Cellier, G., Jacobs, J.M., Mangenot, S., Barbe, V., Lajus, A., Vallenet, D., Medigue, C., Fegan, M., Allen, C. and Prior, P. (2011) *Ralstonia syzygii*, the blood disease bacterium and some Asian *R. solanacearum* strains form a single genomic species despite divergent lifestyles. *PLoS ONE*, **6**, e24356.

Remigi, P., Anisimova, M., Guidot, A., Genin, S. and Peeters, N. (2011) Functional diversification of the GALA type III effector family contributes to *Ralstonia solanacearum* adaptation on different plant hosts. *New Phytol.* **192**, 976–987.

Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L. and Kappler, U. (2014) Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex:

proposal to emend the descriptions of *Ralstonia solanacearum* and *Ralstonia syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii* subsp. *syzygii* subsp. nov., *R. solanacearum* phylotype IV strains as *Ralstonia syzygii* subsp. *indonesiensis* subsp. nov., banana blood disease bacterium strains as *Ralstonia syzygii* subsp. *celebesensis* subsp. nov. and *R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum* sp. nov. *Int. J. Syst. Evol. Microbiol.* **64**, 3087–3103.

Sheard, L.B., Tan, X., Mao, H., Withers, J., Ben-Nissan, G., Hinds, T.R., Kobayashi, Y., Hsu, F.F., Sharon, M., Browse, J., He, S.Y., Rizo, J., Howe, G.A. and Zheng, N. (2010) Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature*, **468**, 400–405.

Steinbrenner, A.D., Goritschnig, S. and Staskawicz, B.J. (2015) Recognition and activation domains contribute to allele-specific responses of an Arabidopsis NLR receptor to an oomycete effector protein. *PLoS Pathog.* **11**, e1004665.

Sun, W., Cao, Y., Jansen Labby, K., Bittel, P., Boller, T. and Bent, A.F. (2012) Probing the Arabidopsis flagellin receptor: FLS2–FLS2 association and the contributions of specific domains to signaling function. *Plant Cell*, **24**, 1096–1113.

Szalkowski, A.M. and Anisimova, M. (2013) Graph-based modeling of tandem repeats improves global multiple sequence alignment. *Nucleic Acids Res.* **41**, 1–11.

Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M. and Zheng, N. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature*, **446**, 640–645.

Turner, M., Jauneau, A., Genin, S., Tavella, M.-J., Vailleau, F., Gentzbittel, L. and Jardinaud, M.-F. (2009) Dissection of bacterial wilt on *Medicago truncatula* revealed two type III secretion system effectors acting on root infection process and disease development. *Plant Physiol.* **150**, 1713–1722.

Vailleau, F., Sartorel, E., Jardinaud, M.-F., Chardon, F., Genin, S., Huguet, T., Gentzbittel, L. and Petitprez, M. (2007) Characterization of the interaction between the bacterial wilt pathogen *Ralstonia solanacearum* and the model legume plant *Medicago truncatula*. *Mol. Plant–Microbe Interact.* **20**, 159–167.

Valls, M., Genin, S. and Boucher, C. (2006) Integrated regulation of the type III secretion system and other virulence determinants in *Ralstonia solanacearum*. *PLoS Pathog.* **2**, e82.

Wang, A., Fu, M., Jiang, X., Mao, Y., Li, X. and Tao, S. (2014) Evolution of the F-box gene family in Euarchontoglires: gene number variation and selection patterns. *PLoS ONE*, **9**, e94899.

Win, J., Chaparro-Garcia, A., Belhaj, K., Saunders, D.G., Yoshida, K., Dong, S., Schornack, S., Zipfel, C., Robatzek, S., Hogenhout, S.A. and Kamoun, S. (2012) Effector biology of plant-associated organisms: concepts and perspectives. *Cold Spring Harb. Symp. Quant. Biol.* **77**, 235–247.

Wulff, B.B.H., Heese, A., Tomlinson-Buhot, L., Jones, D.A., la de Peña, M. and Jones, J.D.G. (2009) The major specificity-determining amino acids of the tomato Cf-9 disease resistance protein are at hypervariable solvent-exposed positions in the central leucine-rich repeats. *Mol. Plant–Microbe Interact.* **22**, 1203–1213.

Xu, J., Zheng, H.-J., Liu, L., Pan, Z.C., Prior, P., Tang, B., Xu, J.S., Zhang, H., Tian, Q., Zhang, L.Q. and Feng, J. (2011) Complete genome sequence of the plant pathogen *Ralstonia solanacearum* strain Po82. *J. Bacteriol.* **193**, 4261–4262.

Yang, Z. (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314.

Yang, Z., Nielsen, R., Goldman, N. and Pedersen, A.M. (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, **155**, 431–449.

Yang, Z., Wong, W.S.W. and Nielsen, R. (2005) Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* **22**, 1107–1118.

Zaltsman, A., Krichevsky, A., Loyter, A. and Citovsky, V. (2010) Agrobacterium induces expression of a host F-box protein required for tumorigenicity. *Cell Host Microbe*, **7**, 197–209.

Zheng, N., Schulman, B.A., Song, L.Z., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., Conaway, R.C., Conaway, J.W., Harper, J.W. and Pavletich, N.P. (2002) Structure of the Cul1-Rbx1-Skp1-F box(Skp2) SCF ubiquitin ligase complex. *Nature*, **416**, 703–709.

Zipfel, C. (2009) Early molecular events in PAMP-triggered immunity. *Curr. Opin. Plant Biol.* **12**, 414–420.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Fig. S1 Maximum-likelihood phylogenetic tree of *ripG7*. The branches with bootstrap support greater than 0.80 are shown in red; the clades are collapsed when the average distance to leaves is smaller than 0.01. This is the case for the two subclades of phylotype I *ripG7* sequences.

Fig. S2 Pathogenicity of *Ralstonia solanacearum* strains vs. GMI1000 on *Medicago truncatula*. Hazard ratios are used to represent the pathogenicity of various strains of *R. solanacearum* in comparison with GMI1000. Only CMR15 and CMR66 show a similar level of pathogenicity to GMI1000. Letters are used to represent the groups of strains with similar behaviour after pairwise *t*-test ($P < 0.05$).

Fig. S3 *In planta* growth of *ripG7* mutants complemented with $RipG7_{GMI1000}$ or $RipG7_{CMR15}$. GRS138/$RipG7_{GMI1000}$ and GRS460/$RipG7_{GMI1000}$ show indistinguishable levels of bacterial growth in *Medicago truncatula* compared with wild-type GMI1000. Neither GRS138/$RipG7_{CMR15}$ nor GRS460/$RipG7_{CMR15}$ is significantly different from GRS138 or GRS460 in the level of bacterial colonization in *M. truncatula*.

Fig. S4 *In vitro* secretion assay of RipG7. GRS138 expressing *ripG7* orthologues and *R. solanacearum hrcV* mutant expressing $RipG7_{GMI1000}$ were grown in secretion medium; cell pellets (CP) and culture supernatants (SN) were harvested after 8 h of culture and analysed by immunoblotting with anti-haemagglutinin (anti-HA). Eight *ripG7* orthologues were visualized in both SN and CP, indicating that they are produced and secreted. The type III secretion system-defective mutant (*hrcV* mutant) failed to secrete $RipG7_{GMI1000}$.

Fig. S5 Pathogenicity of $RipG7_{MAFF301558}$ and $RipG7_{CMR66}$ cumulated mutants vs. the corresponding wild-type RipG7. The data are represented as hazard ratios. Neither the 11 cumulated mutations in $RipG7_{MAFF301558}$(11*) nor the five cumulated mutations in $RipG7_{CMR66}$(5*) produce an increase in aggressiveness on *Medicago truncatula* relative to their wild-type versions $RipG7_{MAFF301558}$ and $RipG7_{CMR15}$. Letters are used to represent groups of strains with similar behaviour after pairwise *t*-test ($P < 0.05$).

Table S1 Metadata of the 16 *Ralstonia solanacearum* strains used in this study.

Table S2 List of *Ralstonia solanacearum* strains, oligonucleotides and plasmids used in this study.

Dataset S1 Nucleotide sequences of the 16 *RipG7* natural variants.

Dataset S2 Posterior probabilities of the selection analysis.