The structural repertoire of *Fusarium oxysporum* f. sp. *lycopersici* effectors revealed by experimental and computational studies

Daniel S. Yu[1], Megan A. Outram[1], Ashley Smith[1], Carl L. McCombe[1], Pravin B. Khambalkar[1], Sharmin A. Rima[1], Xizhe Sun[1,2], Lisong Ma[1], Daniel J. Ericsson[1,3], David A. Jones[1], Simon J. Williams[1,§]

[1]Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia
[2]Key Laboratory of Hebei Province for Plant Physiology and Molecular Pathology, College of Life Sciences, Hebei Agriculture University, Baoding, China
[3]ANSTO, Australian Synchrotron, Clayton, Victoria 3168, Australia

[§]Corresponding author: Email: simon.williams@anu.edu.au (+61 2 6125 7862)

Keywords: fungal effectors, plant-microbe interactions, SIX effectors, *Fusarium oxysporum*, FOLD effectors.

Funding: Australian Research Council (DP200100388, FT200100135, DE170101165). Australian Academy of Science (Thomas Davies Grant). Australian National University Future Scheme (35665). Australian Institute of Nuclear Science and Engineering.

Abstract

Plant pathogens secrete proteins, known as effectors, that function in the apoplast and inside plant cells to promote virulence. Effectors can also be detected by cell-surface and cytosolic receptors, resulting in the activation of defence pathways and plant immunity. Our understanding of fungal effector function and detection by immunity receptors is limited largely due to high sequence diversity and lack of identifiable sequence motifs precluding prediction of structure or function. Recent studies have demonstrated that fungal effectors can be grouped into structural classes despite significant sequence variation. Using protein x-ray crystallography, we identify a new structural class of effectors hidden within the secreted in xylem (SIX) effectors from *Fusarium oxysporum f. sp. lycopersici* (*Fol*). The recognised effectors Avr1 (SIX4) and Avr3 (SIX1) represent the founding members of the Fol dual-domain (FOLD) effector class. Using AlphaFold ab initio protein structure prediction, benchmarked against the experimentally determined structures, we demonstrate SIX6 and SIX13 are FOLD effectors. We show that the conserved N-domain of Avr1 and Avr3 is sufficient for recognition by their corresponding, but structurally-distinct, immunity receptors. Additional structural prediction and comparison indicate that 11 of the 14 SIX effectors group into four structural families. This revealed that genetically linked effectors are related structurally, and we provide direct evidence for a physical association between one divergently-transcribed effector pair. Collectively, these data indicate that Fol secretes groups of structurally-related molecules during plant infection, an observation that has broad implications for our understanding of pathogen virulence and the engineering of plant immunity receptors.

Introduction

Fusarium oxysporum is a soil-borne fungal pathogen responsible for destructive vascular wilt diseases in a wide range of plants. Vascular wilting caused by *F. oxysporum* contributes to significant losses in crop production worldwide, ranking within the top 10 important fungal pathogens in scientific and economic importance [1]. Of particular concern is a variant of the banana pathogen, *F. oxysporum* f. sp. *cubense* tropical race 4 (*Foc*TR4), which is able to infect and cause Panama disease on the widely-grown Cavendish cultivar. The emergence and rapid spread of *Foc*TR4 has resulted in significant economic losses to banana growing regions worldwide and, without human intervention, has the potential to eradicate Cavendish as a commercial cultivar [2].

Arguably, the best-characterised *F. oxysporum* pathosystem is *F. oxysporum* f. sp. *lycopersici* (*Fol*), which specifically infects tomato. Previous studies of *Fol*-infected tomato identified a number of fungal proteins within the xylem sap [3]. These secreted in xylem (SIX) effector proteins represent major pathogenicity determinants across different *formae speciales* of *F. oxysporum*. Currently, 14 SIX effectors have been identified in *Fol* consisting of small (less than 300 amino acids in length), secreted, cysteine-rich proteins [4-7]. Most SIX effectors are encoded on the conditionally-dispensable chromosome 14 required for *Fol* pathogenicity [8]. This dispensable chromosome can be horizontally transferred from *Fol* to a non-pathogenic strain of *F. oxysporum*, resulting in a transfer of pathogenicity [5, 8]. To date, all 14 SIX effectors lack sequence similarity with proteins of known function, preventing prediction of function based on their amino acid sequence. Several SIX effectors have been shown to be essential for full virulence including SIX1, SIX2, SIX3, SIX5 and SIX6 from *Fol* [6, 9-12], SIX1 from *F. oxysporum* f. sp. *conglutinans* (*Focn*), which infects cabbage [13], SIX4 from *F. oxysporum* isolate Fo5176, which infects Arabidopsis [14], and SIX1 and SIX8 from *Foc*TR4 [15, 16]. *Fol SIX3* (Avr2) and *SIX5* are adjacent, divergently-transcribed genes with a common promoter and SIX5 has been shown to interact with SIX3 to promote virulence by enabling symplastic movement of SIX3 via plasmodesmata [17]. *Focn SIX8* and *PSE1* (pair with *SIX8* 1) are another such divergently-transcribed effector gene pair that function together to suppress phytoalexin production and plant immunity in Arabidopsis [18]. In *Fol, SIX8* forms a similar gene pair with *PSL1* (*PSE1*-like 1) [18]. Despite their roles in fungal pathogenicity, the virulence functions of most SIX effectors remain unknown.

To combat pathogen attack, plants have evolved resistance proteins that can recognise specific effectors secreted by pathogens leading to disease resistance. Four resistance genes, introgressed into tomato from related wild species, have been cloned. *I* and *I-7* encode

transmembrane receptor proteins containing extracellular leucine-rich repeat (LRR) domains and short cytoplasmic domains (LRR-RPs) [19, 20]. *I-2* encodes a cytoplasmic receptor containing nucleotide binding (NB) and C-terminal LRR domains [21], while *I-3* encodes a transmembrane protein with an extracellular S-receptor-like domain and cytoplasmic serine/threonine kinase domain (SRLK) [22]. *Fol* Avr1 (SIX4), Avr2 (SIX3) and Avr3 (SIX1) are recognised by tomato resistance proteins I, I-2 and I-3, respectively, leading to effector-triggered immunity and disease resistance [6, 23, 24]. To date, the effector recognised by I-7 remains unknown.

By understanding the function of *F. oxysporum* effector proteins, and how specific effectors are detected by resistance proteins, we (and others) hope to develop novel disease management strategies targeting vascular wilt diseases. Protein structure studies of effectors provide one avenue to assist this pursuit. Currently, Avr2 represents the only SIX effector whose protein structure has been determined [25]. Interestingly, the β-sandwich fold of Avr2 revealed that this effector has structural homology to ToxA from *Pyrenophora tritici-repentis* and AvrL567 from *Melampsora lini* [26, 27], despite a lack of sequence similarity. The observation of structural classes for effectors without identifiable domains or homologies to proteins of known function has now been described for four effector structural families, including the so-called MAX (***M****agnaporthe oryzae* **A**vr effectors and To**x**B from *P. tritici-repentis*) [28], RALPH (**R**N**A**se-**L**ike **P**roteins associated with **H**austoria) [29], LARS (***L****eptosphaeria* **A**vi**r**ulence-**S**uppressing) [30] and ToxA-like families [25-27].

Using x-ray crystallography, we have identified a new structural family of fungal effectors we term the FOLD (***Fol*** **d**ual-domain) effectors. Using experimental and theoretical approaches, we subsequently define and present the structural repertoire within the SIX effectors and 14 additional candidate effectors secreted by *Fol* during infection of tomato.

Results

The structures of Avr1 and Avr3 adopt a similar fold

We sought to solve the structures of Avr1 and Avr3 using x-ray crystallography. Both effectors are cysteine-rich and members of the K2PP (Kex2-processed pro-domain) effector class [31, 32]. Using an optimised protein production strategy [33], we produced Avr1 ($Avr1^{18-242}$) and Avr3 ($Avr3^{22-284}$) without their signal peptides as disulfide-bonded proteins in *E. coli* at yields sufficient for crystallisation studies (S1A and S1D Fig). From sparse matrix crystallisation screening, robust three-dimensional crystals were obtained for $Avr3^{22-284}$ (S1B Fig), however, $Avr1^{18-242}$ failed to crystallise. Previously, we demonstrated that the removal of the pro-domain

from the K2PP effector SnTox3 was required to obtain protein crystals [31] and predicted this may also be important for Avr1. Treatment of Avr1 with Kex2 *in vitro* resulted in a predominant Avr1 band of ~20 kDa consistent with a mature Avr1$^{59-242}$ protein, however, lower molecular weight bands were also observed suggesting *in vitro* Kex2 cleavage activity towards Avr1 at additional sites [31]. To address this problem, Avr1 was engineered with an internal thrombin cleavage site replacing the Kex2 site to produce a single Avr1$^{59-242}$ product after thrombin cleavage (S1C Fig). This protein was subsequently used for crystallisation studies resulting in rectangular plate-like crystals (S1E Fig).

The crystal structures of Avr1 and Avr3 were both solved using a bromide-ion-based single-wavelength anomalous diffraction (SAD) approach (S1 Table). The final structures were resolved to 1.65 Å and 1.68 Å, respectively (Fig 1A and 1B). Strikingly, despite sharing relatively low amino-acid sequence identity (19.5%), the structures revealed that Avr1 and Avr3 adopt a highly similar two-domain protein fold. Interpretable, continuous electron density was observed from residue 96 in Avr3 and, although the pro-domain was mainly disordered, some regions of the intact pro-domain could be interpreted in the electron density (residues 26-49) (S2A Fig). We also identified regions of the pro-domain within the Avr1 structure (residues 23-45 were observed in the electron density), despite cleavage of the pro-domain prior to crystallisation, indicating that an association between Avr1 and the pro-domain was maintained post cleavage *in vitro* (S2B Fig). The importance of this association, if any, remains unclear, but for simplicity, the structured regions of the pro-domains were excluded from further analysis.

The N-terminal domains (N-domains) in Avr1 and Avr3 encompass residues 59-139 and 96-191, respectively. Both structures have a similar architecture, consisting of an N-terminal α-helix followed by five β-strands (Fig 1A and 1B). When superimposed using the DALI server [34], they have a root-mean-square deviation (RMSD) of 2.1 Å and the six-cysteine residues within the domain form three disulfide bonds with conserved connectivity (Fig 1C and 2A). The C-terminal domains (C-domains) of Avr1 and Avr3 are also structurally similar, consisting of a β-sandwich domain architecture involving seven or eight β-strands, respectively (Fig 1D). When superimposed they have an RMSD of 2.8 Å. The C-domain of Avr3 contains a single disulfide bond that connects within strand β10, while the C-domain of Avr1 does not contain any cysteine residues. While the individual domains are very similar, superposition of the dual-domain structures returns an overall RMSD of ~3.4 Å. The larger difference is due to a rotation between the N- and C-domains (Fig 1E). In Avr1, a loop joins the two domains, whereas in Avr3 the domains are joined by a rigid, continuous β strand (β5).

The structures of Avr1 and Avr3, when compared with the solved structures of other fungal effectors, demonstrate that they adopt a unique two-domain fold and represent the founding members of a new structural class of fungal effectors we have designated the FOLD (***Fol*** **d**ual-domain) effectors.

SIX6 and SIX13 are predicted FOLD effector family members

We were interested in determining if other SIX effectors belonged to the FOLD effector family. One conserved sequence feature we noted in Avr1 and Avr3 was the spacing of the six cysteines within the N-domain. We analysed the cysteine spacing of the other SIX effectors and found that SIX6 and SIX13 contained a cysteine profile with similarity to that observed in Avr1 and Avr3 (Fig 2A). To determine whether these proteins were members of the FOLD effector family, we produced SIX6 and SIX13 proteins and obtained crystals for both proteins (S3C Fig), however the crystals diffracted poorly, and to date, we have been unable to solve the structures.

Recently, advances in *ab initio* structural prediction, particularly those generated by Google DeepMind's AlphaFold, have been shown to be highly accurate [35]. We decided to benchmark AlphaFold predicted models against our solved Avr1 and Avr3 (downstream of the Kex2 cleavage site (Avr1$^{59-242}$ and Avr3$^{96-284}$) (S4 Fig) to assess if this approach could be used for SIX6 and SIX13. The AlphaFold model of Avr1 returned a relatively poor confidence score (55%), however, the dual domain architecture was correctly predicted (S4A Fig). The RMSD between the model and structure was 6.9 Å, however this decreased significantly when comparing the N- and C-domains alone (1.5 Å for the N-domain; 2.5 Å for the C-domain) (S4A Fig). The AlphaFold model of Avr3 had a high confidence score (92%) and superimposed well to the solved structure (S4B Fig). However, the orientation of the individual domains was also skewed (RMSD = 3.7 Å overall; 1.1 Å for the N-domain; 0.8 Å for the C-domain) (S4B Fig). These results gave us confidence to continue with this approach.

We generated SIX6 and SIX13 models, downstream of the Kex2 cleavage site (SIX6$^{58-225}$, SIX13$^{78-293}$), using AlphaFold and obtained high confidence models (Fig 2B and 2C). The models of both proteins highly suggest their inclusion as members of the FOLD family. The N-domain of both SIX6 and SIX13 contains five β-strands held together by three disulfide bonds with an arrangement identical to Avr1 and Avr3. Comparison of these models with Avr1, as the FOLD effector representative, show that the N-domains of SIX6 and SIX13 have an overall RMSD of 2.9 Å and 2.2 Å, respectively (Fig 2D). Both the C-domains of SIX6 and SIX13 adopt a seven or eight stranded β-sandwich that is stabilised by a single disulfide bond

(β7 and β12 for SIX6; coil and β11 for SIX13). The β-strand topology within the C-domains is conserved, and the structures superimpose closely when compared to Avr1 (RMSD = 2.5 Å with SIX6; 1.8 Å with SIX13) (Fig 2E). One consistent difference between the FOLD effectors is the orientation of the C-domain relative to the N-domain, with the models of SIX6 and SIX13 having an RMSD of 6.2 Å and 3.9 Å, respectively, when compared with Avr1 (Fig 2F). Collectively, sequence analysis and structural modelling support the inclusion of SIX6 and SIX13 into the FOLD effector structural family.

Distinct structural families exist among the other SIX effectors

With the successful utilisation of AlphaFold to identify additional FOLD effectors, we decided to continue with this approach to model and compare the remaining SIX effectors (modelled without signal peptide and putative pro-domain if present) (Fig S5) against the experimentally determined structures of Avr1, Avr2 and Avr3. We split the FOLD effectors (Avr1 and Avr3) into N and C-domains for this analysis as they form distinct structural units. Structural similarity was compared using the PDBeFold server at the European Bioinformatics Institute [36], and measured by Q-score, which takes into account the RMSD and alignment length. Meaningful structural similarity was highlighted by a Q-score cut-off of 0.15 [37].

As expected, this analysis identified the structural similarities between the FOLD effectors at the N-domain, but it also highlighted extended structural similarities within the C-domain (Fig 3A). Based on Q-score cut-offs (>0.15) the C-domains of the FOLD effectors exhibited structural similarity to Avr2, which is a member of the ToxA-like class of effectors. Like the FOLD C-domain, Avr2 consists of a seven stranded β-sandwich fold, however, the topology of the strands differs (S6 Fig) suggesting they are structurally distinct. Interestingly, SIX7$^{49-220}$ and SIX8$^{50-141}$ (modelled downstream of putative pro-domains) also had Q scores > than 0.15 when compared to the FOLD C-domain, but both are more structurally similar to Avr2 (Fig 3B) and support their inclusion in the ToxA-like structural family.

We were interested to see if the AlphaFold models of other SIX effectors could be grouped further into structural classes and performed a structural pairwise alignment (as above), with groupings based on Q-scores. The highest similarities were between SIX7 and SIX8 (discussed above), SIX9$^{19-114}$ and SIX11$^{19-110}$, and SIX5$^{18-119}$ and SIX14$^{18-88}$ (modelled without signal peptides) (Fig 4A). Based on this analysis, we suggest that there are at least two additional families with multiple structurally-related members.

Despite sharing only 14% sequence identity, the AlphaFold models of SIX9 and SIX11 suggest they share structural folds. Here, we define these as members of structural Family 3.

Both models consist of five β-strands and either two or three α-helices (Fig 4B). When compared against existing structures in the PDB, both share structural similarity towards various RNA binding proteins (Q-score of > 0.15), such as the RNA recognition motif (RRM) fold of the Musashi-1 RNA-binding domain (PDB code: 5X3Z) [38] (Fig 4C).

Another family, designated Family 4, includes SIX5$^{18-119}$ and SIX14$^{18-88}$ (modelled without signal peptides) (Fig 4D). The models of these effectors are highly similar despite the proteins sharing only ~23% sequence identity. Both structural models include two α-helices and four to six β-strands with similar secondary-structure topology (S6 Fig). We compared the models of SIX5 and SIX14 against the PDB using PDBeFold and identified structural similarity toward heavy metal associated (HMA) domains including a mercuric ion reductase (PDB code: 2KT2) [39], sHMA94 from *Setaria italica* (PDB code: 7NMM) [40] and the integrated HMA-domain of the Pikp-1 nucleotide-binding, leucine-rich-repeat receptor (NLR) protein from *Oryza sativa* (PDB code: 5A6W) [41] (Fig 4E).

Based on Q-score, some members of Family 4 have structural similarity against members of Family 3 (Fig 4A). However, on close inspection the structural topology differs (S6 Fig). This suggests that Family 3 and Family 4 are distinct structural families.

Structural modelling and comparison of an expanded set of *Fol* effectors and effector candidates

The SIX effectors are a subset of effectors utilised by *Fol* during infection of tomato and we were interested in identifying structural similarities within an expanded set of effectors/putative effectors. Recently, the *Fol* genome has been re-sequenced [42] and reannotated, in combination with RNAseq data from *Fol*-infected tomato plants (Sun et al. in preparation). Several effector candidates have been identified, including some candidates not predicted or predicted incorrectly in the original genome annotation [5]. Based on reannotation and RNAseq data, 26 genes encoding novel effector candidates were identified that were consistently upregulated during *Fol* infection (Sun et al. in preparation). Of these, 14 genes encoded proteins with no recognised domains or motifs based on their amino acid sequences. We assessed whether these effector candidates could be grouped into the four structural families of SIX effectors already identified, by generating structural models with AlphaFold (without signal peptides and predicted pro-domains if present) (S4 Table, S5 Fig) and aligning them against the known or predicted structures of the SIX effectors (S7A Fig). Out of the 14 effector candidates, three appear to join the structural classes defined above.

Of note, the predicted structure of FOXGR_015533 adopted a nine β-strand sandwich and is likely a member of the ToxA-like class (Fig 5A), despite only sharing 13% sequence identity with ToxA and Avr2. PSL1 [18] and FOXGR_015322, here designated PSL2, are sequence related effectors (~85% sequence identity) with models consisting of two α-helices and four to five β-strands (Fig 5B). Both have high Q-score values against Family 4 and are likely members of this class despite sharing <19% sequence identity with other members of this family. As seen previously with other members of Family 4, PSL1 and PSL2 have some structural similarity towards members of Family 3 based on Q-scores. However, the structural topology of both candidates more closely resembles Family 4 (S6 Fig), likely making them members of this class. FOXG_18699 has structural similarity across many structural families including all members of Family 3 and 4, and SIX8 from the ToxA-like family. Despite these similarities, it has highest similarity based on Q-score (0.245) with FOXGR_015522 (S7D Fig) [43]. Both FOXGR_015522 and FOXGR_18699 adopt a ββαβββ structural topology, which differs from the other classes (Fig 5C). This suggests that FOXG_18699 and FOXGR_015522 (29% sequence identity) are members of a new structural family (designated Family 5).

Interaction between effector pairs from two structural families

Avr2 and *SIX5*, and *SIX8* and *PSL1* effector pairs have shared promoters that are divergently-transcribed in *Fol* [17, 18]. We demonstrate that both protein pairs are also likely to be linked structurally, with the Avr2-SIX5 and SIX8-PSL1 pairs each containing a ToxA-like family member (Avr2, SIX8) and a structural family 4 member (SIX5, PSL1). Previously, Avr2 and SIX5 have been shown to interact directly by yeast-two-hybrid (Y2H) analysis [10], however a direct interaction could not be established between homologues of SIX8 and PSL1 (PSE1) from *Focn* by Y2H analysis [18]. Considering the predicted structural similarities, and the potential for false negatives associated with Y2H assays, we were interested in testing whether SIX8 and PSL1 can interact using biophysical approaches. We successfully produced *Fol* SIX8$^{19-141}$ and PSL1$^{18-111}$ proteins using our adapted *E. coli* production system [33]. Based on amino acid sequence analysis and our experience with K2PP effectors, SIX8 contains a pro-domain, which we removed, resulting in a stable ~10 kDa protein (SIX8$^{50-141}$). To determine whether SIX8 and PSL1 directly interact, the purified proteins alone or co-incubated were analysed by size exclusion chromatography (SEC) (Fig 6). The elution profile of PSL1 shows a major peak (~12.25 mL) at a volume consistent with an oligomeric form of the protein (likely dimer), while SIX8 shows a major peak (~15 mL) consistent with a monomer and a minor peak (~13.25 mL) consistent with a dimer (Fig 6A). Strikingly, when incubated together the major

protein peaks migrate to ~12.8 mL (Fig 6A). SDS-PAGE analysis confirmed that both PSL1 and SIX8 are associated with that peak (Fig 6B), indicating that the migration of both proteins on SEC is altered after incubation, consistent with PSL1 and SIX8 forming a complex. These data also indicate that PSL1 self-association is disrupted by SIX8, as it elutes at a later volume. The conservation of the original peaks for PSL1 and SIX8, albeit significantly reduced, suggest that the PLS1-SIX8 complex can disassociate during the SEC experiment. Collectively, these data suggest PLS1 and SIX8 can form a heterodimer.

The I and I-3 immunity receptors recognise the N-domains of Avr1 and Avr3, respectively.

While the identification of the FOLD effector family and other putative structural families within *Fol* was unexpected and of interest, we are ultimately seeking to utilise the structural information they provide in functional experiments. Structural-based similarity searches using the N- and C-domains of the FOLD effectors suggest similarities with carbohydrate-binding proteins, tumour necrosis factors, cystatins and a range of viral proteins. Our functional studies based on these similarities are ongoing. Here, we sought to investigate recognition of FOLD effectors by their respective resistance proteins. Avr1 (SIX4) and Avr3 (SIX1) are recognised by the plant resistance proteins I and I-3 respectively, leading to effector-triggered immunity and disease resistance [6, 23]. Previously, we utilised a protein-mediated phenotyping approach to demonstrate that Avr1 is capable of causing cell-death in *I*-containing tomato cultivars and not in tomato cultivars lacking *I* [33]. With the discovery that Avr1 contains two domains, we sought to understand effector recognition utilising a structure-guided approach. We attempted to produce the N- and C-domains of Avr1 separately in *E. coli*, however, only the N-domain could be produced in a soluble form (S9A Fig). To determine whether the N-domain of Avr1 was sufficient for recognition by *I*, recombinant Avr1 N-domain was infiltrated into *I*-containing tomato cultivars. Cell death was observed when the N-domain of Avr1 was infiltrated into an *I*-containing tomato cultivar, but not when infiltrated into a tomato cultivar lacking *I* (Fig 7A). This was similar to the protein-mediated cell death observed with full length Avr1, first reported by [33], suggesting the N-domain of Avr1 is sufficient for *I*-mediated recognition (replicate data shown in S10A Fig).

Previous transient co-expression assays of *Avr3* and *I-3* failed to give any phenotype in *Nicotiana benthamiana* or *N. tabacum* [22]. With the successful utilisation of a protein-mediated phenotyping approach for the recognition of Avr1 by I, we attempted a similar approach with Avr3. Purified Avr3 protein infiltrated into cotyledons causes cell death in a

tomato cultivar that contained the *I-3* resistance gene (Fig 7B). Importantly, cell death was not observed when the same protein was infiltrated into tomato cultivars lacking *I-3*. We attempted to produce the N- and C-domains of Avr3 separately in *E. coli,* however, as found for Avr1, only the N-domain could be produced (S9B Fig). To determine whether the N-domain of Avr3 was sufficient for recognition by I-3, recombinant Avr3 N-domain was infiltrated into an *I-3*-containing tomato cultivar (Fig 7B). Cell death was observed when the N-domain of Avr3 was infiltrated into an *I-3*-containing tomato cultivar, although the cell-death phenotype was slightly weaker compared to full-length Avr3. No cell death was observed when the same protein was infiltrated into tomato cultivars lacking *I-3*, suggesting that the N-domain of Avr3 is sufficient for *I-3* recognition (replicate data shown in S10B Fig).

Discussion

Pathogenic fungi are in a continuous arms race with their plant hosts. To aid virulence, but avoid detection, effectors evolve rapidly causing significant diversity at the amino acid sequence level [44]. An emerging theme in fungal effector biology is the classification of effectors into families based on structural similarity. Here, we demonstrate that despite their sequence diversity, the *Fol* SIX effectors can be classified into a reduced set of structural families. This observation has implications for functional studies of SIX effectors, and ultimately our understanding of the infection strategies used by *F. oxysporum*.

Expanding the structural classes in fungal effectors

To date, five fungal effector families have been defined based on experimentally-determined structural homology, including the MAX [28], RALPH [29, 45, 46], ToxA-like [25-27], LARS [30, 47] and FOLD effectors, defined here. Effectors that fall within many of these structural families are shared across distantly related fungal species. The ToxA-like family includes effectors from fungi that group to both divisions of higher-fungi (basidiomycetes and ascomycetes) [25-27]. The MAX effector family were originally defined as AVR effectors from *M. oryzae* and ToxB from *P. tritici-repentis* [28] but pattern-based sequence searches suggest they are widely distributed amongst the Dothideomycetes [28, 48]. Similarly, LARS effectors, defined in *Leptosphaeria maculans* and *Fulvia fulva,* have structural homologues predicted in at least 13 different fungal species [30]. For the newly described FOLD effectors, homologues of SIX1 (Avr3), SIX6 and SIX13 are found in many *formae speciales* of *F. oxysporum* and are also present in several fungal genera including *Colletotrichum, Ustilaginoidea*, *Leptosphaeria, Pyrenophora* and *Bipolaris*. Collectively, the identification of

structural families reinforces the notion that fungi, across different genera and species, use a finite number of structurally-distinct effectors during invasion of their plant hosts.

Effector structure prediction

Solving the structures of fungal effectors is not a trivial undertaking. From challenges associated with effector protein production through to hurdles related to structure solution (such as experimental phasing), the research time required to determine an effector structure experimentally ranges from many months to many years (sometimes never). Not surprisingly, any reliable structural modelling methods are welcomed by researchers interested in effector biology. To this end, several recent studies have used effector structure prediction to expand our understanding of plant-microbe interactions [49, 50]. Bauer and colleagues used structural modelling to show that numerous recognised Avr effectors from the barley powdery mildew-causing fungal pathogen *Blumeria graminis (Bgh)* are members of the RALPH effectors class [49]. Seong and Krasileva used structural modelling to predict the folds of ~70% of the *Magnaporthe oryzae* secretome [50]. In doing so, they suggested an expansion in the number of MAX effectors and identified numerous sequence-unrelated groups of structural homologues (putative structural classes) within *M. oryzae*. Both these informative works preceded the recent release of Google DeepMind's AlphaFold, which represents a "game-changer" in *ab initio* structural modelling [35, 51, 52].

Here, we were in a unique position to apply and benchmark AlphaFold against experimentally determined structures for *Fol* effector prediction. We subsequently used AlphaFold to demonstrate that, within the repertoire of effectors we tested, up to five sequence-unrelated structural families are secreted during *Fol* infection (summarised in S11 Fig).

There are numerous caveats in relying solely on AlphaFold to generate structural models of effectors. The accuracy of models generated by AlphaFold can decline in cases with low numbers of homologues (~30 sequences in the multiple sequence alignment) [35]. This may help explain the low confidence prediction for SIX4 (Avr1) (S4A Fig), which is only distributed in a few *ff. spp.* of *F. oxysporum*. This poses a potential issue for predicting the structures of fungal effectors that lack homologues. In our hands, we have had mixed results when comparing several unpublished effector structures generated in our lab to AlphaFold models. In some instances, the models are completely wrong, however, in these cases the AlphaFold predictions reported low confidence scores, an important criterion for assessment of model reliability. Despite these limitations, the potential of machine-learning structural-

prediction programs like AlphaFold heralds an exciting era for a field that has long suffered from a lack of prediction power based on effector sequences alone.

Structural classes: A starting point for functional characterisation

Given their lack of sequence identity to proteins of known function or conserved motifs, structural determination of effectors is often pursued to provide functional insight and understanding of residues involved in recognition. The existence of structural families of effectors raises the question of whether links can now be made concerning their function based on structural similarities. Unfortunately, the FOLD effectors share little overall structural similarity with known structures in the PDB. However, at a domain level, the N-domain of FOLD effectors have structural similarities with cystatin cysteine protease inhibitors (PDB code: 4N6V, PDB code: 5ZC1) [53, 54], while the C-domains have structural similarities with tumour necrosis factors (PDB code: 6X83) [55] and carbohydrate-binding lectins (PDB code: 2WQ4) [56]. Though a functional link has not yet been established, the information gleaned from the FOLD effector structures gives us a starting point for further functional characterisation, with various avenues now being explored.

Interestingly, the predicted models for SIX9 and SIX11 within Family 3 have structural homology with RNA-binding proteins (PDB code: 3NS6, PDB code: 5X3Z) [38, 57], unrelated to RALPH effectors. Despite this structural homology, close inspection of these models suggests RNA binding is unlikely, as in both models the putative RNA binding surface is disrupted by a disulfide bond.

The putative family 4 effectors (SIX5, SIX14, PSL1 and PSL2) have structural homology with heavy metal associated (HMA) domains (Fig 4E). Metal binding within HMA domains is facilitated by conserved cysteine residues [58], however, their absence in the family 4 effectors suggests they are unlikely to have this activity.

The putative family 5 effectors (FOXGR_015522 and FOXG_18699) have structural homology with different proteins within the PDB. FOXGR_015522 is structurally similar to plant defensins (PDB code: 6MRY, PDB code: 7JN6) [59, 60] and K^+ channel-blocking scorpion toxins (PDB code: 1J5J, PDB code: 2AXK) [61, 62]. FOXG_18699 has structural homology with the C-terminal domain of bacterial arginine repressors (PDB code: 1XXB, PDB code: 3CAG) [63, 64].

A structural explanation for functional effector pairs

One interesting outcome of this study is a link between structural families and co-operative interactions between effectors. The ToxA-like effectors, Avr2 and SIX8 are known to form functional effector pairs with SIX5 and PSL1, respectively [10, 18]. According to our modelling work, both SIX5 and PSL1 are members of structural family 4. *Avr2* and *SIX5* are adjacent divergently-transcribed genes on *Fol* chromosome 14 and the protein products have been shown to physically interact [10]. Likewise, *SIX8* and *PSL1* are adjacent divergently-transcribed genes in the *Fol* genome and we demonstrate here a physical interaction between the proteins. Our data also suggests that the self-association of PSL1 is disrupted by the interaction with SIX8 suggesting the homo and hetero interaction may take place at the same interface. Structural studies, currently ongoing, will be required to decipher this interaction further and if the mode of interaction is conserved between Avr2 and SIX5.

Interestingly, two other *SIX* genes also form a divergently-transcribed gene pair on *Fol* chromosome 14. *SIX7* (ToxA-like family) and *SIX12* possess start codons 2,319 base-pairs apart and potentially share a common promoter. While SIX12 did not group with any structural families, the AlphaFold model had a very low prediction confidence (35.5%). On closer inspection of the sequence, we observed that the cysteine spacing in SIX12 closely resembles other family 4 members (S12 Fig). Collectively, this suggests that SIX12 may also be a family 4 member and we therefore speculate that SIX7 and SIX12 may function together, as described for the Avr2/SIX5 and SIX8/PSL1 pairs.

Effector structural classes and receptor recognition

Understanding the structural basis of plant immunity receptor-effector interactions represent a key step towards engineering plant immunity receptors with novel specificities. Recent structures of full-length NLR proteins reveal exquisite details of these direct interactions [65, 66]. To date, the major engineering advances concerning direct effector-receptor interactions involve the MAX effectors and NLRs with integrated HMA domains [67, 68]. Members of the MAX effector family, which contain a structurally conserved core but are divergent in sequence, are directly recognised by the HMA-domains of different NLRs leading to resistance [41, 69-71]. For example, Pikp recognises AVR–PikD, Pikm recognises AVR–PikD and other AVR–Pik variants [72], and the RGA5/RGA4 NLR pair recognises AVR–Pia [73]. It was recently shown the Pikp can recognise the “mismatched” effector AVR-Pia leading to partial disease resistance [74]. This highlights the capacity for a single NLR protein to bait multiple effectors with a structurally conserved core. Conserved effector classes also appear to drive diversification of NLR receptors, as is observed for the *Blumeria graminis* f. sp. *hordei* RALPH

effectors and multi-allelic barley MLA receptors [49], which provides opportunities for engineering NLRs that can recognise multiple effectors.

The FOLD effectors, Avr1 and Avr3, are recognised by different classes of resistance protein; I, a LRR-RP [20] and I-3, a SRLK [22]. While the mode of recognition has not yet been described, we demonstrate here that the N-domains of Avr1 and Avr3 are sufficient for recognition by their respective resistance proteins (Fig 7). This is significant because it demonstrates that different immunity receptor classes can recognise structural homologues. It might also help explain the function of Avr1 during *Fol* infection. When Houterman and colleagues identified Avr1, they demonstrated that it could suppress plant immunity conferred by the *I-2* and *I-3* receptors [23]. Considering our structural understanding of these FOLD effectors, it is plausible that Avr1 achieves suppression of I-3-mediated immunity by preventing Avr3 recognition through competitive inhibition. The LARS effectors represent another example of effectors that can activate and suppress resistance-gene-mediated immunity. AvrLm4-7 can prevent recognition of AvrLm3 and AvrLm9 (all LARS structural homologues [30]), by their cognate Rlm receptors [75, 76]. *Rlm9* encodes a wall-associated kinase [77], but the identify of *Rlm4*, *-7* and *-3* remain unknown. These studies demonstrate that members of at least two different structural effector families can suppress immunity triggered by structurally homologous effectors.

Collectively, the results presented here will aid future studies to understand the molecular basis of *F. oxysporum* effector function and recognition, and by extension, the design and engineering of immunity receptors with novel recognition specificities to help protect plants against *Fusarium* wilt disease.

Materials and methods

Vectors and gene constructs

SIX6, Avr1Thrombin, SIX6Thrombin, SIX8Thrombin, PSL1 and SIX13 coding sequences without the signal peptide containing Golden-Gate compatible overhangs were codon optimised for expression in *E. coli* and synthesised by Integrated DNA Technologies (IDT, Coralville, Iowa, USA) (S2 Table). The Kex2 cleavage motif of Avr1, SIX6 and SIX8 were replaced with a thrombin cleavage motif for pro-domain processing. Avr1 and Avr3 coding sequences were PCR amplified using *Fol* cDNA as a template with primers containing Golden-Gate compatible overhangs. For N- and C-domain sequences, constructs were PCR amplified from the Avr1 and Avr3 gene constructs using primers containing Golden-Gate compatible overhangs. All of the primers were synthesised by IDT (Coralville, Iowa, USA) (S3 Table). All genes were cloned into the modified, Golden-Gate-compatible, pOPIN expression vector [78]. The final expression constructs contained N-terminal 6xHis-GB1-tags followed by 3C protease recognition sites. The Golden-Gate digestion, ligation reactions and PCR were carried out as described by Iverson, Haddock [79]. All constructs were verified by sequencing.

Protein expression and purification

Sequence-verified constructs were co-expressed with CyDisCo in SHuffle T7 Express C3029 (New England Biolabs (NEB), Ipswich, Massachusetts, USA) and purified as previously described [33]. For Avr3, the buffers used after fusion tag cleavage were altered slightly to increase protein stability and a second IMAC step was excluded after the cleavage of the N-terminal fusion tag. During the cleavage step, the protein was dialysed into a buffer containing 10 mM MES pH 5.5 and 300 mM NaCl. The size-exclusion chromatography (SEC) HiLoad 16/600 Superdex 75 pg column (GE Healthcare) was equilibrated with a buffer containing 10 mM MES pH 5.5 and 150 mM NaCl.

For biochemical and crystallisation studies, Avr1, SIX6 and SIX8 with an internal thrombin cleavage site for pro-domain removal were processed with 2 to 4 units of thrombin from bovine plasma (600-2,000 NIH units/mg protein) (Sigma-Aldrich Inc., St. Louis, Missouri, USA) per mg of protein at 4°C until fully cleaved. This generated mature proteins encompassing residues 59-242 for Avr1, 58-225 for SIX6 and 50-141 for SIX8. Fully-cleaved protein was subsequently purified further by SEC using a HiLoad 16/600 or HiLoad 26/600 Superdex 75 pg column (GE Healthcare) equilibrated with a buffer containing 10 mM HEPES pH 8.0 or pH 7.5 and 150 mM NaCl. Proteins were concentrated using a 10 kDa molecular

weight cut-off Amicon centrifugal concentrator (MilliporeSigma, Burlington, Massachusetts, USA), snap-frozen in liquid nitrogen and stored at -80°C for future use.

Crystallisation, diffraction data collection and crystal structure determination

Initial screening to determine crystallisation conditions was performed at a concentration of 9.5 mg/mL for Avr3$^{22-284}$, 10 mg/mL for Avr1$^{18-242}$ and Avr1$^{59-242}$, 15 mg/mL for SIX6$^{17-225}$ and SIX6$^{58-225}$, and 14 mg/mL for SIX13 with and without Kex2 protease in 96-well MRC 2 plates (Hampton Research) at 18°C using the sitting-drop vapour-diffusion method and commercially available sparse matrix screens. For screening, 150 nL protein solution and 150 nL reservoir solution was prepared on a sitting-drop well using an NT8®-Drop Setter robot (Formulatrix, USA). The drops were monitored and imaged using the Rock Imager system (Formulatrix, USA) over the course of a month.

For Avr1$^{18-242}$, SIX6$^{17-225}$ and SIX13$^{22-293}$, no crystals were obtained from the different sparse matrix screens trialled. From initial screening, crystals with the best morphology for Avr3$^{22-284}$ were obtained in (1) 0.2 M sodium citrate tribasic dihydrate and 20% (w/v) PEG 3350 (SG1 screen: condition A5), (2) 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 (SG1 screen: condition D10), and (3) 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 (SG1 screen: condition F5). Crystals were visible after a period of 3 days and continued to grow for 3 weeks after initial setup. Replicate drops with 1 µl protein solution at 9.5 mg/mL and 1 µl reservoir solution were set-up in 24-well hanging-drop vapour-diffusion plates and produced crystals within 4 days that continued to grow over 1 month. No crystal optimisation was needed for Avr3, with the final conditions being (1) 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5, 25% (w/v) PEG 3350, and (2) 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5, 25% (w/v) PEG 3350. For Avr1$^{59-242}$, crystals with the best morphology were obtained in (1) 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.6, 25% (w/v) PEG 4000 (SG1 screen: condition C1) and (2) 0.2 M ammonium sulfate, 30% (w/v) PEG 8000 (SG1 screen: condition D7) within 1 day of initial setup. Crystal optimisation was carried out in 24-well hanging-drop vapour-diffusion plates at 18°C. The final optimised condition for Avr1$^{59-242}$ was 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.5, 17.5% (w/v) PEG 4000 at a protein concentration of 7 mg/mL with microseeding over a period of 3 weeks. For SIX6$^{58-225}$, crystals with the best morphology were obtained in 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 (SG1 screen: condition F5) within 2 days of initial setup. Crystals were optimised using hanging-drop vapour-diffusion plates and the final optimised condition for SIX6$^{58-225}$ was 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5, 20% (w/v) PEG

3350 at a protein concentration of 7 mg/mL over a period of 3 months. For SIX13, Kex2 protease was added to the protein at a 1:200 protease to protein ratio prior to crystal tray setup. Crystals with the best morphology were obtained in (1) 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 (SG1 screen: condition D10), and (2) 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 (SG1 screen: condition F5) within 2 days of initial setup. Crystals were optimised using hanging-drop vapour-diffusion plates and the final optimised condition for SIX13 was 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5, 25% (w/v) PEG 3350 at a protein concentration of 14 mg/mL.

Before X-ray data collection, crystals were transferred into a cryoprotectant solution containing reservoir solution and 15% (v/v) ethylene glycol or 20% (v/v) glycerol for Avr3$^{22-284}$, 10% (v/v) ethylene glycol and 10% (v/v) glycerol for Avr1$^{59-242}$, SIX6$^{58-225}$ and SIX13. For experimental phasing, Avr3$^{22-284}$ and Avr1$^{59-242}$ crystals were soaked in a cryoprotectant solution containing 0.5 M or 1 M sodium bromide for 10 seconds and backsoaked in the cryoprotectant without sodium bromide before being vitrified in liquid nitrogen. The datasets for bromide-soaked crystals were collected on the MX1 beamline at the Australian Synchrotron [80] (S1 Table). The datasets were processed in XDS [81] and scaled with Aimless in the CCP4 suite [82, 83]. The CRANK2 pipeline in CCP4 was used for bromide-based SAD phasing [84, 85]. Models were then refined using phenix.refine in the PHENIX package [86] and model building between refinement rounds was done in COOT [87]. The models were used as a template for molecular replacement against high resolution native datasets collected on the MX2 beamline at the Australian Synchrotron [88]. Automatic model building was done using AutoBuild [89], and subsequent models were refined with phenix.refine and COOT.

Structural modelling and alignments

Structural models were generated with Google DeepMind's AlphaFold using the amino acid sequences of SIX effectors and candidates without the signal peptide, as predicted by SignalP-5.0 [90] and predicted pro-domain if present [31] (S4 Table; S5 Fig). Pairwise alignments of the structural models generated by AlphaFold and the experimentally determined structures of Avr1 (PDB code: 7T6A) and Avr3 (PDB code: 7T69) were generated using the PDBeFold server at the European Bioinformatics Institute [36]. Structural similarity between the pairwise alignments were measured using Q-scores, which takes into account alignment RMSD and alignment length [36]. Structural similarity searches against the PDB were carried out using the PDBeFold server [36] and DALI server [34].

Tomato infiltration assays

Seeds of tomato cultivars IL7-3, M82 and Moneymaker were sown in seed-raising mix and grown in a controlled-environment chamber with a 16-h day/8-h night cycle at 22°C. Purified Avr1 and Avr3 proteins ($Avr1^{18-242}$ and $Avr3^{22-284}$) and N-domain only ($Avr1^{18-139}$ and $Avr3^{22-191}$) were diluted in water to a concentration of 0.1 mg/mL. Cotyledons of 10-14-d old tomato seedlings were syringe infiltrated with 100 µl of protein or buffer (0.1 mM HEPES pH 8, 1.5 mM NaCl or 0.1 mM MES pH 5.5, 1.5 mM NaCl). Cotyledons were harvested and photographed at 4 or 7 days post-infiltration (dpi).

Interaction studies between PSL1 and SIX8

To investigate whether PSL1 and SIX8 interacted *in vitro* ~250 µg of PSL1 and SIX8 individually, and ~350 µg PSL1 and 250 µg of SIX8 were injected onto a Superdex 75 Increase 10/300 (Cytiva) column pre-equilibrated in 20 mM HEPES pH 7.5, 150 mM NaCl, after a 30 min room temperature incubation. Samples across the peaks were then analysed by Coomassie-stained SDS-PAGE.

Acknowledgements

This work was supported by the Australian Research Council (ARC DP200100388 D.J./S.W.) and the Australian Academy of Science (Thomas Davies Grant). S.W. was funded by an ARC Future Fellowship (FT200100135) and supported by the ANU Future Scheme (35665). L.M. was funded by an ARC Discovery Early Career Researcher Award (DE170101165). A.S. was a recipient of the AINSE Honours Scholarship Program, and D.Y. and C.M. held an AINSE Postgraduate Research Award. P.K. was supported by a Netaji Subhas ICAR International Fellowship. The authors acknowledge the use of the ANU crystallisation facility. We thank Jana Sperschneider for assistance in generating some of the AlphaFold models. The authors acknowledge use of the ANSTO, Australian Synchrotron MX facility and thank the staff for their support. The coordinates and structure factors for Avr1 and Avr3 have been deposited in the Protein Data Bank with accession number 7T6A and 7T69, respectively.

Figure and Figure legends

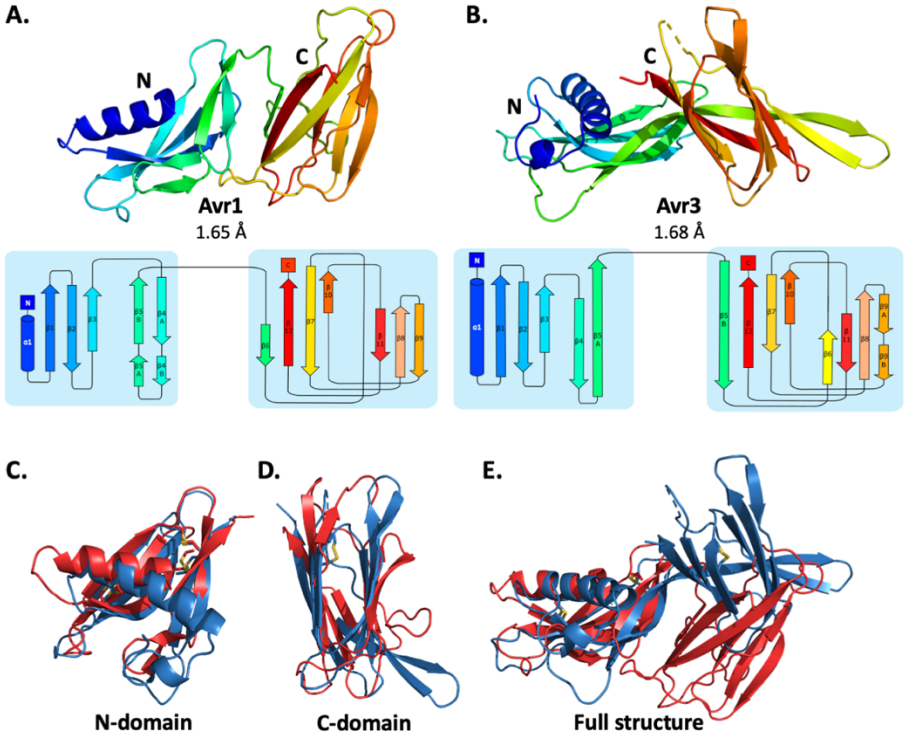


Fig 1. Crystal structures of Avr1 and Avr3 from *Fol* adopt a similar structural fold that is unique among fungal effectors. Ribbon diagrams of Avr1 in the top panels showing the dual-domain structural fold. Bottom panels showing secondary structure topology map of Avr1 **(A)** and Avr3 **(B)**. For both, the N-domain is shown on the left and the C-domain is shown on the right. The colours of the secondary structural elements match the colours depicted on the crystal structure. **(C-E)** Structural alignments of Avr1 (shown in red) and Avr3 (shown in blue) showing **(C)** N-domains alone, **(D)** C-domains alone and **(E)** full structures. Disulfide bonds are shown in yellow. Structural alignment was performed using the pairwise alignment function on the DALI server [34].

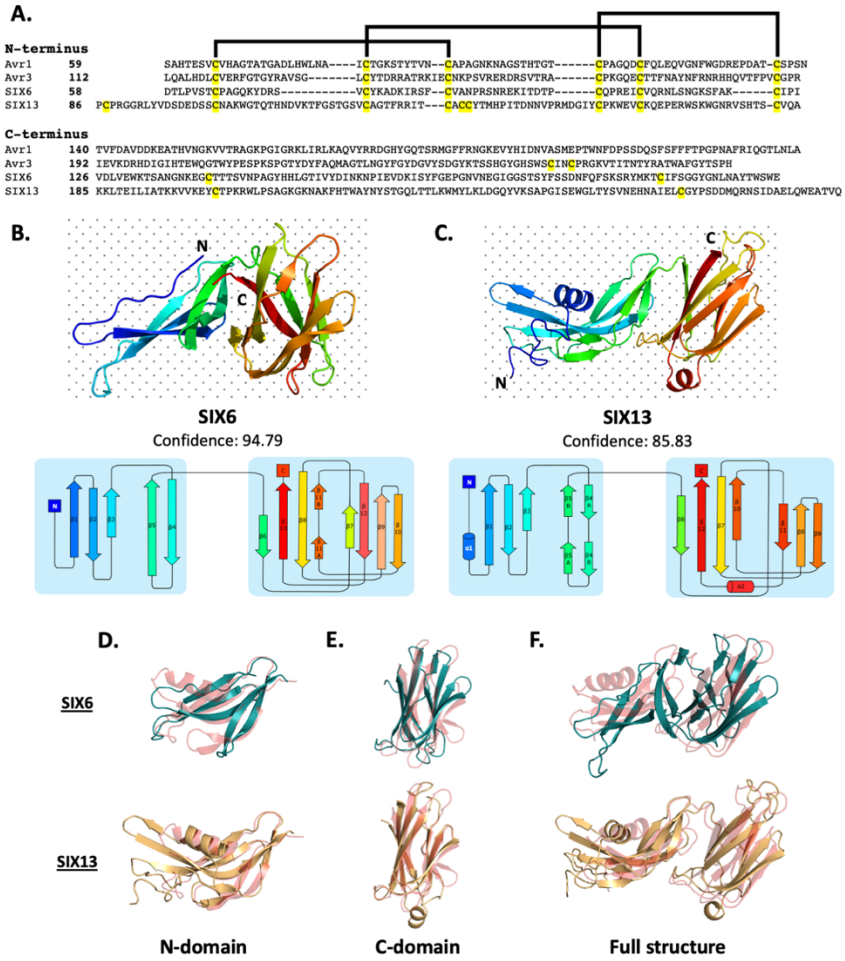


Fig 2. SIX6 and SIX13 are predicted FOLD effector family members. (A) Amino acid sequence alignment of Avr1, Avr3, SIX6 and SIX13 show a common cysteine spacing at the N-terminus. The alignment is split into the N-terminus (N-domain; top panel) and C-terminus (C-domain; bottom panel). Cysteine residues are highlighted in yellow, with the disulfide bonding connectivity, as determined by the crystal structures of Avr1 and Avr3, shown with black lines. Ribbon diagrams of the **(B)** SIX6 and **(C)** SIX13 model predicted by AlphaFold showing the dual-domain structural fold, transitioning from blue (N-terminus) to red (C-terminus). Bottom panel: Secondary structure topology maps of SIX6 and SIX13. The N-domain is shown on the left and the C-domain is shown on the right. The colours of the secondary structural elements match the colours depicted on the crystal structure. **(D-F)** Structural alignments of SIX6 (shown in teal) and SIX13 (shown in orange) with Avr1 (shown in red) showing **(D)** N-domains alone, **(E)** C-domains alone and **(F)** full structures. Structural alignment was performed using the pairwise alignment function on the DALI server [34].

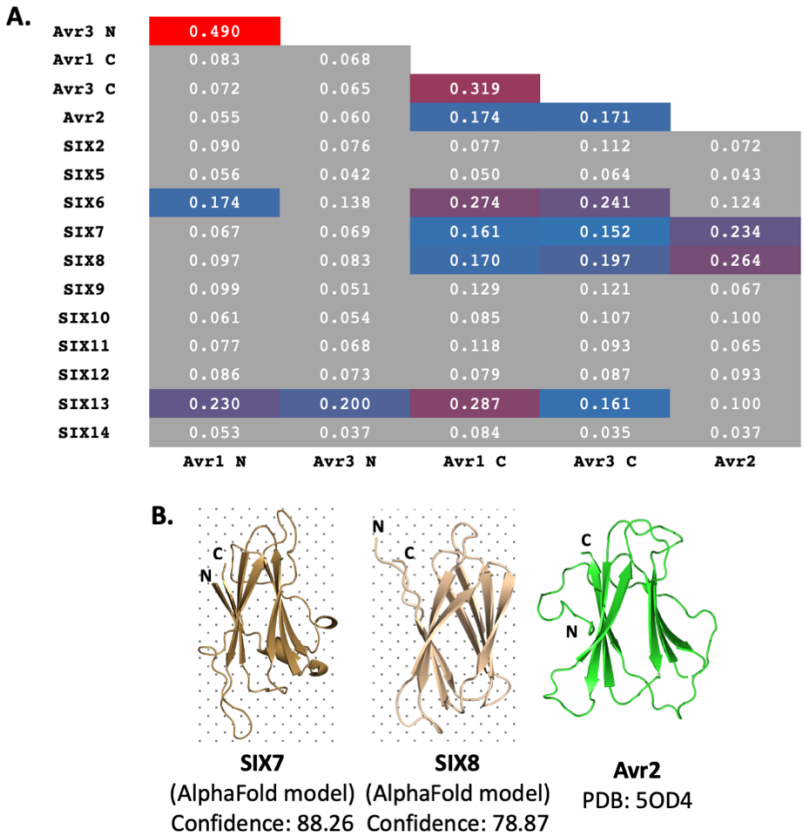


Fig 3. SIX7 and SIX8 are putative ToxA-like effectors. (A) Heat maps showing the structural similarity of the AlphaFold predicted models of the SIX effectors compared to experimentally determined SIX effector structures. Avr1 and Avr3 were split into structural units for the analysis (N or C domain). Structural similarity was determined by Q-scores, which take into account the RMSD and alignment length, using a cutoff of 0.15. Structural similarity (above the cutoff) is shown in a blue to red spectrum, with lower structural similarities shown in blue and higher structural similarities in red. Structural similarity searches were performed using the PDBeFold server [36]. **(B)** AlphaFold models of SIX7 and SIX8, putative members of the ToxA-like effector family. The crystal structure of Avr2 [25], another member of the ToxA-like effector family, is shown in green for comparison.

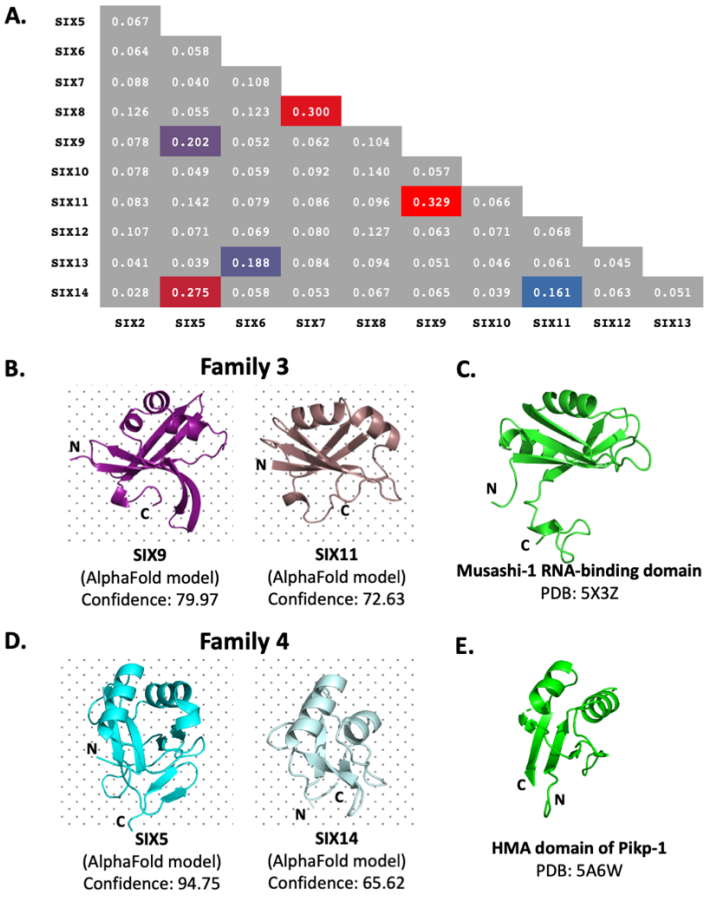


Fig 4. Identification of new putative structural families within the SIX effectors (A) A heat map showing the structural similarity of AlphaFold models of the SIX effectors in a structural pairwise alignment. Structural similarity was measured with Q-scores, which takes into account the RMSD and alignment length, with a cutoff of 0.15. Structural similarity (above the cutoff) is shown in a blue to red spectrum, with lower structural similarities shown in blue and higher structural similarities in red. **(B)** Cartoon representation of the AlphaFold models of SIX9 and SIX11, which are predicted to be structurally similar. The new putative structural class is assigned as 'Family 3'. Structural similarity searches show that SIX9 and SIX11 models have structural similarity to RNA interacting proteins. **(C)** The crystal structure of the Musashi-1 RNA-binding domain is shown for comparison [38]. **(D)** Cartoon representation of the AlphaFold models of SIX5 and SIX14, which are predicted to be structurally similar. The new putative structural class is assigned as 'Family 4'. Structural similarity searches show that SIX5 and SIX14 have structural similarity to heavy metal associated (HMA) domains. **(E)** The crystal structure of the HMA domain of Pikp-1 is shown for comparison [41]. Structural similarity searches were performed using the PDBeFold server [36].

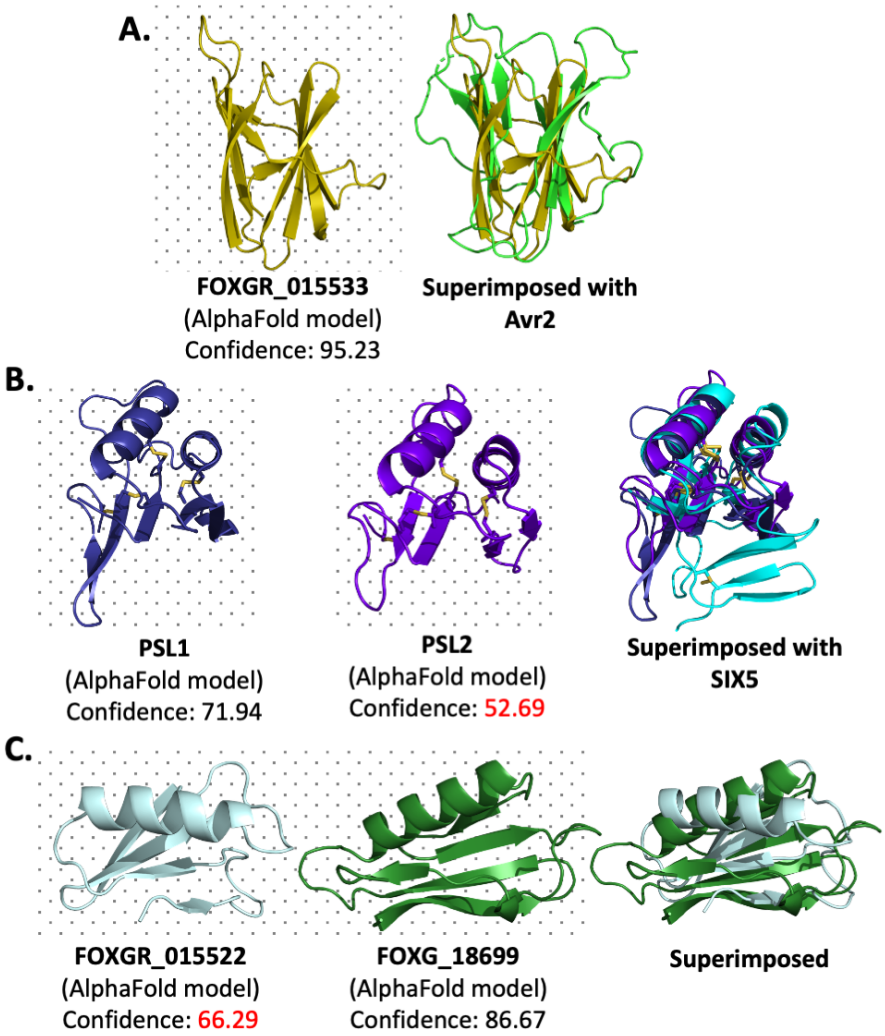


Fig 5. Extension of new putative structural families within *Fol* effectors and effector candidates. (A) Left panel: AlphaFold model of FOXGR_015533 is predicted to adopt a ToxA-like fold. Right panel: FOXGR_015533 superimposed with Avr2, a known ToxA-like structural effector. **(B)** Left and middle panel: AlphaFold models of PSL1 and PSL2 is predicted to be structurally similar to Family 4, represented by SIX5. Right panel: PSL1 and PSL2 superimposed with SIX5 model. **(C)** Left and middle panel: AlphaFold models of FOXGR_015522 and FOXG_18699 are predicted to be structurally similar. The new putative structural class is assigned as 'Family 5'. Right panel: FOXG_18699 superimposed with FOXGR_015522. Low confidence AlphaFold models are shown in red text.

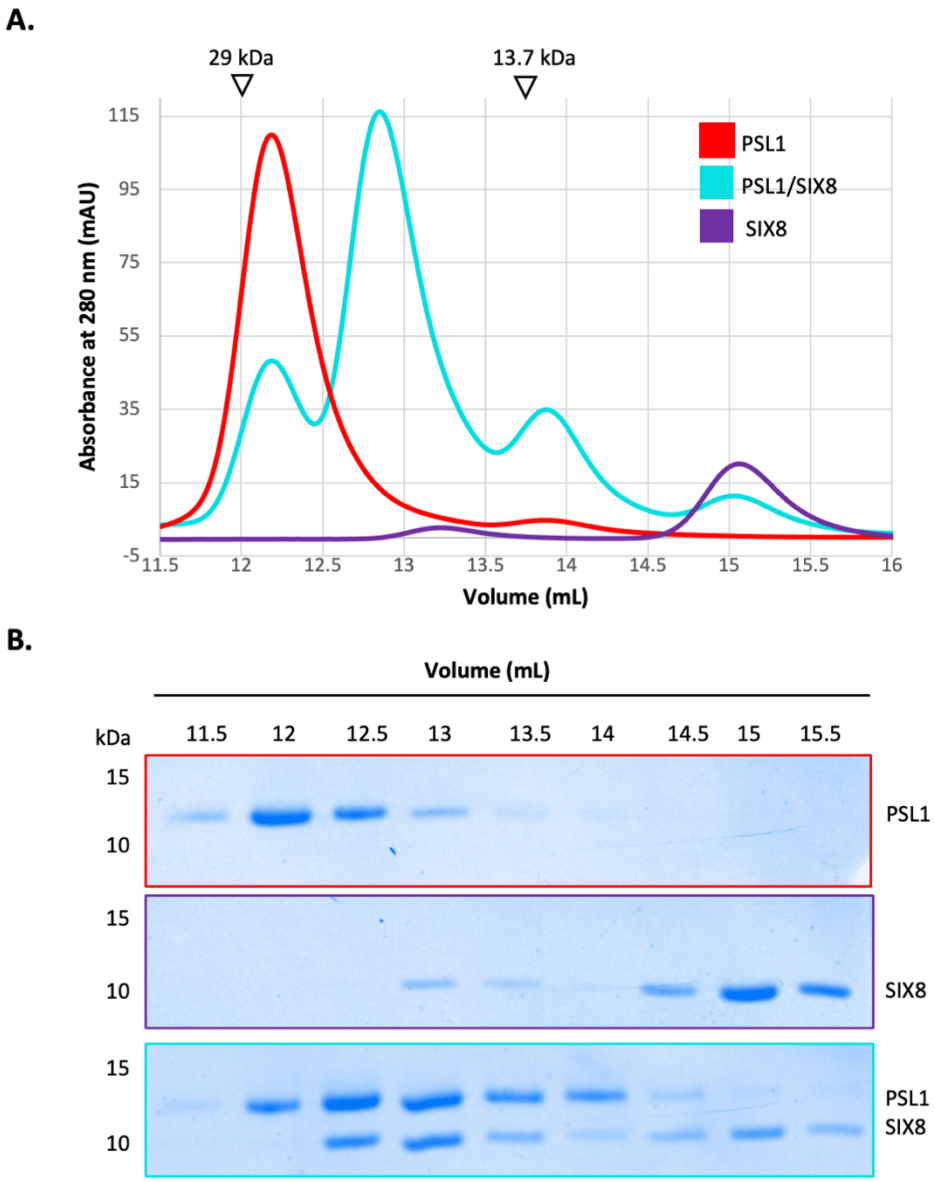


Fig 6. PSL1 (FOXGR_025399) and SIX8 interact *in vitro*. (A) Size exclusion chromatograms of PSL1 alone (red), SIX8 alone (purple), and PSL1 and SIX8 (cyan) following a 30 min incubation injected onto a Superdex S75 Increase 10/300 using 20 mM HEPES pH 7.5, 150 mM NaCl. Equal concentrations of the protein were used, but the absorbance of SIX8 is ~0.3 at 280 nm resulting in a smaller absorbance (peak height). Indicated sizes above the chromatogram are based on protein standards run under similar conditions as presented in the manufacturer's column guidelines. **(B)** Coomassie-stained SDS-PAGE gels depicting samples taken from 500 µL fractions corresponding to the volumes indicated above the gels. The left numbers on the gels indicate the molecular weights, and the right-hand side indicates the proteins present. The expected sixes of PSL1 and SIX8 are 10.8 kDa and 9.9 kDa, respectively. Full gels are available in Supplementary figure 8.

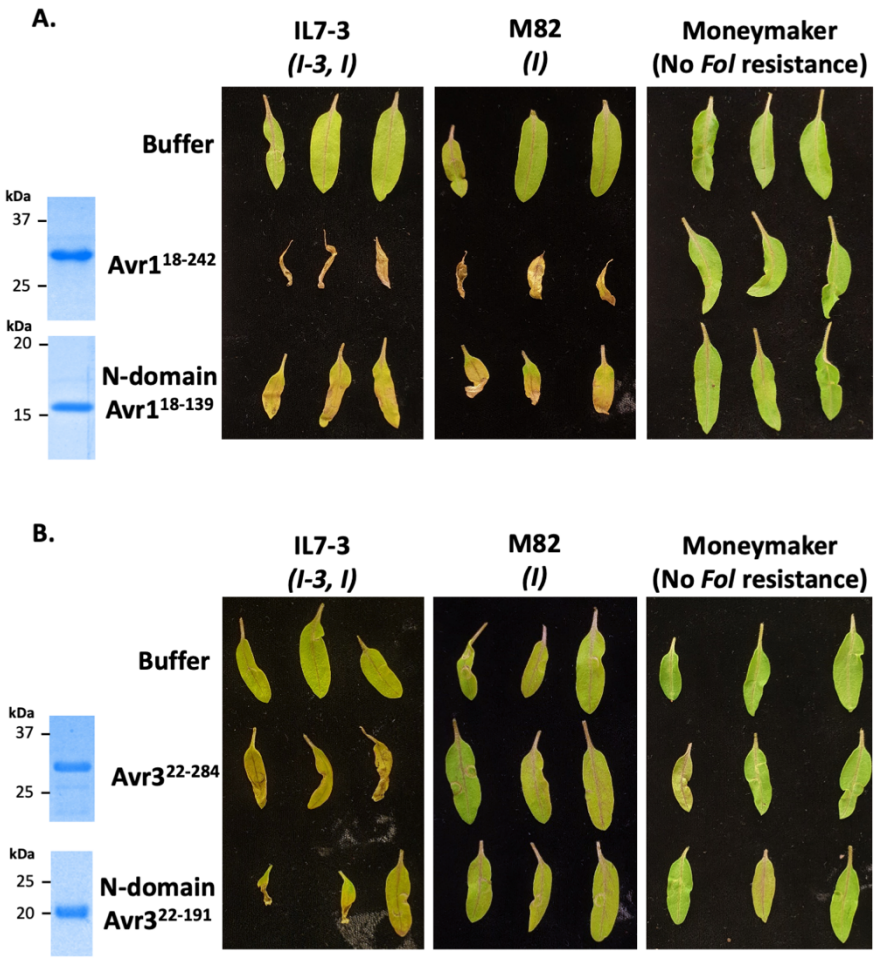
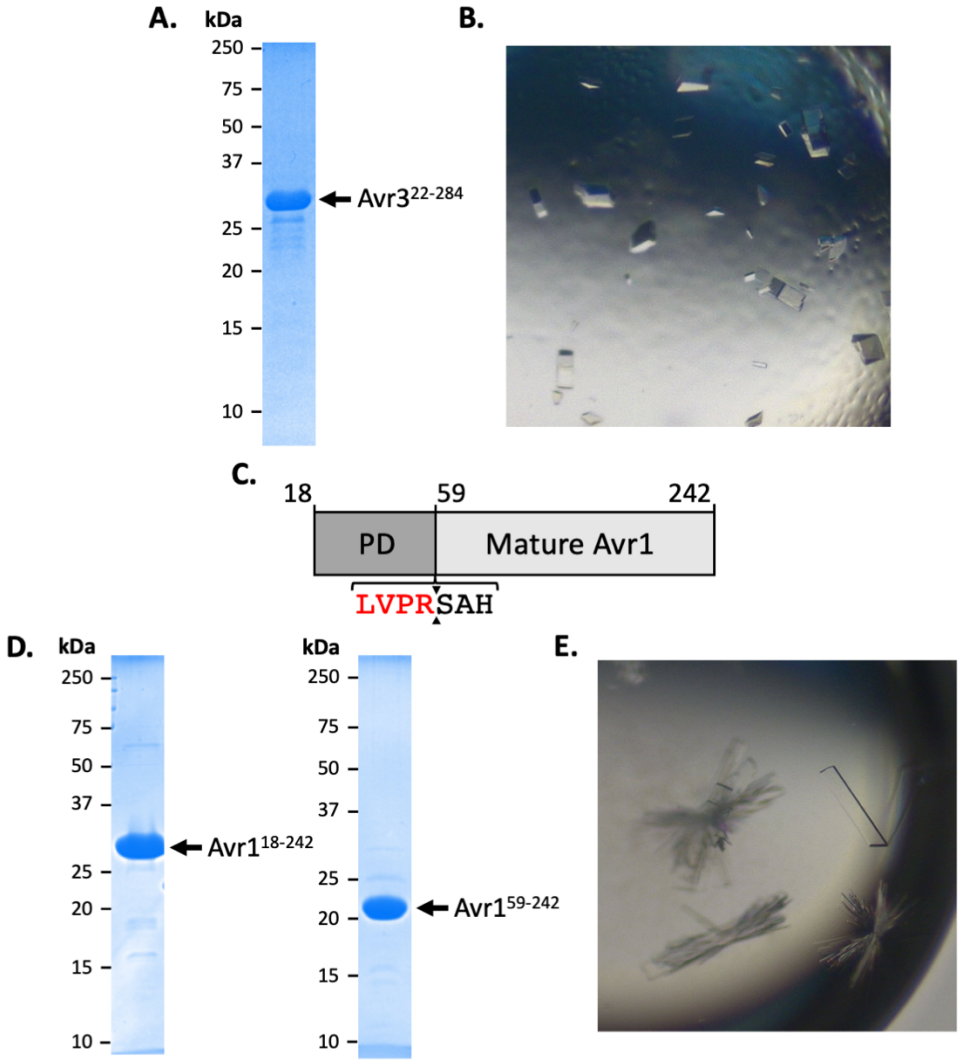
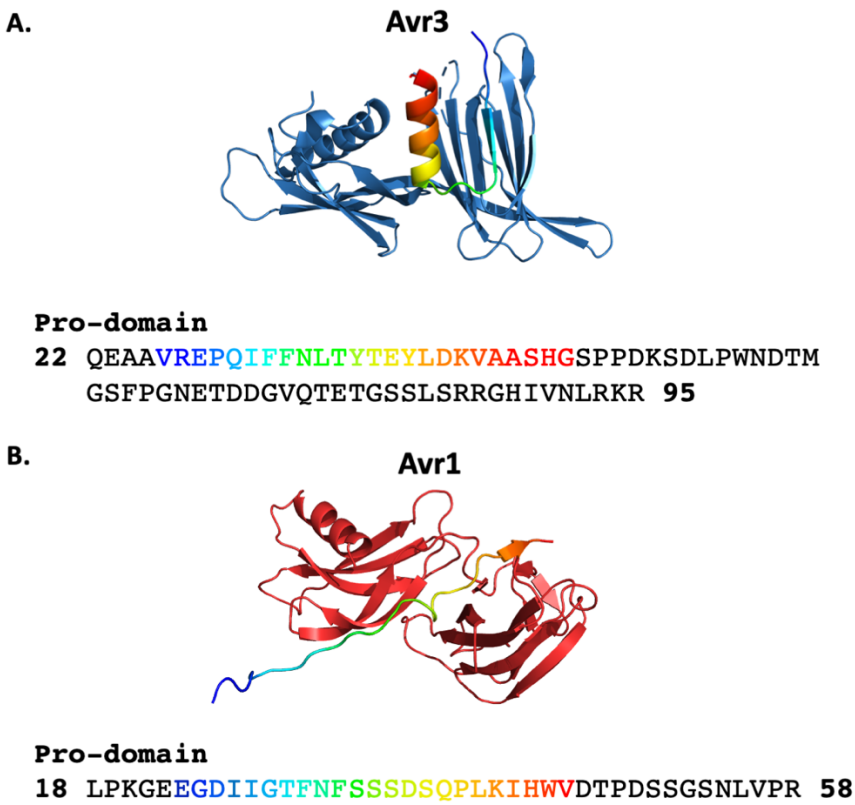


Fig 7. The I and I-3 immunity receptors can recognise the N-domain of Avr1 and Avr3, respectively. (A) The N-domain of Avr1 is sufficient for *I*-mediated cell death. Left panel: Coomassie-stained gel depicting the purified Avr1$^{18\text{-}242}$ and Avr1 N-domain used in the cotyledon infiltrations. **(B)** The N-domain of Avr3 is sufficient for *I-3*-mediated cell death. Left panel: Coomassie-stained gel depicting the purified Avr3$^{22\text{-}284}$ and Avr3 N-domain used in the cotyledon infiltrations. In both **(A)** and **(B)**, proteins (at a concentration of 0.1 mg/mL) and a buffer control were syringe-infiltrated into 10-14-day old tomato cotyledons from cultivars IL7-3 (containing *I-3* and *I*), M82 (containing *I*) and Moneymaker (no *Fol* resistance). Cotyledons were harvested and imaged 7 days post infiltration.

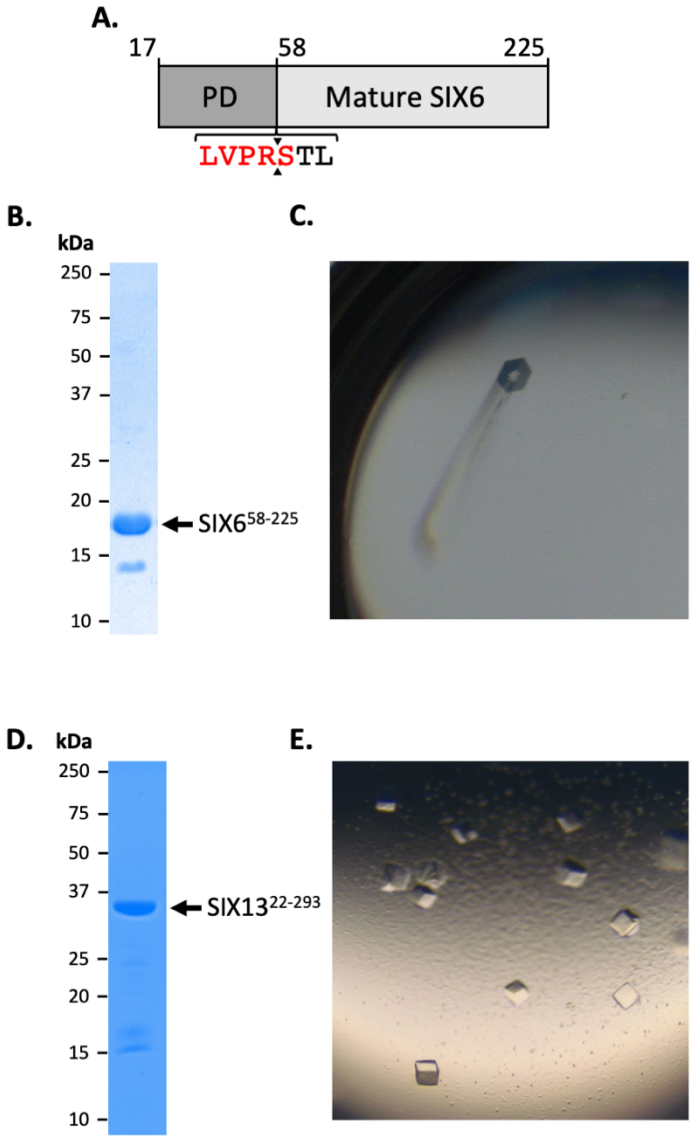
Supplementary Figures and Figure Legends



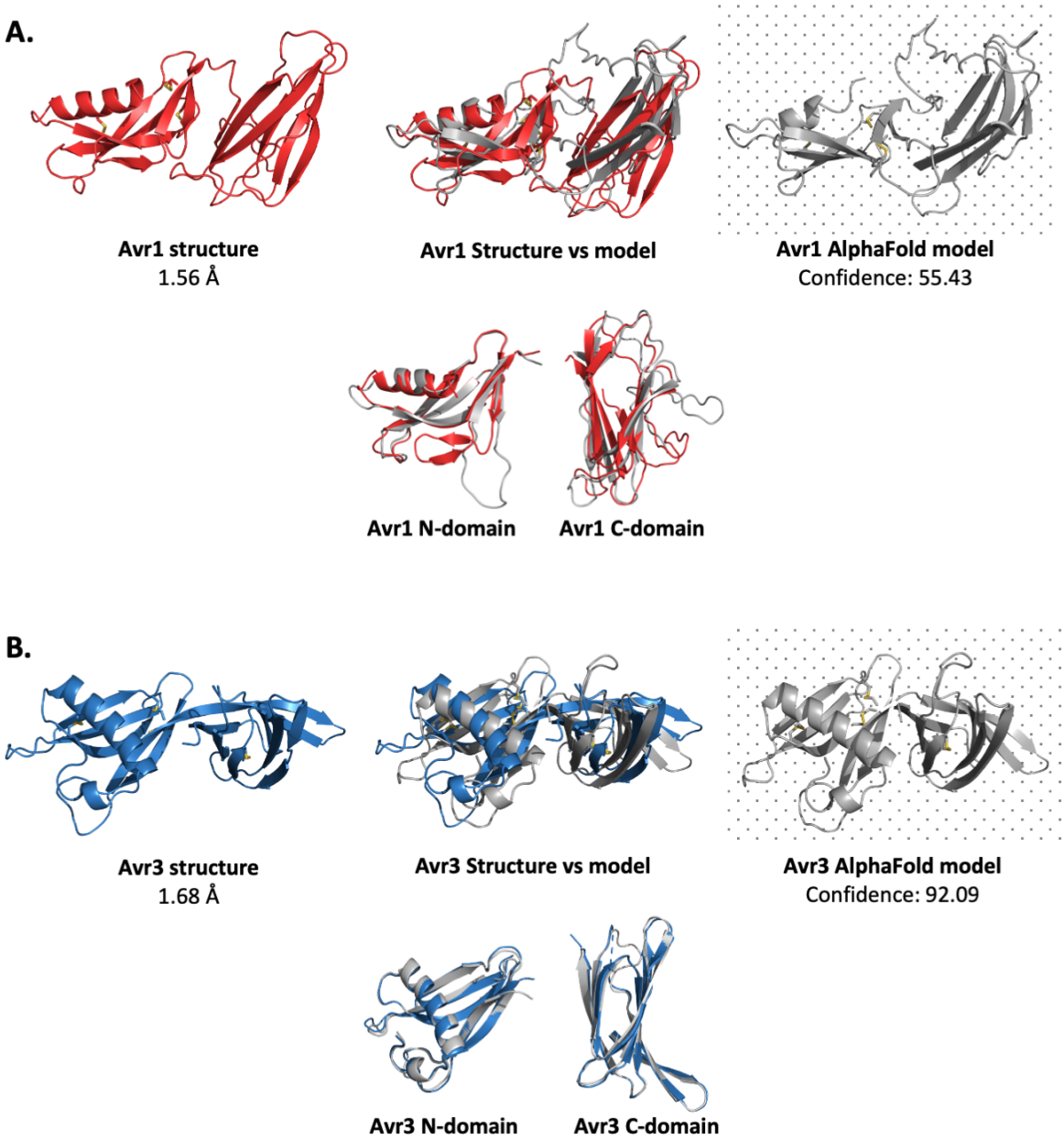
S1 Fig. Crystallisation of Avr1 and Avr3. (A) Coomassie-stained gel showing purified Avr3$^{22-284}$ used for crystallisation studies. **(B)** Optimised crystals of Avr3. Crystallisation was carried out using the hanging-drop vapour-diffusion method with a 1:1 ratio of protein to reservoir solution containing 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350. **(C)** Schematic of Avr1 engineered with an internal thrombin cleavage site replacing the Kex2 cleavage motif. The thrombin recognition motif is shown in red, with the cleavage site shown by black triangles **(D)** Coomassie-stained gel showing full length Avr1 (left panel) and mature Avr1 cleaved *in vitro* with thrombin (right panel). **(E)** Optimised crystals of Avr1$^{59-242}$. Crystallisation was carried out using the hanging-drop vapour-diffusion method with microseeding using a 1:1 ratio of protein to reservoir solution containing 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.5, and 17.5% (w/v) PEG 4000.



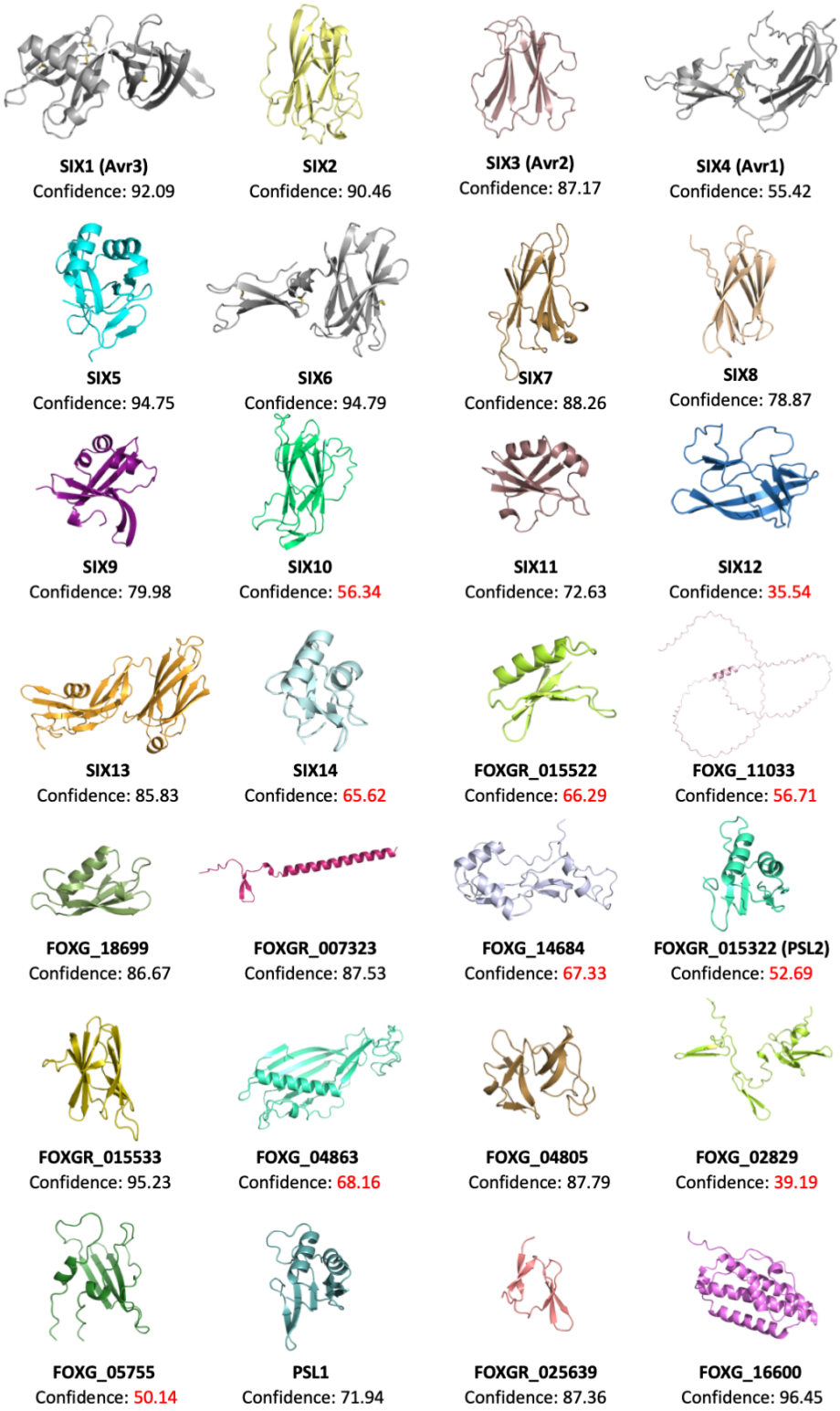
S2 Fig. The pro-domain is present in the crystal structures of Avr1 and Avr3. Top panels: The crystal structure of **(A)** Avr3 and **(B)** Avr1 with the pro-domain shown in rainbow. Bottom panels: The amino acid sequence of the pro-domain of Avr3 and Avr1 with residues observed in the electron density shown in rainbow text. Residues with no density observed are shown in black.



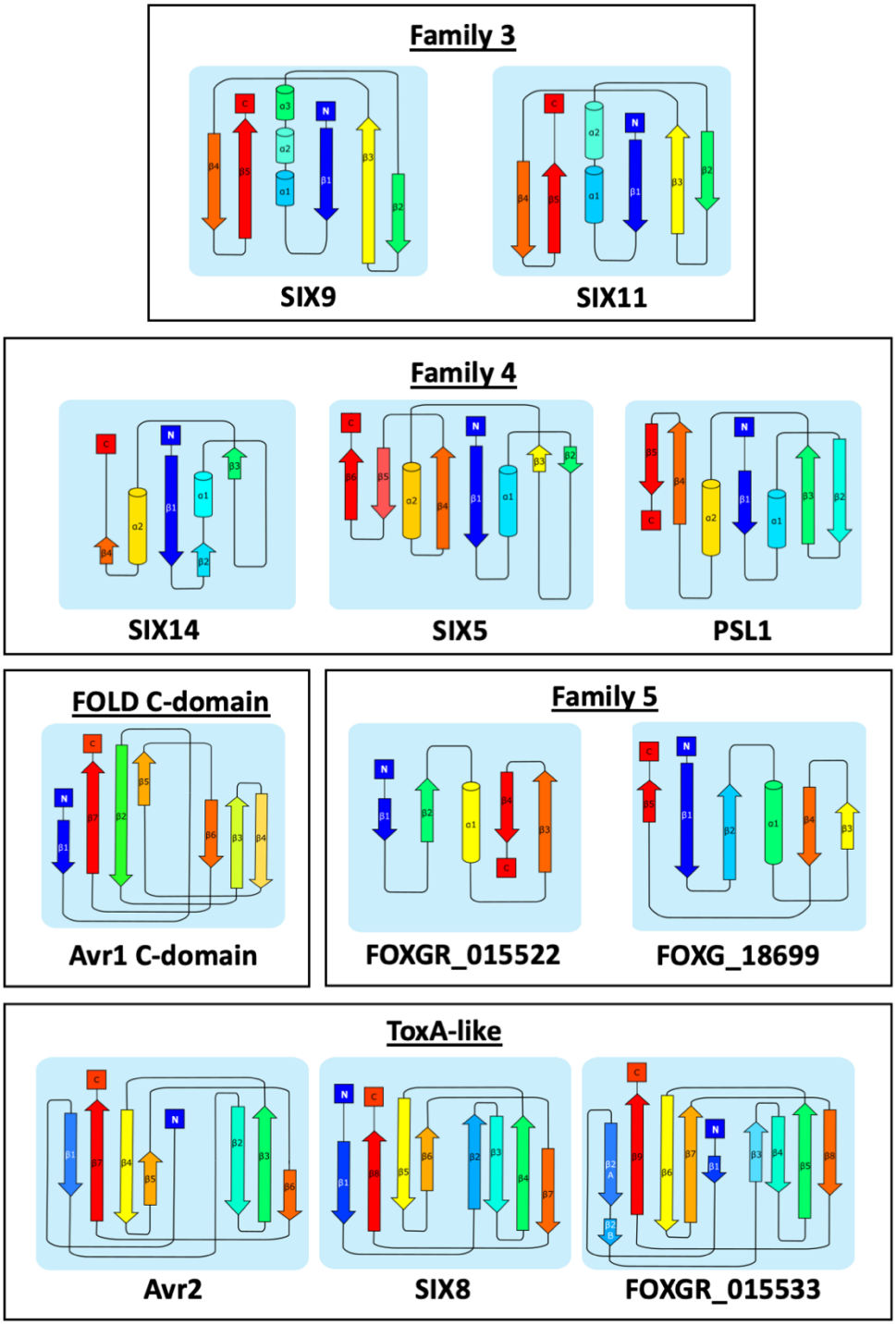
S3 Fig. Crystallisation of SIX6 and SIX13. (A) Schematic of SIX6 engineered with an internal thrombin cleavage site replacing the Kex2 cleavage motif. **(B)** Coomassie-stained gel showing mature SIX6 cleaved *in vitro* with thrombin. **(C)** Optimised crystals of SIX6. Crystal drops were set up using the hanging-drop vapour-diffusion method of 1:1 ratio of protein to reservoir solution containing 0.2 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5, 20% (w/v) PEG 3350. **(D)** Coomassie-stained gel showing SIX13 protein. **(E)** Optimised crystals of SIX13. Kex2 protease was added to the protein at a 1:200 protease to protein ratio prior to crystal tray setup. Crystal drops were set up using the hanging-drop vapour-diffusion method of 1:1 ratio of protein to reservoir solution containing 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5, 25% (w/v) PEG 3350.



S4 Fig. Comparison of AlphaFold models against the experimentally solved structures of Avr1 and Avr3. The crystal structures of **(A)** Avr1 and **(B)** Avr3 (top left panels) and AlphaFold models [35] (top right panels). Crystal structures and AlphaFold models of the full structures (top middle panels) and the separate domains (bottom panels) were superimposed using the pairwise function on the DALI server [34].



S5 Fig. AlphaFold models of all SIX effectors and effector candidates. Signal peptides were identified using SignalP-5.0 [90] and removed prior to amino acid sequences being input into AlphaFold [35]. Any putative pro-domains were identified by searching for a Kex2-like protease site [32] and removed. The sequence inputs used can be found in S4 Table.



S6 Fig. Secondary structure topology maps of representative SIX structural family members. The β-strands and α-helices are represented by arrows and cylinders, respectively. The secondary structural elements are coloured in rainbow, from blue at the N-terminus to red at the C-terminus.

A.

	Avr1	Avr3	SIX6	SIX13	Avr2	SIX7	SIX8	SIX9	SIX11	SIX5	SIX14	SIX2	SIX10	SIX12
	FOLD				ToxA-like			Family 3		Family 4				
FOXG_14684	0.028	0.041	0.036	0.021	0.058	0.070	0.077	0.032	0.091	0.045	0.059	0.040	0.048	0.020
FOXG_04805	0.035	0.035	0.044	0.026	0.046	0.050	0.074	0.078	0.049	0.064	0.050	0.064	0.067	0.063
FOXG_04863	0.029	0.037	0.032	0.028	0.030	0.042	0.049	0.042	0.040	0.027	0.018	0.031	0.038	0.034
FOXG_05755	0.040	0.048	0.062	0.043	0.139	0.148	0.129	0.077	0.057	0.048	0.043	0.096	0.074	0.149
FOXG_16600	0.008	0.013	0.000	0.009	0.004	0.009	0.041	0.019	0.026	0.008	0.030	0.005	0.005	0.000
FOXG_02829	0.040	0.030	0.028	0.029	0.031	0.041	0.039	0.031	0.034	0.032	0.018	0.029	0.020	0.041
FOXG_11033	0.005	0.000	0.000	0.003	0.000	0.004	0.000	0.006	0.008	0.009	0.011	0.000	0.000	0.000
FOXG_18699	0.060	0.057	0.094	0.062	0.109	0.092	0.177	0.166	0.185	0.151	0.186	0.064	0.077	0.071
FOXGR_015522	0.076	0.082	0.085	0.060	0.094	0.063	0.142	0.153	0.154	0.091	0.118	0.053	0.078	0.109
FOXGR_025639	0.053	0.058	0.046	0.032	0.083	0.071	0.083	0.073	0.079	0.063	0.050	0.050	0.067	0.076
FOXGR_007323	0.021	0.018	0.021	0.014	0.029	0.016	0.028	0.026	0.026	0.044	0.031	0.027	0.027	0.024
FOXGR_015322	0.062	0.074	0.056	0.062	0.068	0.061	0.083	0.148	0.193	0.210	0.251	0.077	0.066	0.054
FOXGR_015533	0.107	0.099	0.141	0.112	0.235	0.223	0.366	0.084	0.082	0.065	0.078	0.112	0.125	0.147
PSL1	0.035	0.064	0.078	0.047	0.060	0.049	0.073	0.160	0.175	0.275	0.279	0.079	0.073	0.052

B.

	ToxA	ToxB	BEC1054	AvrLm4-7	Tox3	AvrP	CfAvr4	AvrM	NLP
	ToxA-like	MAX	RALPH	LARS	Tox3-like	Zinc finger	CBM14-like	WY-like	Actinoporin-like
FOLD: Avr1	0.103	0.062	0.037	0.030	0.041	0.045	0.052	0.008	0.034
FOLD: Avr3	0.088	0.052	0.033	0.040	0.043	0.055	0.028	0.008	0.042
FOLD: SIX6	0.108	0.081	0.056	0.039	0.036	0.052	0.120	0.000	0.039
FOLD: SIX13	0.091	0.067	0.035	0.040	0.049	0.037	0.048	0.005	0.044
ToxA-like: Avr2	0.304	0.121	0.073	0.064	0.047	0.070	0.076	0.000	0.047
ToxA-like: SIX7	0.304	0.127	0.081	0.048	0.076	0.074	0.071	0.010	0.068
ToxA-like: SIX8	0.389	0.201	0.099	0.061	0.092	0.084	0.089	0.000	0.084
Family 3: SIX9	0.077	0.122	0.060	0.150	0.083	0.094	0.063	0.024	0.046
Family 3: SIX11	0.097	0.137	0.084	0.145	0.041	0.065	0.095	0.014	0.055
Family 4: SIX5	0.052	0.110	0.061	0.054	0.043	0.038	0.065	0.020	0.043
Family 4: SIX14	0.052	0.151	0.073	0.009	0.013	0.038	0.070	0.032	0.036
SIX2	0.085	0.174	0.076	0.063	0.173	0.061	0.074	0.000	0.084
SIX10	0.107	0.086	0.079	0.052	0.065	0.051	0.072	0.000	0.058
SIX12	0.135	0.110	0.084	0.039	0.117	0.075	0.085	0.000	0.055

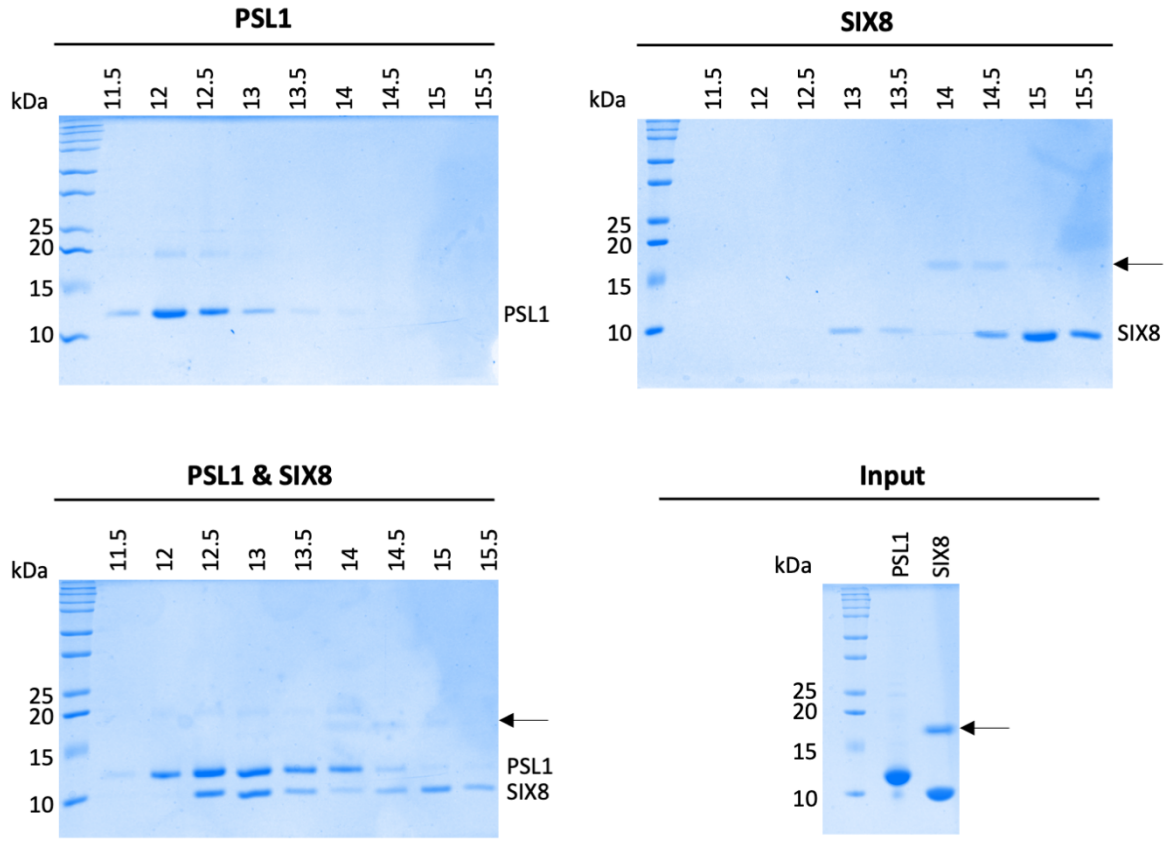
C.

	ToxA	ToxB	BEC1054	AvrLm4-7	Tox3	AvrP	CfAvr4	AvrM	NLP
	ToxA-like	MAX	RALPH	LARS	Tox3-like	Zinc finger	CBM14-like	WY-like	Actinoporin-like
FOXG_14684	0.070	0.062	0.025	0.062	0.066	0.043	0.059	0.017	0.038
FOXG_04805	0.060	0.078	0.066	0.034	0.044	0.070	0.093	0.000	0.034
FOXG_04863	0.039	0.042	0.041	0.077	0.046	0.048	0.032	0.013	0.033
FOXG_05755	0.149	0.135	0.095	0.040	0.073	0.083	0.097	0.013	0.044
FOXG_16600	0.035	0.000	0.040	0.045	0.004	0.014	0.011	0.063	0.015
FOXG_02829	0.041	0.039	0.051	0.026	0.031	0.027	0.041	0.000	0.017
FOXG_11033	0.006	0.000	0.011	0.000	0.000	0.005	0.005	0.008	0.004
FOXG_18699	0.126	0.107	0.085	0.130	0.061	0.082	0.096	0.017	0.051
FOXGR_015522	0.117	0.158	0.079	0.145	0.066	0.094	0.094	0.025	0.037
FOXGR_025639	0.075	0.173	0.097	0.060	0.060	0.157	0.098	0.000	0.041
FOXGR_007323	0.019	0.047	0.038	0.016	0.012	0.043	0.041	0.077	0.014
FOXGR_015322	0.054	0.109	0.067	0.121	0.046	0.044	0.075	0.043	0.050
FOXGR_015533	0.265	0.158	0.121	0.066	0.092	0.068	0.085	0.000	0.058
PSL1	0.059	0.114	0.078	0.091	0.051	0.051	0.068	0.030	0.047

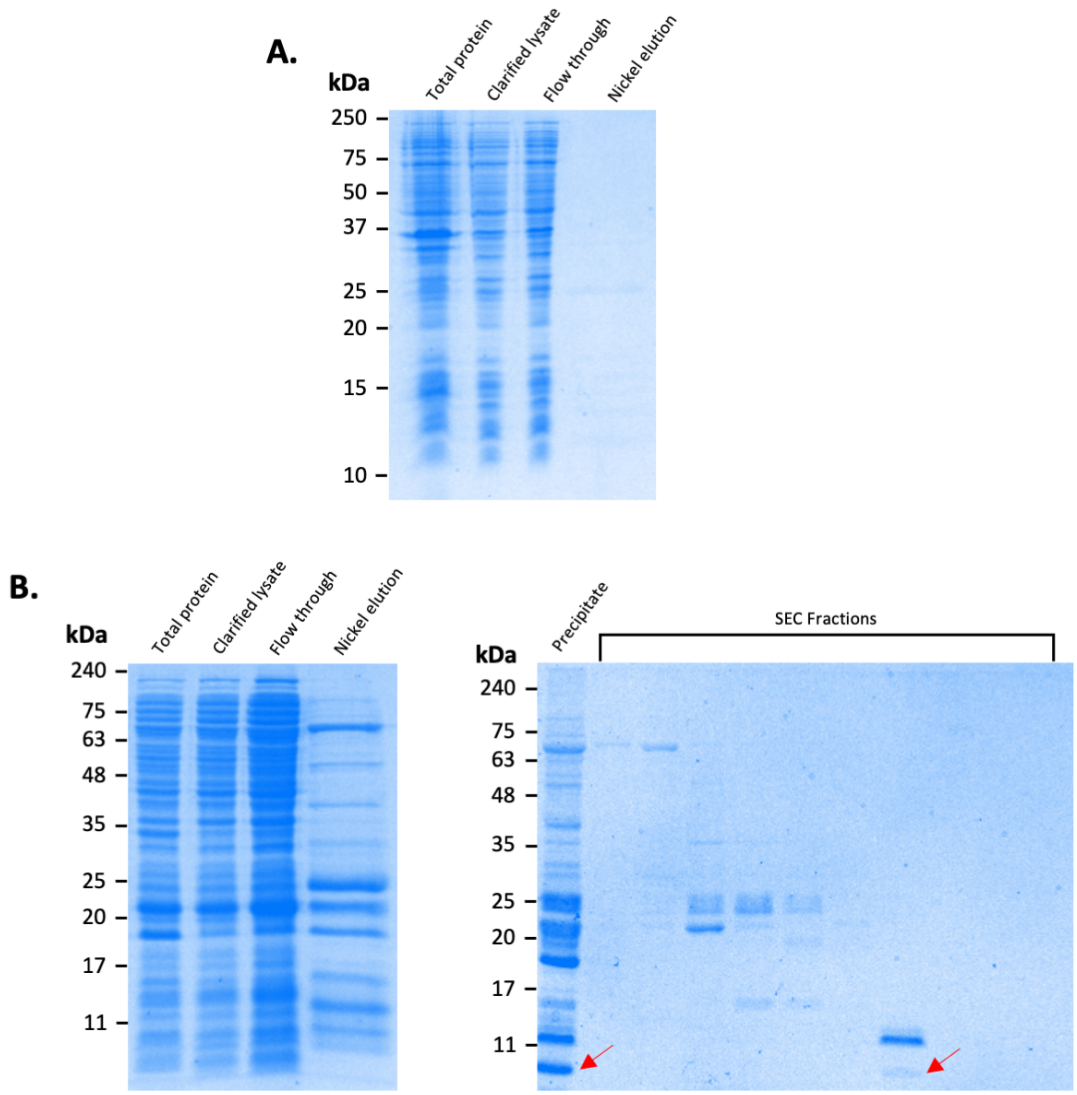
D.

	FOXG_14684	FOXG_04805	FOXG_04863	FOXG_05755	FOXG_16600	FOXG_02829	FOXG_11033	FOXG_18699	FOXGR_015522	FOXGR_025639	FOXGR_007323	FOXGR_015322	FOXGR_015533
FOXG_04805	0.039												
FOXG_04863	0.019	0.034											
FOXG_05755	0.054	0.082	0.039										
FOXG_16600	0.023	0.010	0.029	0.016									
FOXG_02829	0.027	0.042	0.019	0.037	0.006								
FOXG_11033	0.008	0.000	0.000	0.009	0.005	0.000							
FOXG_18699	0.085	0.074	0.087	0.090	0.034	0.044	0.011						
FOXGR_015522	0.095	0.070	0.061	0.068	0.029	0.045	0.014	0.245					
FOXGR_025639	0.059	0.104	0.032	0.085	0.000	0.059	0.000	0.105	0.128				
FOXGR_007323	0.030	0.051	0.009	0.000	0.104	0.022	0.000	0.035	0.043	0.072			
FOXGR_015322	0.075	0.052	0.042	0.042	0.030	0.032	0.011	0.195	0.146	0.065	0.030		
FOXGR_015533	0.050	0.076	0.047	0.091	0.000	0.061	0.000	0.108	0.125	0.090	0.032	0.084	
PSL1	0.064	0.062	0.031	0.060	0.026	0.036	0.008	0.200	0.125	0.074	0.037	0.609	0.081

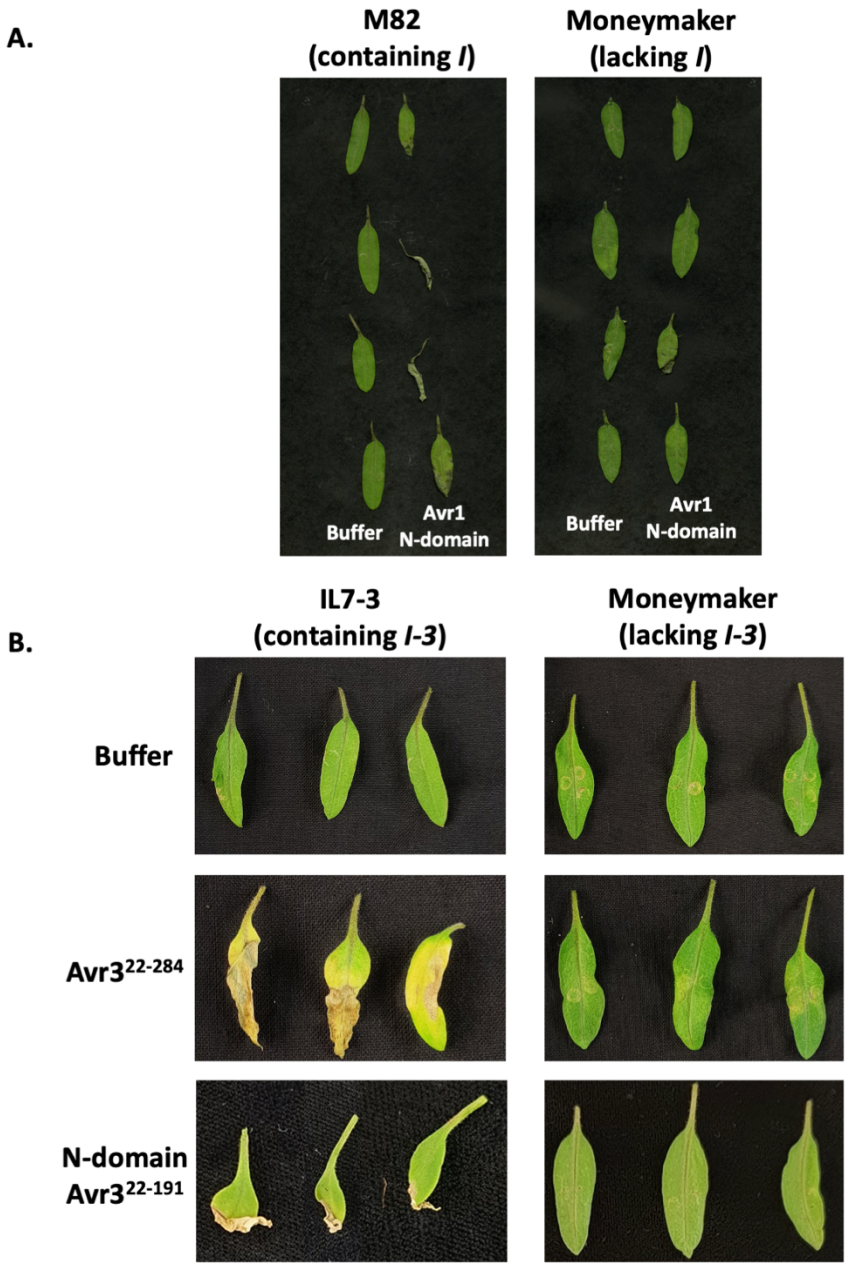
S7 Fig. (A) A heat map showing the structural similarity of AlphaFold models of the *Fol* effector candidates in a structural pairwise alignment against putative SIX effector structural families. Heat maps showing the structural similarity of **(B)** putative SIX effector structural families and **(C)** *Fol* effector candidates against solved effector structures and known effector families. **(D)** A heat map showing the structural similarity of AlphaFold models of the *Fol* effector candidates in a structural pairwise alignment. Structural similarities were measured with Q-scores with a cutoff of 0.15. Q-scores are shown in a blue to red spectrum, with lower structural similarities shown in blue and higher structural similarities in red.



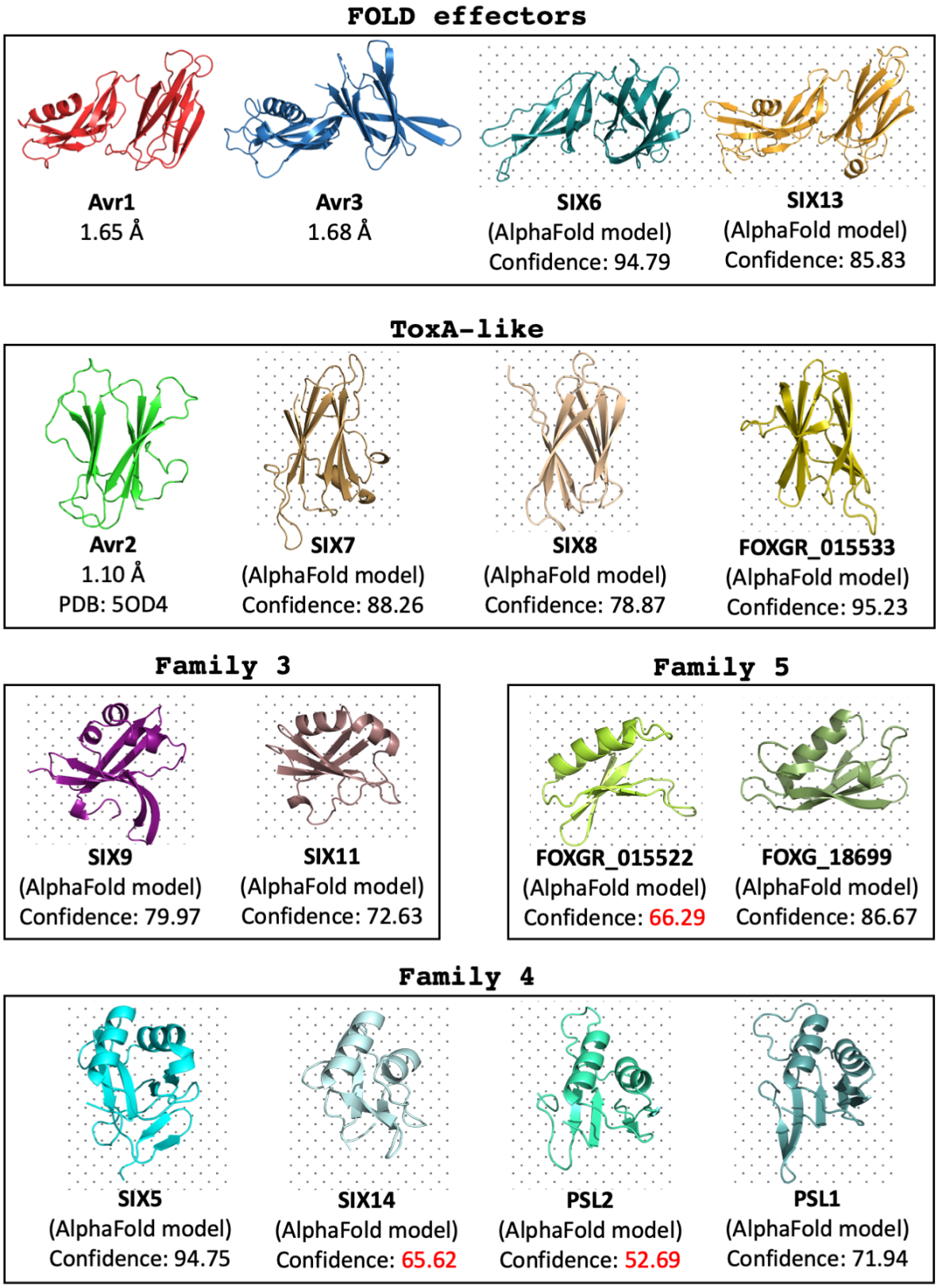
S8 Fig. PSL1 (FOXGR_025399) and SIX8 interact in vitro over size exclusion chromatography. Coomassie-stained SDS-PAGE gels showing full gels presented in Figure 6. Numbers above the gels indicate the volume across the size exclusion column (Superdex 75 Increase) when the sample was taken. The expected sixes of PSL1 and SIX8 are 10.8 kDa and 9.9 kDa, respectively. The black arrow indicates a contaminant band present in the purified SIX8 protein, which is unaffected by the interaction with PSL1.



S9 Fig. The C-domain of Avr1 and Avr3 cannot be produced in a pure, soluble form. Coomassie-stained SDS-PAGE gels showing total and soluble (Clarified lysate) proteins, and proteins not captured (Flow through) and captured (Nickel elution) by immobilised metal affinity chromatography (IMAC) from (A) BL21(DE3) E. coli expressing 6xHisGB1-Avr1$^{140-242}$ and (B) SHuffle E. coli expressing 6xHisGB1-Avr3$^{192-284}$. (C) Coomassie-stained SDS-PAGE gel showing insoluble precipitate and size exclusion chromatography fractions of Avr3$^{192-284}$ after tag cleavage. Red arrow point to bands corresponding to of Avr3$^{192-284}$.



S10 Fig. Replicate data showing the I and I-3 immunity receptors can recognise the N-domain of Avr1 and Avr3, respectively. (A) Avr1 N-domain (at a concentration of 0.1 mg/mL) and a buffer control were syringe-infiltrated into 14-day old tomato cotyledons from cultivars M82 (containing *I*) and Moneymaker (lacking *I*). Cotyledons were harvested and imaged 4 days post infiltration. **(B)** Avr3$^{22-284}$ and Avr3 N-domain (at a concentration of 0.1 mg/mL) and a buffer control were syringe-infiltrated into 14-day old tomato cotyledons from the cultivars IL7-3 (containing *I-3*) and Moneymaker (lacking *I-3*). Cotyledons were harvested and imaged 7 days post infiltration.



S11 Fig. Summary of the different putative SIX effector and candidate effector structural families. Models were generated using AlphaFold [35] and grouped based on the structural similarity of structures and models assessed with PDBeFold with a 0.15 Q-score cutoff [36].

```
SIX5     18  RDHQYCACQSGSGDSIDIDATTQLQNDNSKSYLWAQTSPAYWFADRHKPGPRFAGIYLKAA
SIX12    27  ---SSCLSVGPKGISNQNACVCGGQCVMKDLVVARR-----KVCCEYT-VQIQGGWPVLAQ
PSL1     18  EDWDRCRCMKYPETGTPNDCATIKACGSGKHRAISISEEKGDIWCEKTDVAISG--PEFY-
PSL2     18  EDWDQCRCMKYPSTGTPNDCATIKACGSGKHRAISIYKN-GDIWCEKTDVAING--PEFY-
SIX14    18  QRILGCRMPNGSLNPSPNICNQAGGSFRSGS----------RGCC--TRNTRDG--PVVTE

SIX5     79  NGKIDGDTFYNLCINNGGADSTCFDCSKSHQVRNVIYCDAA
SIX12    79  SRCVYGSTGA-----NGG--SCSGDNVSLAWWLNYEPEVKSTDPTCIFAKPKLCHS
PSL1     76  RTC-YGLLQD--PKPNSEADSCCTRWVDGVVVQSDGCFK
PSL2     75  RTC-YDLLQD--PKPNSEADSCCIKGDRA----SDGCFK
SIX14    65  SRFISGCNK------NGGFVSSKEILATSC
```

S12 Fig. Amino acid sequence alignment of SIX12 against Family 4 members reveals a similar cysteine spacing. All protein sequences have their signal peptides removed. The cysteine residues are highlighted in yellow and groups of two or more amino acid residues shared with SIX12 are highlighted in grey.

Supplementary tables

S1 Table. X-ray data collection, structure solution and refinement statistics for Avr1 and Avr3

	Avr1 Bromide soak (SAD)	Avr1 Native (MR)	Avr3 Bromide soak (SAD)	Avr3 Native (MR)
Data Collection				
Detector	Dectris EIGER 2 9M	Dectris EIGER 16M	Dectris EIGER 2 9M	Dectris EIGER 16M
Wavelength (Å)	0.91946	0.95373	1.45864	0.95373
Space group	P 1 21 1	P 1 21 1	C 2 2 21	C 2 2 21
Unit cell	69.87 38.24 80.10 90 103.56 90	70.00 40.34 81.30 90 104.54 90	54.68 79.93 117.12 90 90 90	54.86 80.13 117.37 90 90 90
Average mosaicity (°) [b]	0.00	0.08	0.00	0.06
Resolution (Å)	46.11 - 2.12 (2.18 - 2.12)	39.35 - 1.65 (1.68 - 1.65)	45.13 - 2.46 (2.56 - 2.46)	45.27 - 1.68 (1.71 - 1.68)
Total no. of reflections	1247183 (94443)	295538 (14086)	1426660 (157248)	280003 (13354)
No. of unique reflections	23868 (1868)	53375 (2645)	9668 (1079)	29849 (1458)
Completeness (%)	99.8 (98.1)	99.9 (100)	99.9 (99.4)	99.8 (96.8)
Multiplicity	52.3 (50.6)	5.5 (5.3)	147.6 (145.7)	9.4 (9.2)
Anomalous completeness	99.5 (96.8)	-	99.8 (98.7)	-
Anomalous multiplicity	26.0 (25.2)	-	77.0 (75.3)	-
Mean $I/s(I)$	24.0 (4.9)	13.1 (1.5)	25.1 (5.4)	14.5 (1.8)
*R*merge	0.157 (0.850)	0.074 (0.990)	0.215 (1.258)	0.072 (0.901)
*R*meas [c]	0.158 (0.858)	0.082 (1.010)	0.216 (1.262)	0.076 (0.953)
*R*pim [d]	0.022 (0.117)	0.034 (0.470)	0.018 (0.102)	0.025 (0.308)
$CC_{1/2}$ [b]	0.999 (0.950)	0.999 (0.490)	1.0 (0.976)	0.999 (0.832)
Matthews coefficient ($Å^3$ Da^{-1}) [e]	2.60	2.78	2.13	2.13
Refinement				
Resolution range (Å)	-	39.35 - 1.65	-	45.27 - 1.68
R_{work} (%) [g]	-	0.1685 (0.2255)	-	0.1688 (0.1846)
R_{free} (%) [h]	-	0.2110 (0.2899)	-	0.2177 (0.2704)
No. of non-H atoms				
Total	-	3647	-	1882
Macromolecules	-	3214	-	1713
Ligand	-	20	-	10
Water	-	413	-	159
Average B-factor ($Å^2$)	-	26.06	-	28.27
RMSD from ideal geometry				
Bond lengths (Å)	-	0.005	-	0.012
Bond angles (°)	-	0.74	-	1.18
Ramachandran plot, residues in (%)[i]				
Favoured regions	-	97.77	-	95.57
Allowed regions	-	2.23	-	4.43
Outlier regions	-	0.00	-	0.00

[a] The values in parentheses are for the highest-resolution shell.
[b] Calculated with AIMLESS [82].
[c] Rmeas = $\Sigma_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\ \Sigma i|Ii(hkl)-|/\ \Sigma hkl\Sigma iIi(hkl)$, where $Ii(hkl)$ is the intensity of the *i*th measurement of an equivalent reflection with indices *hkl*.
[d] Rpim = $\Sigma_{hkl}\{1/[N(hkl)-1]\}^{1/2}\ \Sigma i|Ii(hkl)-|/\ \Sigma hkl\Sigma iIi(hkl)$.
[e] Calculated with MATTHEWS_COEF within the CCP4 suite [83].
[f] Generated by Crank pipeline in the CCP4 suite [83-85].
[g] $R_{work} = \Sigma_{hkl}\ ||F_{obs}|-|F_{calc}||/\Sigma_{hkl}\ |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.
[h] R_{free} is equivalent to R_{work} but calculated with reflections (5%) omitted from the refinement process.
[i] Calculated with MolProbity.

S2 Table. Gene sequences used in this study

S2 Table. Gene sequences used in this study

Name	Sequence
SIX6	TAGGTCTCCAATGGGTCCCTTAGCCCAAACAGAATCCGAGTCGGCAGACGTCGCTGAACATAC AATCAATTATATCGACATTGCCCCTGAAGAATTTGAACCGCCCAAAGCTAATTTGTCATCTCTG GTGAGTCGTGACACGCTTCCTGTCAGTACCTGTCCTGCGGGTCAGAAATACGATCGTTCCGTGT GTTACAAGGCAGACAAAATTCGTAGCTTTTGTGTCGCAAACCCTCGTAGCAACCGTGAGAAGA TTACCGACACACCGTGTCAGCCCCGTGAAATCTGTGTGCAACGCAATCTTTCCAACGGCAAGA GTTTCGCTAAGTGTATCCCCATTGTAGACCTGGTGGAATGGAAGACATCCGCAAATGGGAATA AAGAAGGCTGTACTACAACGTCCGTGAATCCGGCTGGGTACCATCACCTTGGTACTATTGTTTA CGATATCAATAAGAATCCTATCGAAGTTGATAAAATCTCGTACTTCGGCGAGCCGGGAAATGT AAACGAGGGCATTGGTGGCAGCACAAGCTATTTTAGTAGTGACAACTTTCAATTTTCTAAGTCC CGCTACATGAAAACTTGTATTTTCAGTGGTGGGTACGGGAATCTTAACGCCTATACGTGGAGCT GGGAATCTTGGAGACCGT
SIX6Thrombin	TAGGTCTCCAATGGGTCCCCTGGCCCAGACGGAATCGGAGAGTGCTGACGTAGCTGAACATAC GATTAACTATATCGACATTGCGCCTGAAGAGTTCGAACCGCCGAAGGCTAATCTGTCATCTTTA GTACCTCGTTCAACCTTGCCGGTGTCAACGTGTCCAGCAGGCCAGAAGTATGACCGTAGCGTA TGTTACAAAGCGGACAAAATTCGCTCCTTCTGTGTCGCTAACCCGCGCAGCAATCGTGAAAAG ATCACGGATACTCCATGTCAGCCGCGCGAAATCTGTGTCCAACGTAACTTATCAAACGGGAAG AGTTTCGCTAAATGTATCCCTATTGTTGATCTTGTAGAATGGAAAACCAGTGCGAACGGCAACA AAGAGGGGTGTACCACCACGTCCGTAAACCCAGCCGGCTATCACCATTTAGGGACCATCGTCT ATGATATTAATAAAAATCCAATTGAGGTGGACAAGATCAGCTATTTTGGCGAACCAGGCAACG TCAACGAAGGAATTGGTGGTTCTACCTCATATTTTTCGTCAGATAATTTTCAATTCTCTAAGAGC CGTTACATGAAAACCTGTATTTTTTCAGGTGGGTATGGAAATCTTAACGCCTATACCTGGAGTT GGGAATCTTGGAGACCGT
Avr1Thrombin	TAGGTCTCCAATGTTGCCTAAAGGAGAGGAGGGTGACATTATTGGTACTTTCAATTTCTCGTCC AGCGACAGCCAACCCCTTAAAATCCACTGGGTCGATACGCCGGACTCATCTGGGAGCAATCTT GTTCCCCGTTCCGCTCACACGGAGAGTGTATGCGTTCACGCCGGGACCGCTACAGGTGCTGAT CTGCATTGGTTGAATGCGATCTGCACCGGGAAGTCTACATACACAGTGAATTGCGCCCCGGCA GGCAACAAGAATGCTGGGTCTACGCACACAGGAACATGTCCGGCAGGTCAGGACTGTTTCCAA TTAGAGCAGGTCGGAAACTTTTGGGGGGACCGTGAGCCAGATGCTACCTGTAGCCCGTCCAAT ACGGTATTTGACGCCGTAGATGACAAGGAAGCTACGCATGTAAACGGCAAAGTTGTTACACGC GCGGGGAAGCCGGGCATTGGGCGCAAGCTTATTCGTCTTAAGGCTCAGGTCTATCGTCGTGAT GGTCACTATGGTCAGACCTCGCGCATGGGATTCTTTCGTAACGGCAAAGAGGTTTACCATATC GACAACGTTGCCTCGATGGAACCCACTTGGAATTTTGACCCATCGAGTGACCAATCCTTTAGCT TCTTTTTCACACCGGGACCCAACGCTTTCCGTATTCAAGGAACGCTTAATCTGGCCTCTTGGAG ACCGT
SIX8Thrombin	TAGGTCTCCAATGACCCCGATTGACAAATCGTTAGATCAAGCGGCAACTATCGAAGAAACTGT CCACCAACCTCACTCCCATGATGAGCGTGCCTTAGTTCCACGTGGCAGCGATACGAGTGGGAT TTTGCTGGCGTGTATTACCGGCGCAGGATCTGCGTTTCAGGCGTACGCTGGATGCTACTTAACA GCTTTCCGTAATGACCCTCGCACTTTAACTTTGCGTATGGATAAAACCCGTGGAGAACGTATTT CCAATGTTCTGGTTATCTTGTCAGGGGGTGCATTGAGTCACGCCGTGGAAGAAGTAGTACAGA TTGCGCCTGGAGCGGTCCGCAATTTGGCAACATTAGGAGCTTCGACTGTCCAATTCCTTCATAA TTTTCGTTCTTGGAGACCGT
PSL1	TAGGTCTCCAATGGAAGATTGGGATCGTTGTCGTTGCATGAAATACCCCGAAACTGGGACGCC GAATGACTGTGCCACCATTAAAGCGTGCGGTTCTGGCAAGCACCGTGCCATCTCTATTTCGGA GGAGAAAGGAGACATTTGGTGCGAAAAGACTGATGTAGCAATCAGTGGACCGGAGTTCTACC GCACTTGCTATGGGTTGTTGCAGGACCCGAAGCCTAACTCGGAGGCAGATTCTTGTTGTACTC GTTGGGTAGACGGCGTTGTGGTACAATCTGACGGGTGCTTCAAATCTTGGAGACCGT

All gene sequences have been codon optimised for expression in *E. coli*. The coding sequences have been underlined.

S2 Table continued. Gene sequences used in this study

Name	Sequence
SIX13	TAGGTCTCCAATGGAACTTGAAGTTTCCGATTTAAGTGATCAACCGCCGTCAGTGGAGAATACT TACCGCGACCAAGCGTTCAATGAGGAGGAGTTGTTAAAGGTCGTGGACGAATTATCCGTGGA GCGTACCGAACACACCGAACGCGCGCTTGTGAGTGAAGCGGCAGTCCAAAAGCGCCAGGACG ACGAACATCCAAACGGTCCTTGTCCTCGCGGAGGACGTTTGTACGTGGATTCTGACGAAGACT CTTCTTGTAATGCGAAGTGGGGAACACAGACACATAACGATGTTAAGACTTTTGGCTCCACCG GATCAGTCTGTGCCGGTACTTTTCGCCGCATTACGTGTGCTTGTTGTTACACGATGCATCCTATC ACGGACAACAACGTTCCACGCATGGATGGTATTTACTGCCCTAAGTGGGAGGTGTGCAAACAA GAGCCTGAACGCTGGTCTAAGTGGGGCAATCGTGTAAGCCACACATCTTGCGTACAGGCGAA GAAGCTGACCGAAATCCTGATCGCTACTAAGAAAGTGGTGAAGGAATACTGCACCCCAAAACG TTGGTTACCATCTGCGGGAAAAGGCAAGAACGCGAAATTCCACACGTGGGCGTACAATTATTC GACGGGTCAGCTTACAACCTTGAAGTGGATGTATTTGAAGTTGGACGGTCAGTATGTCAAGAG TGCGCCTGGGATCAGCGAGTGGGGGCTTACCTATTCAGTCAATGAACATAACGCGATTGAGTT ATGTGGTTACCCAAGCGATGATATGCAACGCAATTCGATCGATGCTGAGCTTCAATGGGAGGC TACCGTACAGTCTTGGAGACCGT

All gene sequences have been codon optimised for expression in *E. coli*. The coding sequences have been underlined.

S3 Table. Primers used in this study

Name	Sequence
Avr1_Fw	TAGGTCTCCAATGCTTCCAAAGGGGGAGGAGGGTG
Avr1_Rv	ACGGTCTCCAAGAAGCTAAGTTAAGTGTACCTTGAATGCGA
Avr3_Fw	TAGGTCTCCAATGCAAGAGGCTGCGGTTCGGGA
Avr3_Rv	ACGGTCTCCAAGAGGCGTTTGGATATACCAGCCCACAC
Avr3_N Rv	ACGGTCTCCAAGATCGGGGCCCACAGACAGG
Avr3_C Fw	TAGGTCTCCAATGATTGAGGTGAAGGACAGACATGATATAG
SIX6_N Fw	TAGGTCTCCAATGGGTCCCTTAGCCCAAACAGAATCCG
SIX6_N Rv	ACGGTCTCCAAGAGGGGATACACTTAGCGAAACTCTTG
SIX6_C Fw	TAGGTCTCCAATGATTGTAGACCTGGTGGAATGGAAGAC
SIX6_C Rv	ACGGTCTCCAAGATTCCCAGCTCCACGTATAGGCGTTAAG
Avr1_N Fw	TAGGTCTCCAATGTTGCCTAAAGGAGAGGAGG
Avr1_N Rv	ACGGTCTCCAAGAATTGGACGGGCTACAGGTA
Avr1_C Fw	TAGGTCTCCAATGTCCAATACGGTATTTGACGC
Avr1_C Rv	ACGGTCTCCAAGAGGCCAGATTAAGCGTTCCT

S4 Table. Amino acid sequence inputs for AlphaFold

S4 Table. Amino acid sequence inputs for AlphaFold

Name	Sequence
SIX1$^{96-284}$	EPFGEESRNDRVTQDMLQALHDLCVERFGTGYRAVSGLCYTDRRATRKIECNKPSVRERDRSVTRA CPKGQECTTFNAYNFRNRHHQVTFPVCGPRIEVKDRHDIGIHTEWQGTWYPESPKSPGTYDYFAQ MAGTLNGYFGYDGVYSDGYKTSSHGYGHSWSCINCPRGKVTITNTYRATWAFGYTSPH
SIX2$^{98-221}$	GSCFSFPTPARGSCMIDYCWRDDNGVIYSRGITITGSNGASNPTSMRSNDPANLSLNSVFNDGYN GWFPHGHACSNSDTQIYTNHRLLQGVNGVAYVDHVRCENCNFRNVNCLSDVLKNNLIAYSNGVA SQSRCT
SIX3$^{36-163}$	LPVEDADSSVGQLQGRGNPYCVFPGRPTSSTSFTTSFSTEPLGYARMLHRDPPYERAGNSGLNHRI YERSRVGGLRTVIDVAPPDGHQAIANYEIEVRRIPVATPNAAGDCFHTARLSTGSRGPATISWDAD ASYTYYLTISED
SIX4$^{59-242}$	SAHTESVCVHAGTATGADLHWLNAICTGKSTYTVNCAPAGNKNAGSTHTGTCPAGQDCFQLEQV GNFWGDREPDATCSPSNTVFDAVDDKEATHVNGKVVTRAGKPGIGRKLIRLKAQVYRRDGHYGQ TSRMGFFRNGKEVYHIDNVASMEPTWNFDPSSDQSFSFFFTPGPNAFRIQGTLNLA
SIX5$^{18-119}$	RDHQYCACQSGSGDSIDIDATTQLQNDNSKSYLWAQTSPAYWFADRHKPGPRFAGIYLKAANGKI DGDTFYNLCINNGGADSTCFDCSKSHQVRNVIYCDAA
SIX6$^{58-225}$	DTLPVSTCPAGQKYDRSVCYKADKIRSFCVANPRSNREKITDTPCQPREICVQRNLSNGKSFAKCIPI VDLVEWKTSANGNKEGCTTTSVNPAGYHHLGTIVYDINKNPIEVDKISYFGEPGNVNEGIGGSTSYF SSDNFQFSKSRYMKTCIFSGGYGNLNAYTWSWE
SIX7$^{49-220}$	EVTFDITQNVNTFTSAASTPWTEGVGLSNIRYQWRAYYSTRQRTTFVEVRVFGTAEAQVVLLPDAP GTSRYRAIDSNVFRPNEEVTGGGLAGWGQVTTVCLQTWGRRGDITYRLRIQS
SIX8$^{50-141}$	DTSGILLACITGAGSAFQAYAGCYLTAFRNDPRTLTLRMDKTRGERISNVLVILSGGALSHAVEEVVQ IAPGAVRNLATLGASTVQFLHNFR
SIX9$^{19-114}$	QTTQVGCRALDTKNDGLLTELLLNPSARGAADPDLRYGFWDAKWRKCCNKYKECDKYYTFSYNH PYPWAYRQRRGTIRGQQFDFACVNWRTGACK
SIX10$^{15-149}$	IPDSGVSTGTKDLSKRDDAYIFDVTFRVGPAGANVAPFSGSVYVQDGLTPLVRSGSGSSISDRGYNA FRGIVYFTFTHGYNQYSASTRFGVYVDTGLIVDSNGRPIYGTAPRKACIDYSPHGPTDVCSVTITRSK
SIX11$^{19-110}$	INICCSSFAGHTCTKDQYNNHRQNVILNQIIDKDGVNCVRKGAGPGRWTRKGDWSEWYDCQQ WNGPEQHQIEVGECTLFCVTPSGILNRPCI
SIX12$^{27-127}$	SSCLSVGPKGISNQNACVCGGQCVMKDLVVARRKVCCEYTVQIQGGWPVLAQSRCVYGSTGAN GGSCSGDNVSLAWWLNYEPEVKSTDPTCIFAKPKLCHS
SIX13$^{78-293}$	QDDEHPNGPCPRGGRLYVDSDEDSSCNAKWGTQTHNDVKTFGSTGSVCAGTFRRITCACCYTM HPITDNNVPRMDGIYCPKWEVCKQEPERWSKWGNRVSHTSCVQAKKLTEILIATKKVVKEYCTPK RWLPSAGKGKNAKFHTWAYNYSTGQLTTLKWMYLKLDGQYVKSAPGISEWGLTYSVNEHNAIEL CGYPSDDMQRNSIDAELQWEATVQ
SIX14$^{18-88}$	QRILGCRMPNGSLNPSPNICNQAGGSFRSGSRGCCTRNTRDGPVVTESRFISGCNKNGGFVSSKEI LATSC

S4 Table continued. Amino acid sequence inputs for AlphaFold

Name	Sequence
PSL1$^{18-111}$	EDWDRCRCMKYPETGTPNDCATIKACGSGKHRAISISEEKGDIWCEKTDVAISGPEFYRTCYGL LQDPKPNSEADSCCTRWVDGVVVQSDGCFK
FOXG_11033$^{24-226}$	APEAAPGYTTYGDYKGAGENLPSYPSYGSYGAKPKPKPKPAPAPKKYTNYGSYNYKKYSSYGH YKREAEPEAAPEAAPEAAPEAEAAPGYTTYGDYKGAGENLPSYPSYGSYGSKPKPKPKPAPAPK KYTNYGSYNYKKYSSYGHYKREAEPEAAPEAEAEPETYSKYGSYPKKYTHYGSYNYKKYSSYGTY KRAKEFINSLF
FOXG_05755$^{56-157}$	NGVPDPGNFAASCHGLQVISDDLDLTGKPDFERCNDTAYEARQYFSGEYTTVEVRRTDYPDL GKEVQISATANYTSTNDNIVNGHLKFGDFQTKFISTPIEG
FOXG_18699$^{21-96}$	CKRTCSASNDAGTTCSYSCTQVCSSISAKQARDTFLAALQSGGNSCSAVGTSGVSCRKTAKFG SCYDHHWSCGSGC
FOXG_04863$^{106-300}$	MYDSSDDKGGLSDITRNAWSKFCNSPYGNNGGVTTRFILDGQWGAVGRLSGWSMRDALIH SMWQTADGIGKKNGYTVYNGCYGFTWQESKPGKANSACGGRSGKACPYNDDCPLAGMEC TGLKWGTWMPSIIRMNVYNRDGSLRADAYQARISSQAVGSGGCSKAQTISAYVADFIPIVGP YFATGIRINCLYQS
FOXG_04805$^{39-132}$	QNGQNGGRPVPSGECCVANTSLKQDACTASNGQAGRCVPGGNNCGGRLSCVAQANLQCD ANVIERGKDLCRAKAANGLFDGGNIIQNLSQAKVN
FOXG_02829$^{18-149}$	APSSPSDIQARSCVCKKVGDDWICTGTKCYDKVKRDLVPRQCSCHKIGDEWLCGGPKCPRDL PEENKLAKRQCSCHKVGDEWLCGGPKCPRSLPAEESGLEKRQCSCKKVAGEWICSGRKCPRD LSHLMGEE
FOXG_16600$^{17-164}$	SPISKRAVFSQTTYDDLSISGGTAGNAQQEALQKLGGLPTDLSTVEKSDLDFLNSVNQIANDAE DEAFNPAIDAASGEAADALQRGKIKNKVLKLTATILKLQAQQAQGEDVADKLAEENKKLQNNI SQDKDEAGKASTFLAFDATTS
FOXG_14684$^{38-168}$	DGTCPRPMCTTPASQGPKDPPACGDSYAACKFDQFPCDEYFSPKVTDTHHCYCILANKKAM DAYCQERGFKSGTNPWKYYYAVECHGAVSNQVCNKDCRDQGRGKGRIDKAHPNGACACDK PNPPYDTCKP
FOXGR_007323$^{18-86}$ *	SLVRRVDVNVPAMTNADGVVVPFDTAGVVQPAKKRDLEQKKRDLAQRKRHISRKRRAVSQE KQKQQQKQ
FOXGR_025639$^{18-61}$ *	APVVRGPGGRLVQEGAGCTLVQGRSVCDDGFGNTFFEDDPFSSK
FOXGR_015322$^{18-106}$ * (PSL2)	EDWDQCRCMKYPSTGTPNDCATIKACGSGKHRAISIYKNGDIWCEKTDVAINGPEFYRTCYDL LQDPKPNSEADSCCIKGDRASDGCFK
FOXGR_015533$^{19-114}$ *	QTCAIAPDPQRNADAFSATSHSGNIDIAFRDHVVFARPSAGTATGVLRLSNGDSYRKIYRIAGP NNVAQFYWLDASSQCKTNLAITQMTNAAWYKE
FOXGR_015522$^{19-79}$ * (SIX15)	TIYCRDVSPPRDTRSWCKTNTPAWQGCQRFCSEHCRSTPRDYPDGCMYHLQVGGDYDCFCK

* Effector candidates identified in the reannotation of the *Fol* genome by Sun et al. (in preparation) and not predicted in the original genome annotation by Ma et al. (2010).

Literature Cited

1. Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, et al. The Top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol. 2012;13(4):414-430.
2. Ordonez N, Seidl MF, Waalwijk C, Drenth A, Kilian A, Thomma BPHJ, et al. Worse comes to worst: bananas and Panama Disease - when plant and pathogen clones meet. PLoS Pathog. 2015;11(11):e1005197.
3. Rep M. Small proteins of plant-pathogenic fungi secreted during host colonization. FEMS Microbiol Lett. 2005;253(1):19-27.
4. Houterman PM, Speijer D, Dekker HL, de Koster CG, Cornelissen BJC, Rep M. The mixed xylem sap proteome of *Fusarium oxysporum*-infected tomato plants. Mol Plant Pathol. 2007;8(2):215-221.
5. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature. 2010;464(7287):367-373.
6. Rep M, van der Does HC, Meijer M, van Wijk R, Houterman PM, Dekker HL, et al. A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for *I-3*-mediated resistance in tomato. Mol Microbiol. 2004;53(5):1373-1383.
7. Schmidt SM, Houterman PM, Schreiver I, Ma L, Amyotte S, Chellappan B, et al. MITEs in the promoters of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. BMC Genomics. 2013;14:119.
8. Vlaardingerbroek I, Beerens B, Rose L, Fokkens L, Cornelissen BJ, Rep M. Exchange of core chromosomes and horizontal transfer of lineage-specific chromosomes in *Fusarium oxysporum*. Environ Microbiol. 2016;18(11):3702-3713.
9. Gawehns F, Houterman PM, Ait Ichou F, Michielse CB, Hijdra M, Cornelissen BJC, et al. The *Fusarium oxysporum* effector Six6 contributes to virulence and suppresses I-2-mediated cell death. Mol Plant Microbe Interact. 2014;27(4):336-348.
10. Ma L, Houterman PM, Gawehns F, Cao L, Sillo F, Richter H, et al. The *AVR2-SIX5* gene pair is required to activate I-2-mediated immunity in tomato. New Phytol. 2015;208(2):507-518.
11. van der Does HC, Lievens B, Claes L, Houterman PM, Cornelissen BJC, Rep M. The presence of a virulence locus discriminates *Fusarium oxysporum* isolates causing tomato wilt from other isolates. Environ Microbiol. 2008;10(6):1475-1485.
12. Gawehns F, Ma L, Bruning O, Houterman PM, Boeren S, Cornelissen BJC, et al. The effector repertoire of *Fusarium oxysporum* determines the tomato xylem proteome composition following infection. Front Plant Sci. 2015;6:967.
13. Li E, Wang G, Xiao J, Ling J, Yang Y, Xie B. A *SIX1* homolog in *Fusarium oxysporum* f. sp. *conglutinans* is required for full virulence on cabbage. PLoS One. 2016;11(3):e0152273.
14. Thatcher LF, Gardiner DM, Kazan K, Manners JM. A highly conserved effector in *Fusarium oxysporum* is required for full virulence on Arabidopsis. Mol Plant Microbe Interact. 2012;25(2):180-190.
15. An B, Hou X, Guo Y, Zhao S, Luo H, He C, et al. The effector SIX8 is required for virulence of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 to Cavendish banana. Fungal Biol. 2019;123(5):423-430.

16. Widinugraheni S, Nino-Sanchez J, van der Does HC, van Dam P, Garcia-Bastidas FA, Subandiyah S, et al. A SIX1 homolog in *Fusarium oxysporum* f. sp. *cubense* tropical race 4 contributes to virulence towards Cavendish banana. PLoS One. 2018;13(10):e0205896.
17. Cao L, Blekemolen MC, Tintor N, Cornelissen BJC, Takken FLW. The Fusarium oxysporum Avr2-Six5 Effector Pair Alters Plasmodesmatal Exclusion Selectivity to Facilitate Cell-to-Cell Movement of Avr2. Mol Plant. 2018;11(5):691-705.
18. Ayukawa Y, Asai S, Gan P, Tsushima A, Ichihashi Y, Shibata A, et al. A pair of effectors encoded on a conditionally dispensable chromosome of *Fusarium oxysporum* suppress host-specific immunity. Commun Biol. 2021;4(1):707.
19. Gonzalez-Cendales Y, Catanzariti AM, Baker B, McGrath DJ, Jones DA. Identification of *I-7* expands the repertoire of genes for resistance to *Fusarium* wilt in tomato to three resistance gene classes. Mol Plant Pathol. 2016;17(3):448-463.
20. Catanzariti AM, Do HTT, Bru P, de Sain M, Thatcher LF, Rep M, et al. The tomato *I* gene for *Fusarium* wilt resistance encodes an atypical leucine-rich repeat receptor-like protein whose function is nevertheless dependent on *SOBIR1* and *SERK3/BAK1*. Plant J. 2017;89(6):1195-1209.
21. Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, et al. Dissection of the *Fusarium I-2* gene cluster in tomato reveals six homologs and one active gene copy. Plant Cell. 1998;10(6):1055-1068.
22. Catanzariti AM, Lim GTT, Jones DA. The tomato *I-3* gene: a novel gene for resistance to *Fusarium* wilt disease. New Phytol. 2015;207(1):106-118.
23. Houterman PM, Cornelissen BJC, Rep M. Suppression of plant resistance gene-based immunity by a fungal effector. PLoS Pathog. 2008;4(5):e1000061.
24. Houterman PM, Ma L, van Ooijen G, de Vroomen MJ, Cornelissen BJC, Takken FLW, et al. The effector protein Avr2 of the xylem-colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. Plant J. 2009;58(6):970-978.
25. Di X, Cao L, Hughes RK, Tintor N, Banfield MJ, Takken FLW. Structure-function analysis of the *Fusarium oxysporum* Avr2 effector allows uncoupling of its immune-suppressing activity from recognition. New Phytol. 2017;216(3):897-914.
26. Sarma GN, Manning VA, Ciuffetti LM, Karplus PA. Structure of Ptr ToxA: an RGD-containing host-selective toxin from *Pyrenophora tritici-repentis*. Plant Cell. 2005;17(11):3190-3202.
27. Wang CI, Guncar G, Forwood JK, Teh T, Catanzariti AM, Lawrence GJ, et al. Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. Plant Cell. 2007;19(9):2898-2912.
28. de Guillen K, Ortiz-Vallejo D, Gracy J, Fournier E, Kroj T, Padilla A. Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. PLoS Pathog. 2015;11(10):e1005228.
29. Spanu PD. Cereal immunity against powdery mildews targets RNase-Like Proteins associated with Haustoria (RALPH) effectors evolved from a common ancestral gene. New Phytol. 2017;213(3):969-971.
30. Lazar N, Mesarich CH, Petit-Houdenot Y, Talbi N, de la Sierra-Gallay IL, Zélie E, et al. A new family of structurally conserved fungal effectors displays epistatic interactions with plant resistance proteins. bioRxiv. 2021:2020.2012.2017.423041.
31. Outram MA, Sung YC, Yu D, Dagvadorj B, Rima SA, Jones DA, et al. The crystal structure of SnTox3 from the necrotrophic fungus *Parastagonospora nodorum* reveals a

unique effector fold and provides insight into Snn3 recognition and pro-domain protease processing of fungal effectors. New Phytol. 2021;231(6):2282-2296.
32. Outram MA, Solomon PS, Williams SJ. Pro-domain processing of fungal effector proteins from plant pathogens. PLoS Pathog. 2021;17(10):e1010000.
33. Yu D, Outram MA, Creen E, Smith A, Sung YC, Darma R, et al. Optimised production of disulfide-bonded fungal effectors in E. coli using CyDisCo and FunCyDisCo co-expression approaches. Mol Plant Microbe Interact. 2021.
34. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010;38(Web Server issue):W545-549.
35. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583-589.
36. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr. 2004;60(Pt 12 Pt 1):2256-2268.
37. Krissinel E. Enhanced fold recognition using efficient short fragment clustering. J Mol Biochem. 2012;1(2):76-85.
38. Iwaoka R, Nagata T, Tsuda K, Imai T, Okano H, Kobayashi N, et al. Structural insight into the recognition of r(UAG) by Musashi-1 RBD2, and construction of a model of Musashi-1 RBD1-2 bound to the minimum target RNA. Molecules. 2017;22(7):1207.
39. Ledwidge R, Hong B, Dotsch V, Miller SM. NmerA of Tn501 mercuric ion reductase: structural modulation of the pKa values of the metal binding cysteine thiols. Biochemistry. 2010;49(41):8988-8998.
40. Bentham AR, Petit-Houdenot Y, Win J, Chuma I, Terauchi R, Banfield MJ, et al. A single amino acid polymorphism in a conserved effector of the multihost blast fungus pathogen expands host-target binding spectrum. bioRxiv. 2021:2021.2003.2015.435478.
41. Maqbool A, Saitoh H, Franceschetti M, Stevenson CEM, Uemura A, Kanzaki H, et al. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. eLife. 2015;4:e08709.
42. Li J, Fokkens L, Conneely LJ, Rep M. Partial pathogenicity chromosomes in *Fusarium oxysporum* are sufficient to cause disease and can be horizontally transferred. Environ Microbiol. 2020;22(12):4985-5004.
43. Simbaqueba J, Catanzariti AM, Gonzalez C, Jones DA. Evidence for horizontal gene transfer and separation of effector recognition from effector function revealed by analysis of effector genes shared between cape gooseberry- and tomato-infecting *formae speciales* of *Fusarium oxysporum*. Mol Plant Pathol. 2018;19(10):2302-2318.
44. Stergiopoulos I, de Wit PJGM. Fungal effector proteins. Annu Rev Phytopathol. 2009;47:233-263.
45. Pennington HG, Jones R, Kwon S, Bonciani G, Thieron H, Chandler T, et al. The fungal ribonuclease-like effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. PLoS Pathog. 2019;15(3):e1007620.
46. Pedersen C, Ver Loren van Themaat E, McGuffin LJ, Abbott JC, Burgis TA, Barton G, et al. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics. 2012;13:694.
47. Blondeau K, Blaise F, Graille M, Kale SD, Linglin J, Ollivier B, et al. Crystal structure of the effector AvrLm4-7 of *Leptosphaeria maculans* reveals insights into its translocation into plant cells and recognition by resistance proteins. Plant J. 2015;83(4):610-624.

48. Petit-Houdenot Y, Langner T, Harant A, Win J, Kamoun S. A clone resource of *Magnaporthe oryzae* effectors that share sequence and structural similarities across host-specific lineages. Mol Plant Microbe Interact. 2020;33(8):1032-1035.
49. Bauer S, Yu D, Lawson AW, Saur IML, Frantzeskakis L, Kracher B, et al. The leucine-rich repeats in allelic barley MLA immune receptors define specificity towards sequence-unrelated powdery mildew avirulence effectors with a predicted common RNase-like fold. PLoS Pathog. 2021;17(2):e1009223.
50. Seong K, Krasileva K. Computational structural genomics unravels common folds and novel families in the secretome of fungal phytopathogen *Magnaporthe oryzae*. Mol Plant Microbe Interact. 2021.
51. Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Zidek A, et al. Highly accurate protein structure prediction for the human proteome. Nature. 2021;596(7873):590-596.
52. Flower TG, Hurley JH. Crystallographic molecular replacement using an in silico-generated search model of SARS-CoV-2 ORF8. Protein Sci. 2021;30(4):728-734.
53. Park SY, Jeong MS, Park SA, Ha SC, Na BK, Jang SB. Structural basis of the cystein protease inhibitor *Clonorchis sinensis* Stefin-1. Biochem Biophys Res Commun. 2018;498(1):9-17.
54. Renko M, Taler-Vercic A, Mihelic M, Zerovnik E, Turk D. Partial rotational lattice order-disorder in stefin B crystals. Acta Crystallogr D Biol Crystallogr. 2014;70(Pt 4):1015-1025.
55. Dietrich JD, Longenecker KL, Wilson NS, Goess C, Panchal SC, Swann SL, et al. Development of orally efficacious allosteric inhibitors of TNFalpha via fragment-based drug design. J Med Chem. 2021;64(1):417-429.
56. Sulak O, Cioci G, Delia M, Lahmann M, Varrot A, Imberty A, et al. A TNF-like trimeric lectin domain from *Burkholderia cenocepacia* with specificity for fucosylated human histo-blood group antigens. Structure. 2010;18(1):59-72.
57. Khoshnevis S, Neumann P, Ficner R. Crystal structure of the RNA recognition motif of yeast translation initiation factor eIF3b reveals differences to human eIF3b. PLoS One. 2010;5(9):e12784.
58. Bull PC, Cox DW. Wilson disease and Menkes disease: new handles on heavy-metal transport. Trends Genet. 1994;10(7):246-252.
59. Bleackley MR, Vasa S, Harvey PJ, Shafee TMA, Kerenga BK, Soares da Costa TP, et al. Histidine-rich defensins from the *Solanaceae* and *Brasicaceae* are antifungal and metal binding proteins. J Fungi 2020;6(3):145.
60. Lay FT, Ryan GF, Caria S, Phan TK, Veneer PK, White JA, et al. Structural and functional characterization of the membrane-permeabilizing activity of *Nicotiana occidentalis* defensin NoD173 and protein engineering to enhance oncolysis. FASEB J. 2019;33(5):6470-6482.
61. Prochnicka-Chalufour A, Corzo G, Satake H, Martin-Eauclaire MF, Murgia AR, Prestipino G, et al. Solution structure of discrepin, a new K+-channel blocking peptide from the alpha-KTx15 subfamily. Biochemistry. 2006;45(6):1795-1804.
62. Korolkova YV, Bocharov EV, Angelo K, Maslennikov IV, Grinenko OV, Lipkin AV, et al. New binding site on common molecular scaffold provides HERG channel specificity of scorpion toxin BeKm-1. J Biol Chem. 2002;277(45):43104-43109.
63. Van Duyne GD, Ghosh G, Maas WK, Sigler PB. Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. J Mol Biol. 1996;256(2):377-391.

64. Cherney LT, Cherney MM, Garen CR, Lu GJ, James MN. Structure of the C-terminal domain of the arginine repressor protein from *Mycobacterium tuberculosis*. Acta Crystallogr D Biol Crystallogr. 2008;64(Pt 9):950-956.
65. Ma S, Lapin D, Liu L, Sun Y, Song W, Zhang X, et al. Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. Science. 2020;370(6521):eabe3069.
66. Martin R, Qi T, Zhang H, Liu F, King M, Toth C, et al. Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. Science. 2020;370(6521):eabd9993.
67. De la Concepcion JC, Franceschetti M, MacLean D, Terauchi R, Kamoun S, Banfield MJ. Protein engineering expands the effector recognition profile of a rice NLR immune receptor. eLife. 2019;8:e47713.
68. Cesari S, Xi Y, Declerck N, Chalvon V, Mammri L, Pugnière M, et al. Design of a new effector recognition specificity in a plant NLR immune receptor by molecular engineering of its integrated decoy domain. bioRxiv. 2021:2021.2004.2024.441256.
69. Ortiz D, de Guillen K, Cesari S, Chalvon V, Gracy J, Padilla A, et al. Recognition of the *Magnaporthe oryzae* effector AVR-Pia by the decoy domain of the rice NLR immune receptor RGA5. Plant Cell. 2017;29(1):156-168.
70. Cesari S, Thilliez G, Ribot C, Chalvon V, Michel C, Jauneau A, et al. The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell. 2013;25(4):1463-1481.
71. Guo L, Cesari S, de Guillen K, Chalvon V, Mammri L, Ma M, et al. Specific recognition of two MAX effectors by integrated HMA domains in plant immune receptors involves distinct binding surfaces. Proc Natl Acad Sci U S A. 2018;115(45):11637-11642.
72. Kanzaki H, Yoshida K, Saitoh H, Fujisaki K, Hirabuchi A, Alaux L, et al. Arms race co-evolution of *Magnaporthe oryzae* AVR-Pik and rice *Pik* genes driven by their physical interactions. Plant J. 2012;72(6):894-907.
73. Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, et al. A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. Plant J. 2011;66(3):467-479.
74. Varden FA, Saitoh H, Yoshino K, Franceschetti M, Kamoun S, Terauchi R, et al. Cross-reactivity of a rice NLR immune receptor to distinct effectors from the rice blast pathogen *Magnaporthe oryzae* provides partial disease resistance. J Biol Chem. 2019;294(35):13006-13016.
75. Plissonneau C, Daverdin G, Ollivier B, Blaise F, Degrave A, Fudal I, et al. A game of hide and seek between avirulence genes *AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans*. New Phytol. 2016;209(4):1613-1624.
76. Ghanbarnia K, Ma L, Larkan NJ, Haddadi P, Fernando WGD, Borhan MH. *Leptosphaeria maculans* AvrLm9: a new player in the game of hide and seek with AvrLm4-7. Mol Plant Pathol. 2018;19(7):1754-1764.
77. Larkan NJ, Ma L, Haddadi P, Buchwaldt M, Parkin IAP, Djavaheri M, et al. The *Brassica napus* wall-associated kinase-like (WAKL) gene *Rlm9* provides race-specific blackleg resistance. Plant J. 2020;104(4):892-900.
78. Bentham AR, Youles M, Mendel MN, Varden FA, De la Concepcion JCB, Mark J. pOPIN-GG: A resource for modular assembly in protein expression vectors. bioRxiv. 2021:2021.2008.2010.455798.

79. Iverson SV, Haddock TL, Beal J, Densmore DM. CIDAR MoClo: improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. ACS Synth Biol. 2016;5(1):99-103.
80. Cowieson NP, Aragao D, Clift M, Ericsson DJ, Gee C, Harrop SJ, et al. MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. J Synchrotron Radiat. 2015;22(1):187-190.
81. Kabsch W. *XDS*. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-132.
82. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013;69(Pt 7):1204-1214.
83. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the *CCP4* suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):235-242.
84. Skubak P, Pannu NS. Automatic protein structure solution from weak X-ray data. Nat Commun. 2013;4:2777.
85. Skubak P, Arac D, Bowler MW, Correia AR, Hoelz A, Larsen S, et al. A new MR-SAD algorithm for the automatic building of protein models from low-resolution X-ray data and a poor starting model. IUCrJ. 2018;5(Pt 2):166-171.
86. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with *phenix.refine*. Acta Crystallogr D Biol Crystallogr. 2012;68(Pt 4):352-367.
87. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *COOT*. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.
88. Aragao D, Aishima J, Cherukuvada H, Clarken R, Clift M, Cowieson NP, et al. MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. J Synchrotron Radiat. 2018;25(Pt 3):885-891.
89. Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, et al. Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard. Acta Crystallogr D Biol Crystallogr. 2008;64(Pt 1):61-69.
90. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37(4):420-423.