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[1,2,4], Lisong Ma[1,4], 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]The Australian Nuclear Science and Technology Organisation, Australian Synchrotron, Clayton, Victoria 3168, Australia

[4]State Key Laboratory of North China Crop Improvement and Regulation, College of Horticulture, Hebei Agricultural University, Baoding 071001, China

[§]Corresponding author: Simon Williams; simon.williams@anu.edu.au, +61 2 6125 7862
Megan Outram: megan.outram@csiro.au

* Current address: Black Mountain Science and Innovation Park, CSIRO Agriculture and Food, Canberra, Australia

Keywords: Fungal effectors, FOLD, *Fusarium oxysporum*, Secreted in xylem (SIX) effectors, Structural biology

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.

Daniel S. Yu	0000-0003-0454-7989
Megan A Outram	0000-0003-4510-3575
Ashley Smith	
Carl L. McCombe	0000-0001-9347-8879
Pravin B. Khambalkar	0000-0003-3641-6553
Sharmin A. Rima	0000-0003-2808-8369
Xizhe Sun	
Lisong Ma	0000-0001-9288-869X
Daniel J. Ericsson	0000-0001-5101-9244
David A. Jones	0000-0001-8809-5822
Simon J. Williams	0000-0003-4781-6261

Abstract

Plant pathogens secrete proteins, known as effectors, that function in the apoplast or inside plant cells to promote virulence. Effector detection by cell-surface or cytosolic receptors results in the activation of defence pathways and plant immunity. Despite their importance, our general understanding of fungal effector function and detection by immunity receptors remains poor. One complication often associated with effectors is their high sequence diversity and lack of identifiable sequence motifs precluding prediction of structure or function. In recent years, several studies have demonstrated that fungal effectors can be grouped into structural classes, despite significant sequence variation and existence across taxonomic groups. 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*** **d**ual-domain (FOLD) effector class, with members containing two distinct domains. We predicted the full SIX effector repertoire of *Fol* using AlphaFold2, and show that SIX6 and SIX13 are also FOLD effectors, which we validated experimentally for SIX6. Based on structural prediction and comparison, we show that FOLD effectors are present within three divisions of fungi, and are expanded in pathogens and symbionts. Further structural comparisons within the *Fol* effectors demonstrated that *Fol* secretes a limited number of structurally related effectors during infection and colonisation of tomato. This analysis also revealed a structural relationship between transcriptionally co-regulated effector pairs. Collectively, these observations have broad implications for our understanding of effector function, pathogen virulence and the engineering of plant immunity receptors.

Author Summary

Fusarium oxysporum is a soil-borne fungal pathogen responsible for destructive vascular wilt diseases in plants. The wide host range of *F. oxysporum* and ability to lay dormant within the soil for many years makes it one of the most destructive fungal pathogens worldwide. During infection, *F. oxysporum* secretes multiple effector proteins to promote virulence and aid in colonisation. In recent years, significant progress has been made in our capacity to identify effectors within fungal genomes. Despite this progress, our structural and mechanistic understanding of how effectors promote virulence remains relatively poor. Here, we combine experimental and computational approaches to define and model the structural repertoire of effector proteins secreted by *F. oxysporum* f. sp. *lycopersici* (*Fol*), the causative agent of vascular wilt disease in tomato. Our analysis identifies a new structural class of fungal effectors we define as FOLD (***Fol*** **d**ual-domain) effectors and show, via structural comparisons, that FOLD proteins are conserved across fungi and expanded in both pathogens and symbionts. We subsequently show that sequence unrelated *Fol* effectors can be grouped into at least 5 structural classes. Collectively, these results show that *F. oxysporum,* and most likely fungal pathogens in general, secrete effectors during plant infection with a smaller range of structural diversity than predicted based on sequence studies alone. This study represents an important advance in our understanding of plant-fungus interactions and will assist the development of novel control and engineering strategies to combat fungal pathogens.

Introduction

Fusarium oxysporum is a soil-borne fungal pathogen responsible for destructive vascular wilt diseases in a wide range of plants. It ranks within the top ten important fungal pathogens in terms of scientific and economic importance [1]. Vascular wilting caused by *F. oxysporum* contributes to significant losses in crop production worldwide. Of particular concern is the tropical race 4 variant of the banana pathogen, *F. oxysporum* f. sp. *cubense* (*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 has the potential to eradicate Cavendish as a commercial cultivar [2].

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 identity 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 also a 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 possess resistance genes that encode immune receptors capable of recognising specific effectors 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 shares structural homology to ToxA from *Pyrenophora tritici-repentis* and AvrL567 from *Melampsora lini* [26, 27], despite a lack of sequence identity. The observation of structural classes for effectors without identifiable domains or homologies to proteins of known function has been demonstrated experimentally 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**NA**s**e-**L**ike **P**roteins associated with **H**austoria) [29], LARS (***L****eptosphaeria* **A**vi**r**ulence-**S**uppressing) [30, 31] and ToxA-like families [25-27].

Here, we present the structures of Avr1, Avr3, SIX6 and SIX8, determined using x-ray crystallography. We identified a new structural family of fungal effectors we term the FOLD (***Fol*** **d**ual-domain) effectors, and show using structural comparisons against the AlphaFold structural database that FOLD effectors are widely distributed in phytopathogenic fungi as well as symbionts. Combining experimental and computational approaches, we present the structural repertoire of sequence unrelated effectors utilised by *Fol* during infection of tomato and demonstrate that many of these effectors fall within a limited number of structural families.

Results

The structures of Avr1 and Avr3 adopt a similar fold

Avr1 and Avr3 are cysteine-rich effectors that belong to the K2PP (Kex2-processed pro-domain) effector class [32, 33]. To help understand their function, and recognition by I and I-3, we sought to solve their structures using x-ray crystallography. Using an optimised protein production strategy [34], we produced Avr1 ($Avr1^{18-242}$) and Avr3 ($Avr3^{22-284}$) as disulfide-bonded proteins in *E. coli* for crystallisation studies (S1A and S1B Fig). Crystals were obtained for $Avr3^{22-284}$ (S1B Fig), however, $Avr1^{18-242}$ failed to crystallise. Previously, we demonstrated that pro-domain removal from the K2PP effector SnTox3 was required to obtain protein crystals [32] 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 at additional sites [32]. To address this, Avr1 was engineered with an internal thrombin cleavage site replacing the Kex2 site to produce a single $Avr1^{59-242}$ product after thrombin cleavage. This protein was subsequently used for crystallisation studies resulting in rectangular plate-like crystals (S1A Fig).

The crystal structures of Avr1 and Avr3 were solved using a bromide-ion-based single-wavelength anomalous diffraction (SAD) approach (S1 Table), and subsequently were refined using a native dataset to a resolution of 1.65 Å and 1.68 Å, respectively (Fig 1A and 1B). Despite sharing low amino-acid sequence identity (19.5%), Avr1 and Avr3 adopt a structurally similar dual-domain protein fold. Interpretable, continuous electron density was observed from residue 96 in Avr3 and 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 (residues 23-45) of Avr1 in the electron density, despite thrombin cleavage of the pro-domain prior to crystallisation. This indicates that an association between respective Avr and pro-domain was maintained post cleavage *in vitro* (S2B Fig). The importance of this association, if any, remains unclear, but for simplicity, 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 topology, consisting of an N-terminal α-helix followed by five β-strands (Fig 1A and 1B). The N-domains have a root-mean-square deviation (RMSD) of 2.1 Å when superimposed using the DALI server [35], 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 with an RMSD of 2.8 Å, and consist of a β-sandwich domain architecture, involving seven or eight β-strands, respectively (Fig 1D). The C-domain of Avr3 contains a single disulfide bond within strand β10 that is not present in Avr1. 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* dual**-domain) effectors.

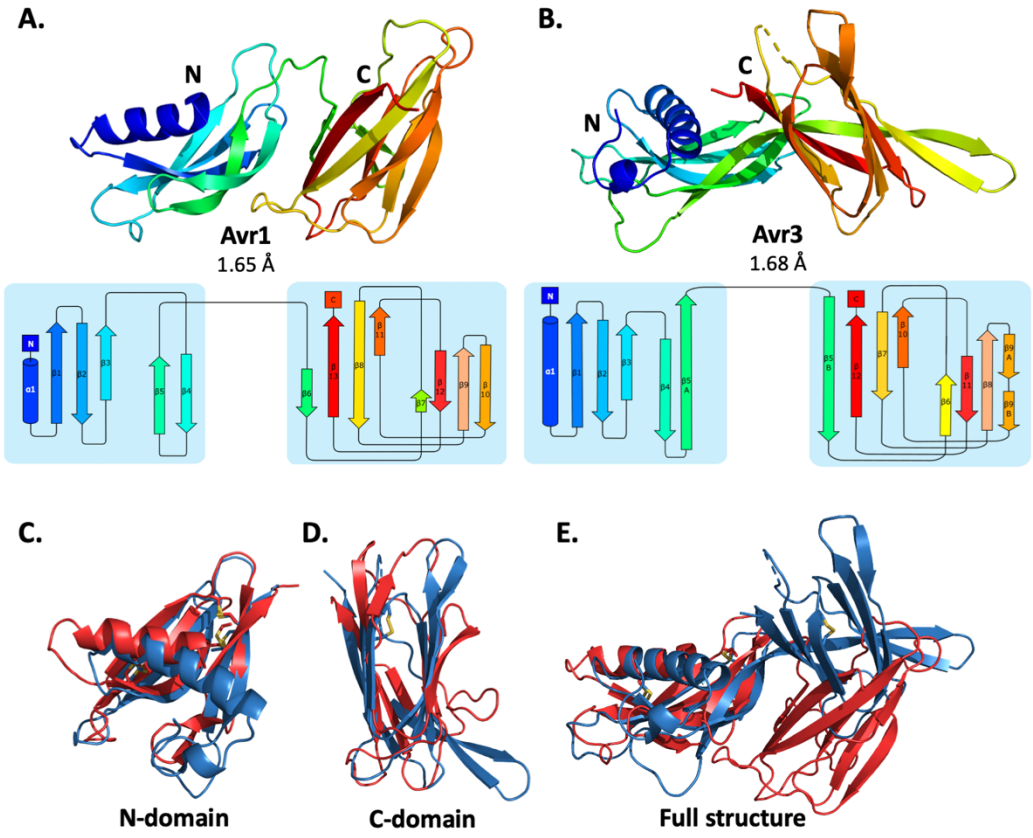


Fig 1. Crystal structures of Avr1 and Avr3 from *Fol* adopt a similar structural fold that is unique among fungal effectors. Ribbon diagrams of Avr1 and Avr3 coloured from N- (blue) to C-terminus (red) in the top panel showing the dual-domain structural fold, and bottom panels showing secondary structure topology map of Avr1 **(A)** and Avr3 **(B)**, respectively. 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. 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 of the DALI server [35].

SIX6 and SIX13 belong to the FOLD effector family

We were interested to determine if other SIX effectors belonged to the FOLD effector family. One conserved sequence feature observed 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 like Avr1 and Avr3 (Fig 2A), suggesting they may be FOLD effectors. With the recent advances in *ab initio* structural prediction by Google DeepMind's AlphaFold2 [36] we predicted the structures of the SIX effectors to determine if, as suggested by our sequence analysis, other SIX effectors are FOLD effector family members.

As an initial step we benchmarked AlphaFold2 predicted models of Avr1 and Avr3 (downstream of the Kex2 cleavage site (Avr1$^{59-242}$ and Avr3$^{96-284}$) against our experimentally determined structures (S3 Fig). The AlphaFold2 model of Avr1 returned a relatively poor average per-residue confidence score (pLDDT =55%) with an RMSD of 6.9 Å between the model and structure, however, the dual domain architecture was correctly predicted with a Z-score of 11.3 identified using a Dali pair-wise structural comparison (S3A Fig and S3E). The AlphaFold2 model of Avr3 returned a high pLDDT score (92%) and superimposed well to the solved structure (S3B Fig), despite a slight skew between the orientation of the individual domains (RMSD = 3.7 Å overall; 1.1 Å for the N-domain; 0.8 Å for the C-domain). This demonstrated that accurate FOLD effector prediction was possible using AlphaFold2.

We subsequently generated SIX6 and SIX13 models, downstream of the predicted Kex2 cleavage site (SIX6$^{58-225}$, SIX13$^{78-293}$), using AlphaFold2 and obtained high average confidence scored models supporting their inclusion in the FOLD family (S4 Fig). To validate this experimentally, we produced SIX6 and SIX13 as described for Avr1/Avr3 and obtained crystals for both proteins (S1 Fig). While the SIX13 crystals diffracted poorly, the SIX6 crystals diffracted x-rays to ~1.9 Å and we solved the structure of SIX6 using the AlphaFold2 generated model as a template for molecular replacement (Fig 2B, S1 Table).

The SIX6 structure confirms its inclusion as a member of the FOLD family. Despite lacking an N-terminal helix, the N-domain contains five β-strands held together by three disulfide bonds with an arrangement identical to Avr1 and Avr3. The C-domain is an eight stranded β-sandwich that is stabilised by a single disulfide bond (unique to SIX6) connecting the β7 and β12 strands. Like Avr1, we identified regions of the pro-domain within the SIX6 structure (residues 29-46 were observed in the electron density), despite cleavage of the pro-domain prior to crystallisation (S2C Fig). In the case of SIX6, two molecules were observed in the asymmetric unit (S2D Fig), but only part of one pro-domain was supported by electron

density. For subsequent structural analysis, we used Chain A of SIX6 and excluded the structured regions of the pro-domain (Fig 2B).

FOLD effectors are distributed across multiple fungal genera

Despite structural similarities, the FOLD effectors are divergent in their amino acid sequences, sharing 15.5 – 22.5% sequence identities between all members (Fig 2A). Homologues of FOLD effectors are dispersed across multiple *formae speciales* of *F. oxysporum* (S5A Fig) [7, 9, 37-40]. Previous structural-based searches performed on effector candidates from *Venturia inaequalis* using Avr1 and Avr3 as templates (which we provided to the authors) found three candidates predicted to be FOLD effectors [41].

To explore this further, we utilised our experimentally determined structures (Avr1, Avr3 and SIX6) to search for other fungal FOLD effectors within the AlphaFold2 protein structure database [42] (https://alphafold.ebi.ac.uk/) using the Foldseek webserver [43]. This analysis identified 124 putative FOLD protein family members across three Divisions of Fungi (Ascomycota, Basidiomycota, and Glomeromycota) (Fig. 2C). Over half of these were found in Ascomycetes (73), with expanded families in species of *Colletotrichum, Diversisora,* and *Rhizophagus* (Fig 2C, S2 Table), as well as many *formae speciales* of *Fusarium oxysporum* and other *Fusarium* species (S2 Table). Expanded families of FOLD proteins were observed in the genus of *Glomeromycota* that form arbuscular mycorrhiza in plant roots, while two putative FOLD effectors were also predicted in the ectomycorrhizal fungus *Piloderma olivaceum* (basidiomycete), which forms mutualistic associations with conifer and hardwood species [44]. Structural superposition of members from the three Divisions confirms the structural similarities between the N and C domains and highlights that the major differences identified are the orientation of the domains relative to each other (Fig. 2D), consistent with our experimental data for Avr1, Avr3 and SIX6.

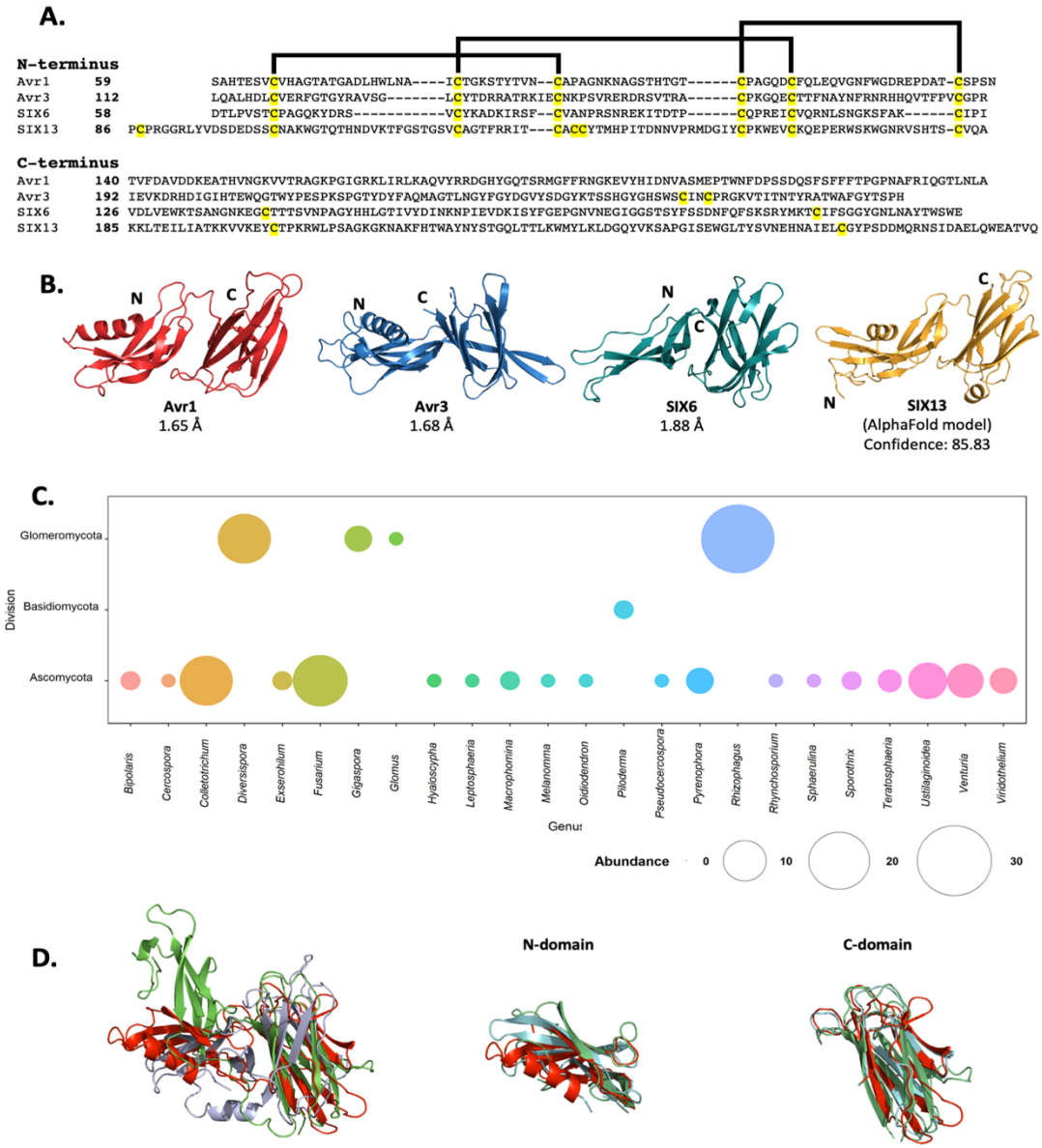


Fig 2. FOLD effector family is distributed within *Fusarium oxysporum* and other fungi. **(A)** Amino acid sequence alignment of Avr1, Avr3, SIX6 and SIX13 show a common cysteine spacing at the N-terminus. The alignment is split into the N-terminus (N-domain; top panel) and C-terminus (C-domain; bottom 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 crystal structure and **(C)** SIX13 model predicted by AlphaFold2 showing the dual-domain structural fold, transitioning from blue (N-terminus) to red (C-terminus). **(C)** Structure-guided search for putative FOLD effectors across fungi using Foldseek webserver. Size of circles represent abundance with genus. **(D)** Superposition (structural alignment) of representative putative FOLD effectors from

the divisions Glomeromycota and Basidiomycota with Avr1 in ribbon representation. Putative FOLD protein from *Rhizophagus clarus* (UniProt: A0A2Z6QDJ0) in light blue, and *Piloderma croceum* (UniProt: A0A0C3C2B2) in green. FOLD structural alignment (right), N-domain only (middle), C-domain only (right).

Distinct structural families exist among the other SIX effectors

With the successful utilisation of AlphaFold2 as a model for molecular replacement (SIX6 structure), and structural similarity searches for FOLD effectors, we decided to perform structural comparisons with the remaining SIX effectors. AlphaFold2 modelling of the effectors was conducted on sequences with the signal peptide and putative pro-domain (if present) (S6 Fig) removed. The models and experimentally determined SIX effector structures (Avr1, Avr2, Avr3 and SIX6) were compared using the DALI server [35] and a Z-score with a cutoff of >2 was used to indicate structure similarity.

The observed structural similarity between the FOLD effectors was high, with scores above 8 for all comparisons (Fig 3A). Avr2, a member of the ToxA-like effector family, exhibited structural similarity with the SIX7$^{49-220}$ and SIX8$^{50-141}$ models (Z-scores > 5) (Fig 3A). Analysis of the models and topology show that SIX7 and SIX8 both consist of a β-sandwich fold, strongly indicating their inclusion of within the ToxA-like structural family (Fig 3C, S7 Fig).

Beyond these described structural families, the Z-scores indicated that two additional, but not yet characterised, structural families exist within the SIX effectors. Here, we define these are structural family 3 and 4, consisting of SIX9$^{19-114}$ and SIX11$^{19-110}$, and SIX5$^{18-119}$ and SIX14$^{18-88}$, respectively (Fig 3D, E). The structures of SIX9 and SIX11 both consist of five β-strands and either two or three α-helices (Fig 3D, S8 Fig), despite sharing only 14% sequence identity. To further our understanding of the putative function of this family we did a structural search against the protein databank (PDB) and found that both structures share structural similarity to various RNA binding proteins, such as the RNA recognition motif (RRM) fold of the Musashi-1 RNA-binding domain (PDB code: 5X3Z) [45].

SIX5 and SIX14 also share limited sequence identity (23%) but the structural predictions show a similar secondary-structure topology consisting of two α-helices and four to six β-strands (Fig 3E, S8 Fig). We compared the models of SIX5 and SIX14 against the PDB using DALI and identified structural similarity toward the *Ustilago maydis* and *Zymoseptoria tritici* KP6 effector (PDB codes: 4GVB and 6QPK) [46], suggesting SIX5 and SIX14 belong to the KP6-like structural family (S7 Fig). Collectively, this analysis demonstrates that 11 of the 14 SIX effectors, group into 4 different structural families.

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

The SIX effectors are only a subset of effectors utilised by *Fol* during infection of tomato. Recently, the *Fol* genome was re-sequenced [47] and reannotated in combination with RNAseq

data from *Fol*-infected tomato plants [48]. A total of 26 genes encoding novel effector candidates were identified that were consistently upregulated during *Fol* infection [48], which were not previously predicted or predicted incorrectly in the original genome annotation [5]. Of these, 14 genes encoded proteins with no recognised domains or motifs based on their amino acid sequences. We assessed whether these 14 effector candidates could be grouped into the four structural families of SIX effectors we identified, by generating structural models using AlphaFold2 (S3 Table, S6 Fig) and structurally aligning them using DALI against SIX effector representatives from each family (Fig 3B). We found the predicted structure of FOXGR_015533 adopts a nine β-stranded sandwich and is likely a member of the ToxA-like class (Fig 3C). PSL1 [18] and FOXGR_015322, here designated PSL2, are sequence related effectors (~85% sequence identity) and show a conserved structure consisting of two α-helices and four or five β-strands (Fig 3E). Both have Z-scores of >2 against Family 4 and are likely members of this family.

Based on this analysis we also suggest an additional structural family. FOXG_18699 and FOXGR_015522 are structurally related (Z-score of 2.2) with a sequence identity of ~29%. While FOXGR_015522 does share some resemblance to Family 4, based on manual alignment (Fig 3F) and domain topology analysis (S8 Fig) these effectors appear to belong to an independent structural family, designated Family 5. Collectively, these data demonstrate that *Fol* utilises multiple structurally related, sequence diverse, effectors during infection of tomato.

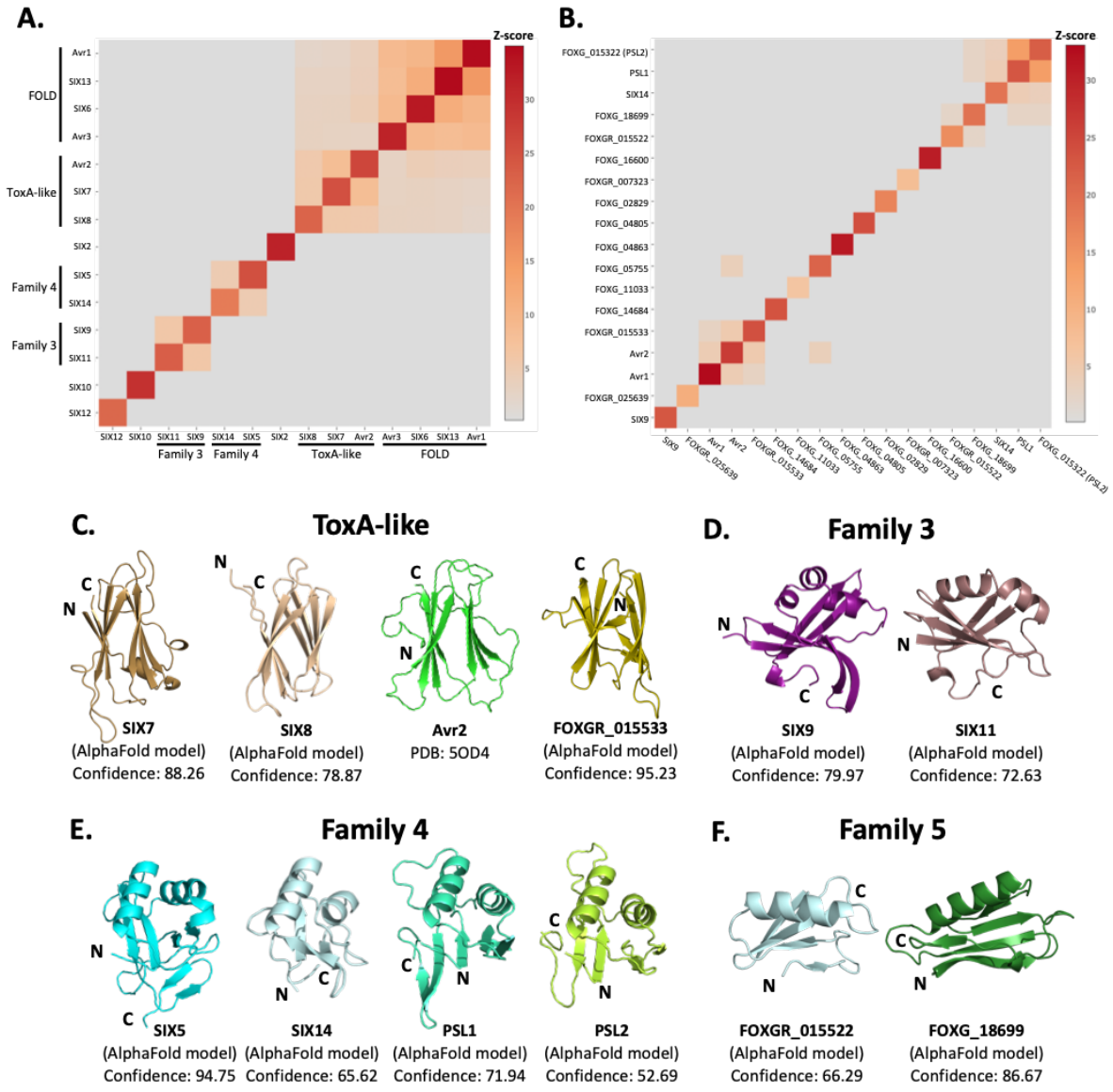


Fig 3. Identification of new putative structural families within the SIX effectors. Heat maps showing the structural similarity of Structures and AlphaFold2 models of the **(A)** SIX effectors and **(B)** effector candidates from *Fol* in a structural pairwise alignment. Structural similarity was measured with Z-scores. A cutoff Z-score of 2 was applied for defining structural families. Z-score scale is shown in a grey to red spectrum. **(C)** Cartoon representation of the ToxA-like effectors from *Fol*. AlphaFold2 models of SIX7, SIX8 and FOXGR_015533 effector candidate are 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. Cartoon representations of **(D)** Family 3, **(E)** Family 4 and **(F)** Family 5 consisting of members that are predicted to be structurally similar. Structural similarity searches were performed using the DALI server [35].

Interaction between effector pairs from two structural families

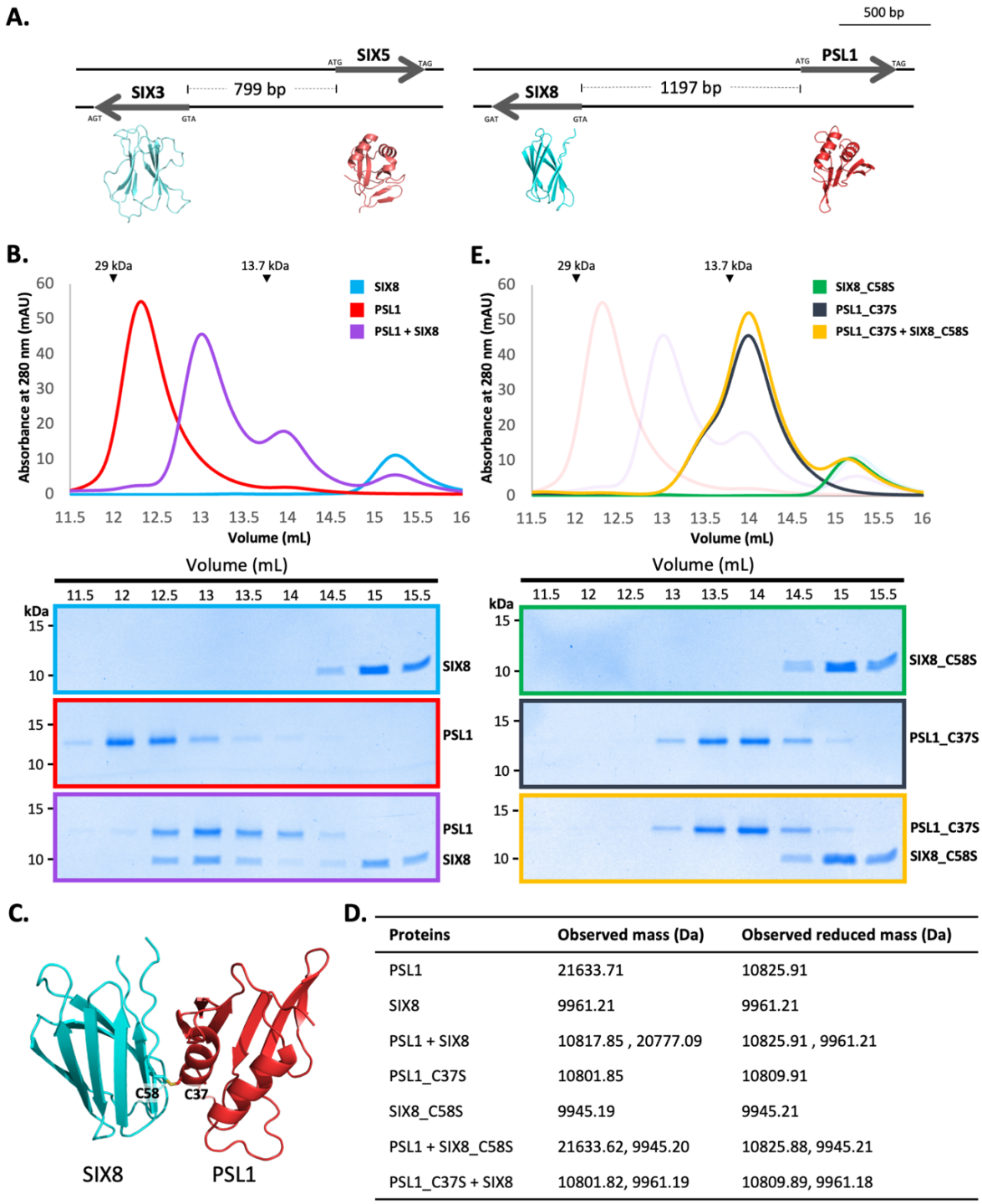
In *Fol, Avr2* and *SIX5*, and *SIX8* and *PSL1* from a similar head-to-head relationship in the genome with shared promoters and are divergently-transcribed (Fig 4A) [17, 18]. Previously, studies concerning Avr2 and SIX5 have demonstrated that the proteins function together and interact directly (via yeast-two-hybrid (Y2H) analysis) [10]. Homologues of SIX8 and PSL1 from *Focn* (SIX8 and PSE1) are also functionally dependent on each other, however in this case an interaction could not be established in yeast [18]. Here we demonstrate that both protein pairs containing a ToxA-like family member (Avr2, SIX8) and a structural family 4 member (SIX5, PSL1). Considering the predicted structural similarities, we were interested in testing whether *Fol* SIX8 and PSL1 interact.

We produced *Fol* SIX8$^{19\text{-}141}$ (S1E Fig) and PSL1$^{18\text{-}111}$ (S1F Fig) in *E. coli*. SIX8 has a putative pro-domain, which we removed resulting in the production of a stable ~10 kDa protein (SIX8$^{50\text{-}141}$). To determine whether SIX8 and PSL1 interact, the purified proteins alone or co-incubated were analysed by size exclusion chromatography (SEC) (Fig 4B). The elution profile of PSL1 shows a major peak (~12.25 mL) at a volume consistent with a dimeric form of the protein, while SIX8 shows a major peak (~15 mL) consistent with a monomer (Fig 4B). Strikingly, when incubated together the major protein peaks migrate to ~12.8 mL. SDS-PAGE analysis confirmed that presence of PSL1 and SIX8, indicating that the migration of both proteins on SEC is altered after incubation. These data are consistent with PSL1 and SIX8 forming a heterodimer.

To understanding the structural basis of the interaction; we attempted to solve the structure of the complex, but we were unable to obtain crystals of PLS1 and SIX8 alone or in complex. We subsequently utilised Alphafold2-Multimer [49] through ColabFold [50], to model the interaction. Manual inspection of the top 5 models (S10A Fig, top model shown Fig. 4C) demonstrated that the thiol side chain of a free cysteine in PSL1 (cys 37) and SIX8 (cys 58) co-localised in the dimer interface, suggesting that an inter-disulfide bond may mediated the interaction. To test this, we performed intact mass spectrometry of SIX8 and PSL1 (alone and post incubation) under non-reduced and reducing conditions. The mass observed from the incubated SIX8 and PSL1 non-reduced sample contained two species. The predominant species was consistent with the combined molecular weight of SIX8 and PSL1 (20777 Da), the other represented PSL1 (~10818 Da) which contains 4 intra-disulfide bonds (Fig 4D, S9G-H Fig). Interestingly, MS analysis of PSL1 alone sample demonstrated that the protein forms an exclusive dimer mediated by an intermolecular disulfide bond (Fig 4C, S9G-H Fig). Collectively, these data demonstrated that the SIX8-PSL1 heterodimer is mediated via a

disulfide bond, and that SIX8 can disrupt the PSL1 homodimer. We also showed that SIX8 and PSL1 failed to form a heterodimer with an unrelated protein containing a free cysteine, suggesting specificity in the interaction (S9I-L Fig).

To confirm the involvement of the predicted residues involved we produced recombinant protein of cysteine mutants of PSL1 (PSL1_C37S$^{18-111}$) and SIX8 (SIX8_C58S$^{50-141}$) and repeated the analysis (Fig 4E). The elution profile of PSL1_C37S shows a major peak at ~14 mL consistent with the monomeric form of the protein demonstrating that this residue is required for PSL1 dimerisation (Fig 4B). When PSL1_C37S was incubated with SIX8_C37S or SIX8 alone, the heterodimer was not resolved via SEC (Fig 4D, S10B Fig). This was further confirmed using MS (Fig 4C), and demonstrates that these residues mediate the heterodimer. Interestingly, we were able to crystallise SIX8_C58S$^{50-141}$ (S1E Fig) and PSL1_C37S (S1F Fig). The SIX8_C58S$^{50-141}$ crystals diffracted to a resolution of 1.28 Å, and we solved the structure of SIX8$^{50-141}$ using the AlphaFold2 model as a template (S10C Fig). The SIX8 structure consists of seven β-strands arranged in a β-sandwich and the structure confirms the inclusion of SIX8 within the ToxA-like structural family (S10D Fig).



Proteins	Observed mass (Da)	Observed reduced mass (Da)
PSL1	21633.71	10825.91
SIX8	9961.21	9961.21
PSL1 + SIX8	10817.85 , 20777.09	10825.91 , 9961.21
PSL1_C37S	10801.85	10809.91
SIX8_C58S	9945.19	9945.21
PSL1 + SIX8_C58S	21633.62, 9945.20	10825.88, 9945.21
PSL1_C37S + SIX8	10801.82, 9961.19	10809.89, 9961.18

Fig 4. PSL1 and SIX8 interact *in vitro* mediated by an intermolecular disulfide bond. (A) Schematic representation of the Avr2 (SIX3) – SIX5 and SIX8 – PSL1 loci within *Fol*. AlphaFold2 models or experimentally solved structures are shown underneath. **(B)** Top panels: Size exclusion chromatograms of PSL1 alone (red), SIX8 alone (blue), PSL1 and SIX8 (purple) (following a 30 min incubation) separated over a Superdex S75 Increase SEC column. Equal concentrations of the protein were used (note the absorbance of SIX8 @ 280nM is ~0.3 resulting in a smaller absorbance and peak height). Indicated sizes above the chromatogram

are based on protein standards run under similar conditions as presented in the manufacturer's column guidelines. Bottom panels: Coomassie-stained SDS-PAGE gels depicting samples taken from 500 µL fractions corresponding to the volumes indicated above the gels, with molecular weights (left) and proteins (right) annotated. **(C)** Model of the SIX8-PSL1 complex generated by AlphaFold2-Multimer (top model shown), co-localisation of cys 58 from SIX8 and cys 37 from PSL1 shown in stick **(D)** Observed masses of PSL1 and SIX8 protein mixtures by intact mass spectrometry (MS). Samples were treated with or without the reducing agent DTT prior to MS. The deconvoluted mass spectra of all proteins can be found in S9 Fig. **(E)** As for **(B)** but with PSL1_C37S (black), SIX8_C58S (green), and PSL1_C37S and SIX8_C58S (yellow)

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 [51]. An emerging theme in fungal effector biology is the classification of effectors into families based on structural similarity [52]. 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, 53, 54], ToxA-like [25-27], LARS [31] 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 and Sordariomycetes [28, 55]. Similarly, LARS effectors, defined in *Leptosphaeria maculans* and *Fulvia fulva,* have structural homologues predicted in at least 13 different fungal species [31]. Based on sequence homologues alone, FOLD effectors are well dispersed in fungi with homologues amongst the Sordariomycetes including many *formae speciales* of *F. oxysporum*, *Colletotrichum* and *Ustilaginoidea*. Based on structural comparison of the Alphafold2 structural database we show that is extended to fungi in three Divisions, including plant pathogens and symbionts.

Effector structure prediction

Experimentally determining 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 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 [56-59].

Work by Bauer and colleagues, prior to the release of Alphafold2, 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 [56]. Seong and Krasileva used similar structural modelling approaches to predict the folds of ~70% of the *Magnaporthe oryzae* secretome [57]. 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*. Making use of AlphaFold2, Yan and colleagues show that structurally conserved effectors, including the MAX effector family, from *M. oryzae* are temporally co-expressed during the infection process [58]. In the largest comparison study to date, Seong and Krasileva carried out a large comparative structural genomics study of fungal effectors utilising AlphaFold2 [59]. Their findings support the hypothesis that the structurally conserved effector families are the result of divergent evolution and support previous finding that the structural landscape of effectors is more limited than what is suggested by sequence-diversification.

Here, we were in a unique position to apply and benchmark AlphaFold2 against experimentally determined structures for *Fol* effector prediction. We subsequently used AlphaFold2 to demonstrate that, within the repertoire of effectors we tested, up to five sequence-unrelated structural families are secreted during *Fol* infection. There are numerous caveats in relying solely on AlphaFold2 to generate structural models of effectors. The accuracy of models generated by AlphaFold2 can decline in cases with low numbers of homologues (~30 sequences in the multiple sequence alignment) [36]. 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 experimentally determined in our lab to AlphaFold2 models. In some instances, the models are wrong, for example AvrSr50 [60], however, in these cases the AlphaFold2 predictions reported low confidence scores, an important criterion for assessment of model reliability. Despite this, AlphaFold2 models were critical in solving the structure of SIX6 and SIX8, as templates for molecular replacement. This negated the need to derivatise our crystals, a process that we had struggled with for SIX6 crystals, significantly reducing the time and research effort to determine the experimental structures.

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) [61, 62], while the C-domains have structural similarities with tumour necrosis factors (PDB code: 6X83) [63] and carbohydrate-binding lectins (PDB code: 2WQ4) [64]. 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) [45, 65], 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 KP6 effectors and heavy metal associated (HMA) domains. Metal binding within HMA domains is facilitated by conserved cysteine residues [66], 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) [67, 68] and K^+ channel-blocking scorpion toxins (PDB code: 1J5J, PDB code: 2AXK) [69, 70]. FOXG_18699 has structural homology with the C-terminal domain of bacterial arginine repressors (PDB code: 1XXB, PDB code: 3CAG) [71, 72].

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 PSE1 (PSL1-homolouge), 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. The AlphaFold2-multimer models of the SIX8 and PSL1 heterodimer, drew our attention to the inter-disulfide bond between SIX8 and PSL1 required for the interaction, which we confirmed experimentally. While these residues are conserved in *Focn* SIX8 and PSE1, the Avr2 structure and SIX5 model lack free cysteine residues, suggesting a different mode of interaction.

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 AlphaFold2 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 (S11 Fig), which suggests that SIX12 may also be a family 4 member. We therefore speculate that SIX7 and SIX12 may function together, as described for the Avr2/SIX5 and SIX8/PSL1 pairs.

Are experimentally derived effector structures worth the effort?

The potential of machine-learning structural-prediction programs, such as AlphaFold2, heralds an exciting era, especially for a field that has long suffered from a lack of prediction power based on effector sequences. A question now emerges; when prediction model confidence is high, should we bother solving structures experimentally? The answer to such a question will always depend on what the structure is being used for. Ultimately, structural models, whether experimentally or computationally derived, represent information to base and/or develop a hypothesis to subsequently test. Here we demonstrate the power of structure prediction in combination with experimentation, both for validating models and understanding protein:protein interaction interfaces. One interesting observation we made was that while the AphaFold2-multimer models of the SIX8 and PSL1 heterodimer were sufficient to highlight the cysteine residues required for mediating the interaction, the models and interaction interfaces differed significantly (S10A Fig). When the modelling was repeated with the SIX8^{C58S} experimentally derived structure included as a template, the interaction models and heterodimer interface were of higher quality and essentially identical (S10E Fig). This observation can be retrospectively reconciled. The region of SIX8 involved in the interaction with PSL1 was modelled incorrectly by AlphaFold2 when compared to the structure (S10D

Fig). Collectively, these data highlight that some models are good enough, but others maybe better.

Materials and methods

Vectors and gene constructs

SIX6, Avr1Thrombin, SIX6-TEV, SIX8Thrombin, SIX8_C58SThrombin, PSL1, PSL1_C37S and SIX13 coding sequences (without their signal peptides as determined by SignalP-5.0) were codon optimised for expression in *E. coli* and synthesised with Golden-Gate compatible overhangs by Integrated DNA Technologies (IDT, Coralville, Iowa, USA) (S4 Table). The Kex2 cleavage motif of Avr1 and SIX8 were replaced with a thrombin cleavage motif, and TEV protease cleavage motif for SIX6 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. All of the primers were synthesised by IDT (Coralville, Iowa, USA) (S5 Table). All genes were cloned into a modified, Golden-Gate-compatible, pOPIN expression vector [73]. 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 [74]. 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 [34]. 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 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. SIX6 with an internal TEV protease cleavage site for pro-domain removal was processed with TEV protease (produced in-house) until fully cleaved. Mature proteins encompass 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 or

3 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.

Intact mass spectrometry

For untreated samples, proteins were adjusted to a final concentration of 6 µM in 0.1% (v/v) formic acid (FA) for HPLC-MS analysis. For reduced samples, DTT was added to the protein to a final concentration of 10 mM. Proteins were incubated at 60°C for 30 minutes and adjusted to 6 µM in 0.1% (v/v) FA. Intact mass spectrometry on all proteins was carried out as described previously [34]. Data were analysed using the Free Style v.1.4 (Thermo Fisher Scientific) protein reconstruct tool across a mass range of m/z 500 – 2000 and compared against the theoretical (sequence based) monoisotopic mass.

Circular dichroism (CD) spectroscopy

The CD spectra of purified effectors of interest were recorded on a Chirascan spectrometer (Applied Photophysics Ltd., UK) at 20ºC. Samples were diluted to 10 µM in a 20 mM sodium phosphate buffer at pH 8.0. Measurements were taken at 1 nm wavelength increments from 190 nm to 260 nm. A cell with a pathlength of 1 mm, a bandwidth of 0.5 nm and response time of 4 s were used, with 3 accumulations. The data were averaged and corrected for buffer baseline contribution, and visualised using the webserver CAPITO tool with data smoothing [75].

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$^{18-242}$, Avr1$^{59-242}$, SIX8$^{50-141}$ and PSL1$^{18-111}$, 15 mg/mL for SIX6$^{17-225}$ and SIX6$^{58-225}$, 25 mg/mL for SIX8_C58S$^{19-141}$, 18 mg/mL for SIX8_C58S$^{50-141}$ and PSL1_C37S$^{18-111}$, 14 mg/mL for SIX8-PSL1 complex and 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$^{17-225}$, SIX8$^{50-141}$, PSL1$^{18-111}$, SIX8-PSL1 complex 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 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). 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 were obtained in 0.2 M ammonium tartrate and 20% (w/v) PEG 3350 (SG1 screen: condition G9) 40 days after initial setup. Crystals were picked directly from the sparse matrix screen. For SIX8_C58S$^{50-141}$, crystals were obtained in 0.17 M ammonium sulfate, 15% (w/v) glycerol and 25.5% (w/v) PEG 4000 (JCSG screen: condition D9) a week after initial setup. Crystals were picked directly from the sparse matrix screen. 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. For PSL1_C37S$^{18-111}$, crystals were obtained in 70% (w/v) MPD and 0.1 M HEPES pH 7.5 within 3 days after initial setup. Crystal optimisation was carried out in 24-well hanging-drop vapour-diffusion plates at 18°C. The final optimised condition for PSL1_C37S$^{18-111}$ was 62% (w/v) MPD and 0.1 M HEPES pH 7.5 at a protein concentration of 17.5 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. No cryoprotecting was necessary for SIX8_C58S$^{50-141}$ and PSL1_C37S$^{18-111}$ crystals. 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 [76] (S1 Table). The datasets were processed in XDS [77] and scaled with Aimless in the CCP4 suite [78, 79]. The CRANK2 pipeline in CCP4 was used for bromide-based SAD phasing [80, 81]. Models were then refined using phenix.refine in the PHENIX package [82] and model building between refinement rounds was done in COOT [83]. The models were used as a template for molecular replacement against high resolution native datasets collected on the MX2 beamline at the Australian Synchrotron [84]. Automatic model building was done using AutoBuild [85], and subsequent models were refined with phenix.refine and COOT. For SIX6$^{58-225}$ and SIX8_C58S$^{50-141}$, high confidence *ab initio* models were generated with AlphaFold2 (S3 Fig), which was used as a template for molecular replacement against a native dataset collected on the MX2 beamline at the Australian Synchrotron. The resultant structure was refined as described above.

Structural modelling and structural alignment

Structural models were generated with Google DeepMind's AlphaFold2 using the amino acid sequences of SIX effectors and candidates without the signal peptide, as predicted by SignalP-5.0 [86] and predicted pro-domain by searching for a Kex2 cleavage motif (KR, RR or LxxR) if present [32] (S3 Table; S6 Fig). For AlphaFold2 predictions the full databases were used for multiple sequence alignment (MSA) construction. All templates downloaded on July 20, 2021 were allowed for structural modelling. For each of the proteins, we produced five models and selected the best model (ranked_0.pdb). Pairwise alignments of the structural models generated by AlphaFold2 and the experimentally determined structures of Avr1 (PDB code: 7T6A), Avr3 (PDB code: 7T69), SIX6 (PDB code: 8EBB) and SIX8 (PDB code: 8EB9) were generated using the DALI server all against all function [35]. Structural similarity between the pairwise alignments were measured using Z-scores from the DALI server.

Distribution of FOLD family members across fungi

Structure based searches to determine the distribution of FOLD effectors across other phytopathogens was carried out by searching the experimentally determined Avr1, Avr3 and SIX6 structures against available structure databases (Uniprot50, Proteome, Swiss-Prot) using the Foldseek webserver [43] using a 3Di search limited to fungi. An e-value cut off of 0.01

was used, and non-plant associated fungi were removed as well as duplicated results for final analysis. Proteins below 100 amino acids, and above 500 amino acids were filtered out and remaining structural hits were manually inspected for similarity to FOLD effectors.

Interaction studies between PSL1 and SIX8

To investigate whether PSL1 and SIX8 interacted *in vitro* ~140 µg of PSL1$^{18-111}$ and SIX8$^{50-141}$ individually, and ~140 µg PSL1$^{18-111}$ and 140 µg of SIX8$^{50-141}$ together 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. To investigate the residues responsible for the interaction, SIX8_C58S$^{50-141}$ and PSL1_C37S$^{18-111}$ mutants were used instead. Samples across the peaks were then analysed by Coomassie-stained SDS-PAGE. To investigate the mode of interaction, PSL1 and SIX8 proteins and mutants at 10 µM were incubated individually or together for 1 hour at room temperature. An unrelated protein with a free cysteine (AvrSr50^{RKQQC}) [60] was used to assess the specificity of the PSL1-SIX8 interaction. Proteins were analysed by intact mass spectrometry with or without the addition of DTT as described above.

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. This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. The authors acknowledge use of the Australian Synchrotron MX facility and thank the staff for their support. The coordinates and structure factors for Avr1, Avr3, SIX6 and SIX8 have been deposited in the Protein Data Bank with accession number 7T6A, 7T69, 8EBB and 8EB9, respectively.

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 and Avr3 coloured from N- (blue) to C-terminus (red) in the top panel showing the dual-domain structural fold, and bottom panels showing secondary structure topology map of Avr1 **(A)** and Avr3 **(B)**, respectively. 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. 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 of the DALI server [35].

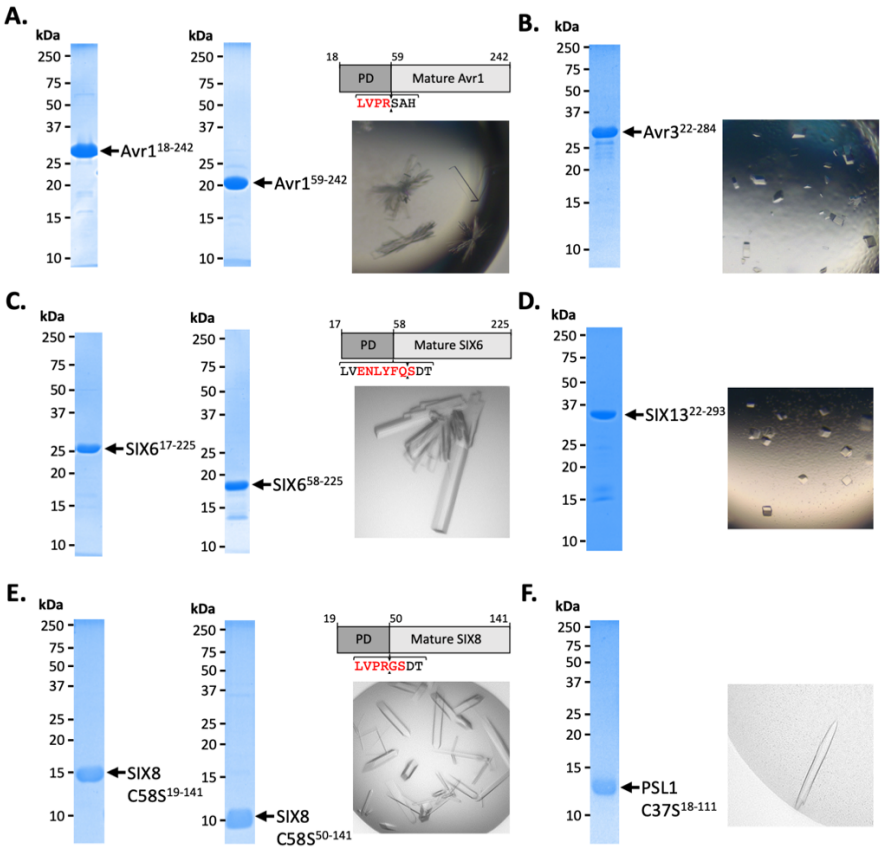
Fig 2. FOLD effector family is distributed within *Fusarium oxysporum* and other fungi. **(A)** Amino acid sequence alignment of Avr1, Avr3, SIX6 and SIX13 show a common cysteine spacing at the N-terminus. The alignment is split into the N-terminus (N-domain; top panel) and C-terminus (C-domain; bottom 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 crystal structure and **(C)** SIX13 model predicted by AlphaFold2 showing the dual-domain structural fold, transitioning from blue (N-terminus) to red (C-terminus). **(C)** Structure-guided search for putative FOLD effectors across fungi using Foldseek webserver. Size of circles represent abundance with genus. **(D)** Superposition (structural alignment) of representative putative FOLD effectors from the divisions Glomeromycota and Basidiomycota with Avr1 in ribbon representation. Putative FOLD protein from *Rhizophagus clarus* (UniProt: A0A2Z6QDJ0) in light blue, and *Piloderma croceum* (UniProt: A0A0C3C2B2) in green. FOLD structural alignment (right), N-domain only (middle), C-domain only (right).

Fig 3. Identification of new putative structural families within the SIX effectors. Heat maps showing the structural similarity of Structures and AlphaFold2 models of the **(A)** SIX effectors and **(B)** effector candidates from *Fol* in a structural pairwise alignment. Structural similarity was measured with Z-scores. A cutoff Z-score of 2 was applied for defining structural families. Z-score scale is shown in a grey to red spectrum. **(C)** Cartoon representation of the ToxA-like effectors from *Fol*. AlphaFold2 models of SIX7, SIX8 and FOXGR_015533 effector candidate are putative members of the ToxA-like effector family. The crystal structure

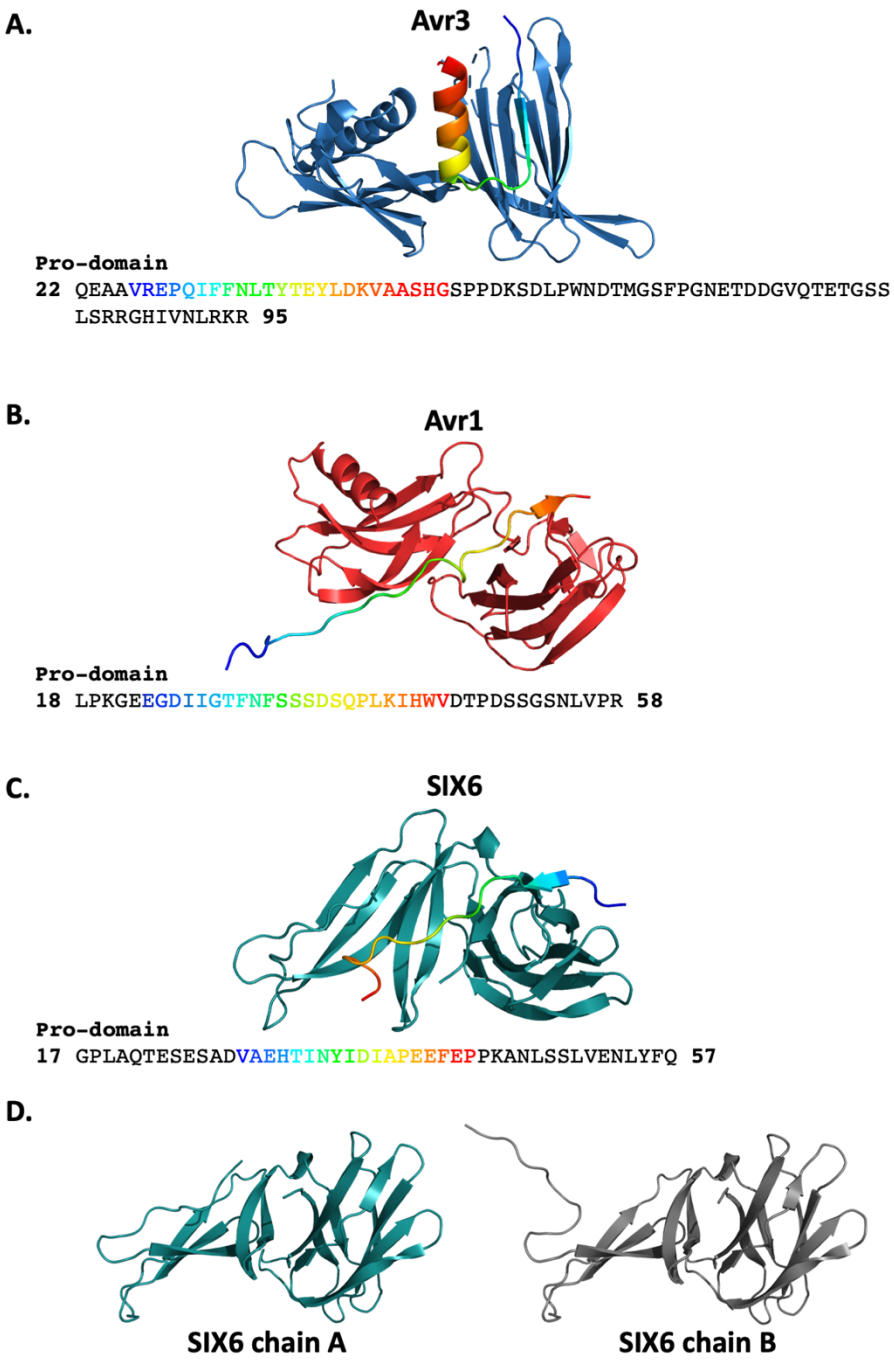
of Avr2 [25], another member of the ToxA-like effector family, is shown in green for comparison. Cartoon representations of **(D)** Family 3, **(E)** Family 4 and **(F)** Family 5 consisting of members that are predicted to be structurally similar. Structural similarity searches were performed using the DALI server [35].

Fig 4. PSL1 and SIX8 interact *in vitro* mediated by an intermolecular disulfide bond. (A) Schematic representation of the Avr2 (SIX3) – SIX5 and SIX8 – PSL1 loci within *Fol*. AlphaFold2 models or experimentally solved structures are shown underneath. **(B)** Top panels: Size exclusion chromatograms of PSL1 alone (red), SIX8 alone (blue), PSL1 and SIX8 (purple) (following a 30 min incubation) separated over a Superdex S75 Increase SEC column. Equal concentrations of the protein were used (note the absorbance of SIX8 @ 280nM is ~0.3 resulting in a smaller absorbance and peak height). Indicated sizes above the chromatogram are based on protein standards run under similar conditions as presented in the manufacturer's column guidelines. Bottom panels: Coomassie-stained SDS-PAGE gels depicting samples taken from 500 µL fractions corresponding to the volumes indicated above the gels, with molecular weights (left) and proteins (weight) annotated. **(C)** Model of the SIX8-PSL1 complex generated by AlphaFold2-Multimer (top model shown), co-localisation of cys 58 from SIX8 and cys 37 from PSL1 shown in stick **(D)** Observed masses of PSL1 and SIX8 protein mixtures by intact mass spectrometry (MS). Samples were treated with or without the reducing agent DTT prior to MS. The deconvoluted mass spectra of all proteins can be found in S9 Fig. **(E)** As for **(B)** but with PSL1_C37S (black), SIX8_C58S (green), and PSL1_C37S and SIX8_C58S (yellow).

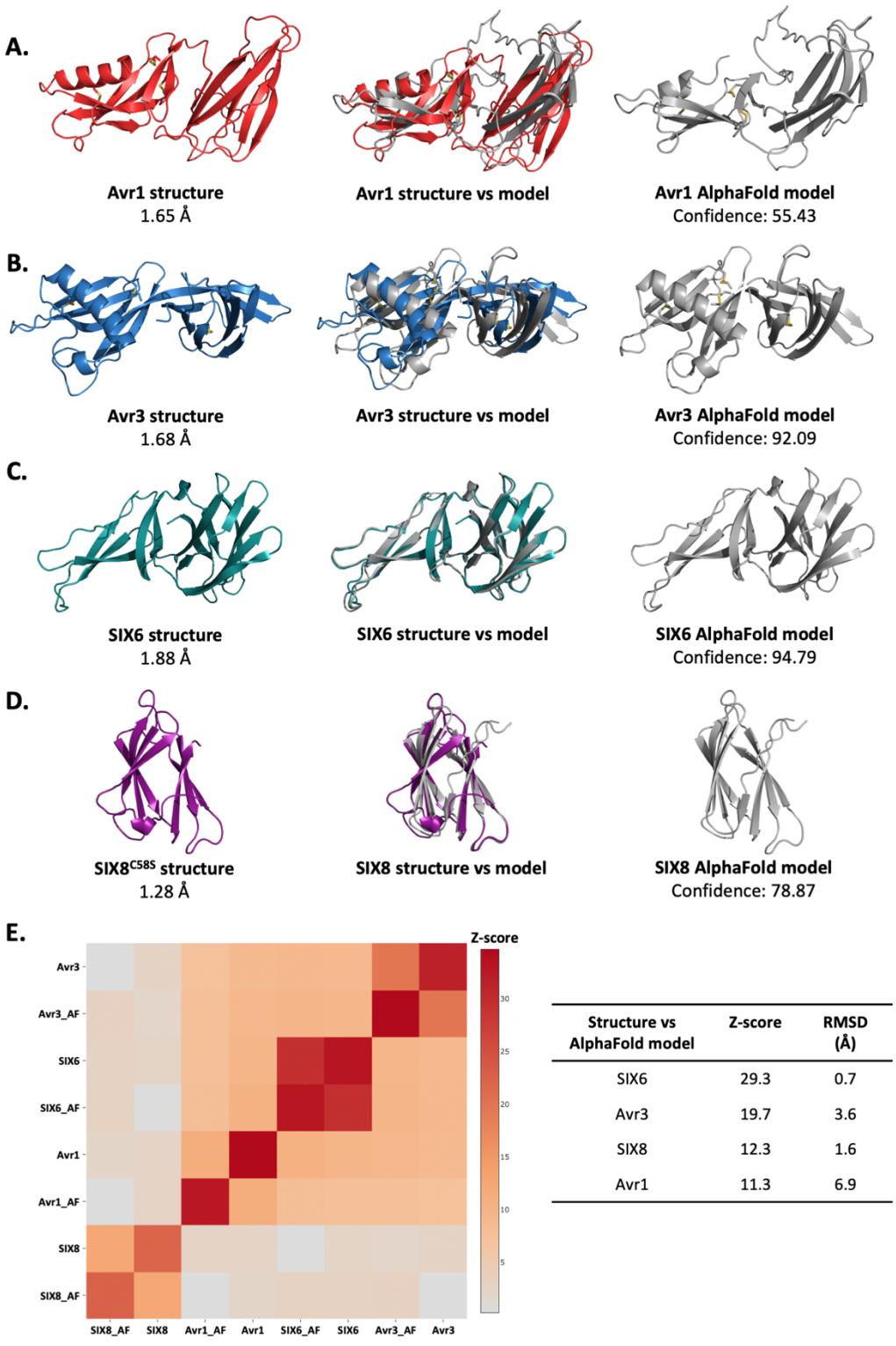
Supplementary Figures



S1 Fig. Crystallisation of Avr1, Avr3, SIX6, SIX8, SIX13 and PSL1. (A) Coomassie-stained gel showing full length Avr1 (left panel) and mature Avr1 cleaved *in vitro* with thrombin (middle panel). Schematic of Avr1 engineered with an internal thrombin cleavage site replacing the Kex2 cleavage motif (top right panel). Optimised crystals of Avr1$^{59-242}$ (Bottom right panel) **(B)** Coomassie-stained gel showing purified Avr3$^{22-284}$ used for crystallisation studies (left panel). Optimised crystals of Avr3 (right panel). **(C)** Coomassie-stained gel showing full length SIX6 (left panel) and mature SIX6 cleaved *in vitro* with TEV protease (middle panel). Schematic of SIX6 engineered with an internal TEV protease cleavage site replacing the Kex2 cleavage motif (top right panel). Optimised crystals of SIX6$^{58-225}$ (bottom right panel). **(D)** Coomassie-stained gel showing SIX13 protein (left panel). Optimised crystals of SIX13 (right panel). Kex2 protease was added to the protein at a 1:200 protease to protein ratio prior to crystal tray setup. **(E)** Coomassie-stained gel showing full length SIX8_C58S (left panel) and mature SIX8_C58S cleaved *in vitro* with thrombin (middle panel). Schematic of SIX8 engineered with an internal thrombin cleavage site replacing the Kex2 cleavage motif (top right panel). Optimised crystals of SIX8$^{50-141}$ (bottom right panel). **(F)** Coomassie-stained gel showing PSL1_C37S protein (left panel). Optimised crystals of PSL1_C37S (right panel).



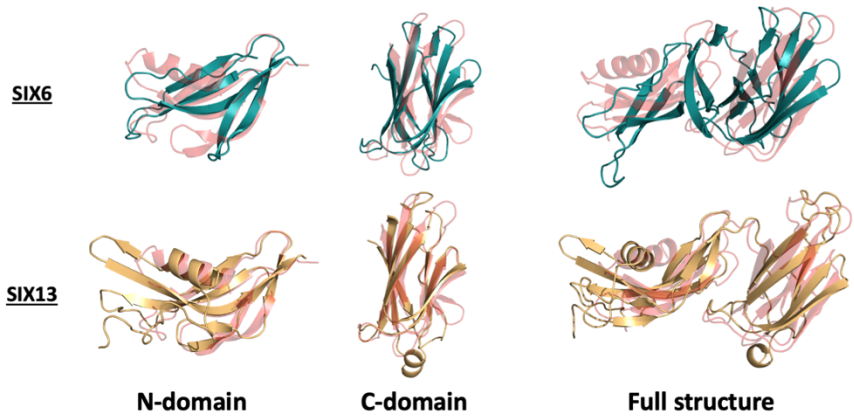
S2 Fig. Continuous electron density of the pro-domain is present in the crystal structures of Avr1, Avr3 and SIX6. Top panels: The crystal structure of **(A)** Avr3, **(B)** Avr1 and **(C)** SIX6 with the pro-domain shown in rainbow. Bottom panels: The amino acid sequence of the pro-domain of Avr3, Avr1 and SIX6 with residues observed in the electron density shown in rainbow text. Residues with no density observed are shown in black. For SIX6, electron density corresponding to the pro-domain was only associated to chain A. **(D)** Different orientations of the N-terminal region of SIX6 between chains A and B. Chain A was used in subsequent structural analysis.



Structure vs AlphaFold model	Z-score	RMSD (Å)
SIX6	29.3	0.7
Avr3	19.7	3.6
SIX8	12.3	1.6
Avr1	11.3	6.9

S3 Fig. Comparison of AlphaFold2 models against the experimentally solved structures of Avr1, Avr3, SIX6 and SIX8. The crystal structures of **(A)** Avr1, **(B)** Avr3, **(C)** SIX6 and **(D)** SIX8 (left panels) and AlphaFold2 models [36] (right panels). Crystal structures and AlphaFold2 models of the full structures (middle panels) were superimposed using the pairwise and all against all functions on the DALI server [35]. **(E)** Heat map of the structural

similarity between crystal structures and AlphaFold2 models (left panel). Z-score and RMSD values are shown in the right panel.

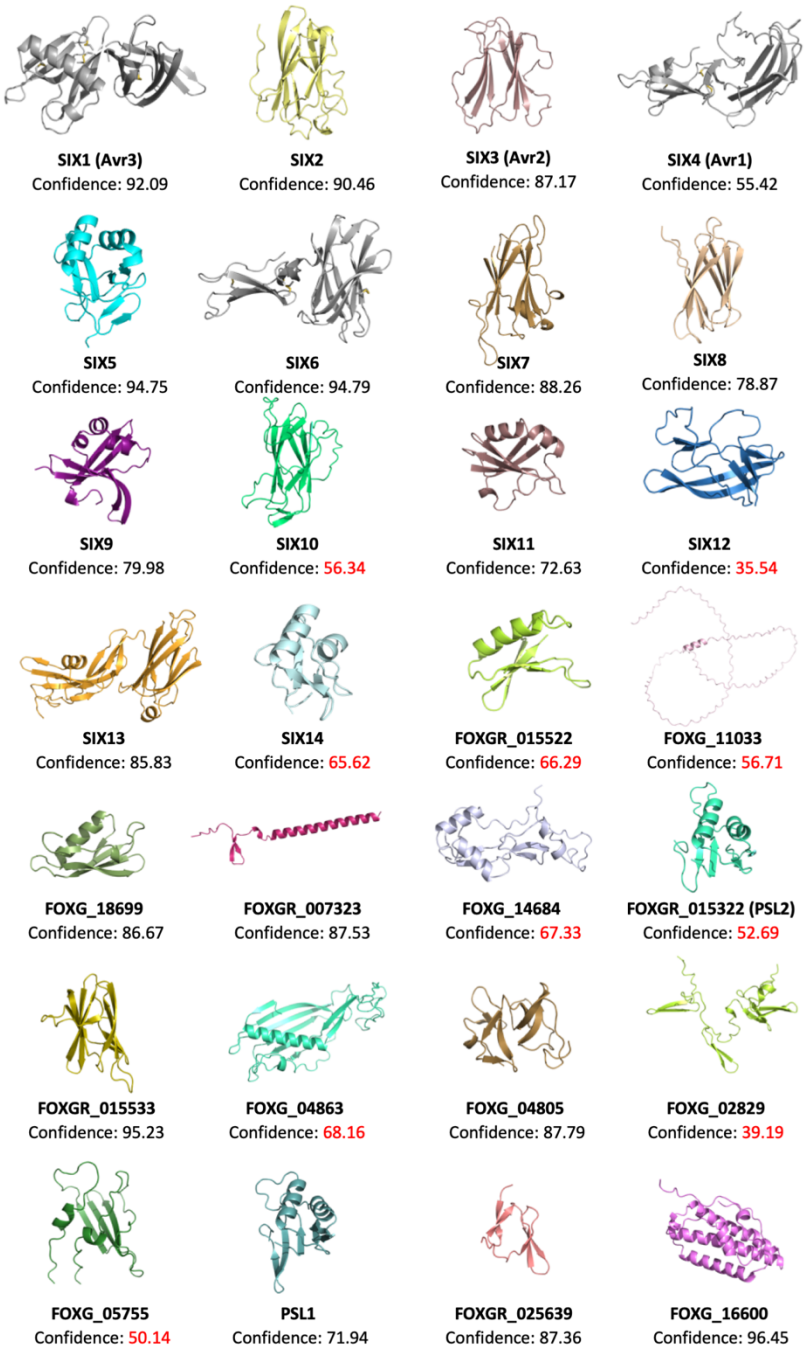


S4 Fig. Structural alignments of SIX6 and SIX13 with Avr1. (A) SIX6 crystal structure and **(B)** SIX13 AlphaFold2 model aligned with Avr1 using the N-domains alone (left panel), C-domains alone (middle panel) and full structure (right panel). Structural alignment was performed using the pairwise alignment function on the DALI server [35].

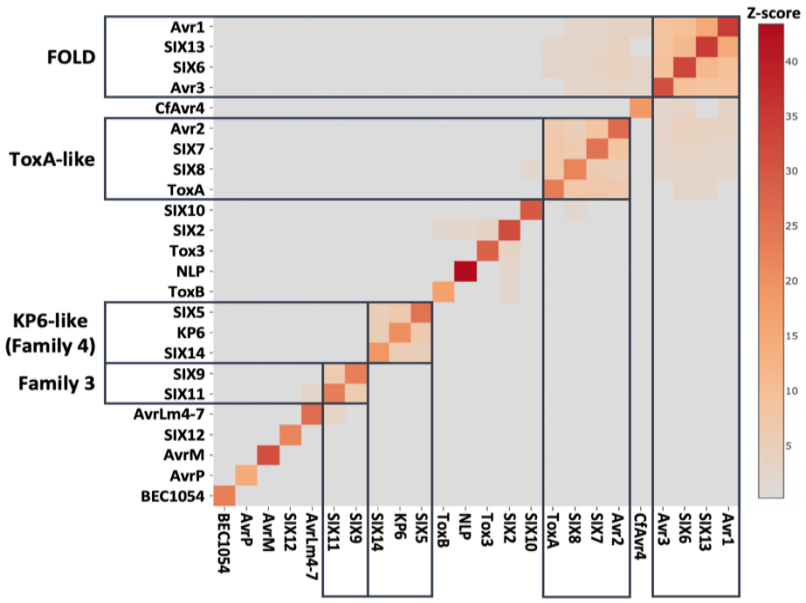
A.

	FOLD effectors			
	Avr1	**Avr3**	**SIX6**	**SIX13**
Fusarium oxysporum f. sp.		*pisi*		
		melonis		
		cubense TR4		
			vasinfectum	
	niveum		*niveum*	
			passiflorae	
		fragariae		*fragariae*
		conglutinans		
				medicaginis
	spinaciae			

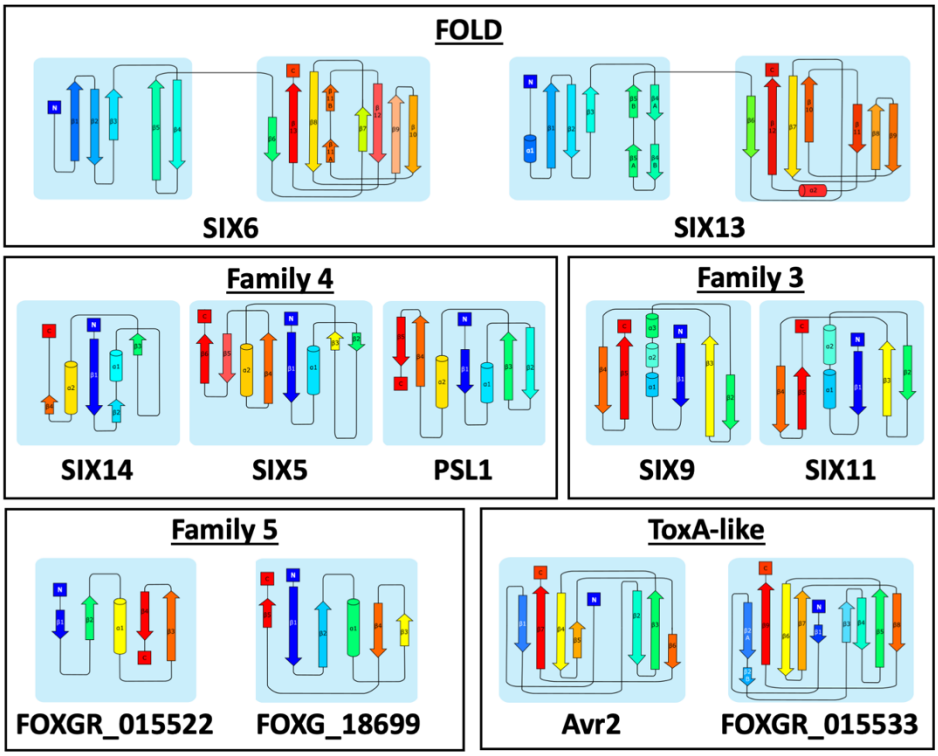
S5 Fig. (A) Homologues of FOLD effectors are dispersed across multiple *formae speciales* of *F. oxysporum.* Functional homologues of Avr1 (SIX4), Avr3 (SIX1), SIX6 and SIX13 reported in literature were assessed [7, 9, 37-40].



S6 Fig. AlphaFold2 models of all SIX effectors and effector candidates. Signal peptides were identified using SignalP-5.0 [86] and removed prior to amino acid sequences being input into AlphaFold2 [36]. Any putative pro-domains were identified by searching for a Kex2-like protease site [33] and removed. The sequence inputs used can be found in S3 Table.

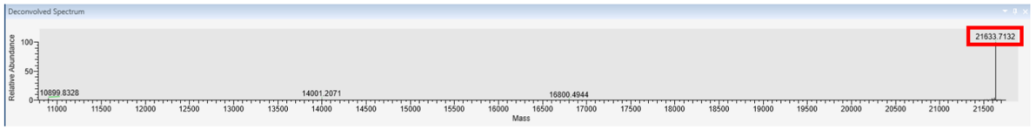


S7 Fig. Structural similarity of SIX effectors against representative solved effector structures from known structural families. The solved structures of Avr1, Avr2, Avr3 and SIX6, and AlphaFold2 models the remaining SIX effectors were compared with the structures of ToxA (ToxA-like), ToxB (MAX), Tox3 (Tox3-like), BEC1054 (RALPH), AvrLm4-7 (LARS), AvrP (Zinc finger), CfAvr4 (CBM14-like), AvrM (WY-like), NLP (Actinoporin-like) and KP6 (KP6-like). Structural alignment was performed using the all against all function on the DALI server [35]. Structural similarity was measured using Z-score. Groupings with Z-scores > 2 were outlined.

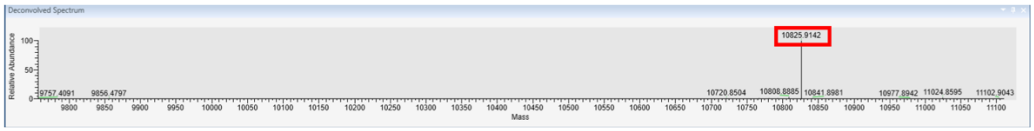


S8 Fig. Secondary structure topology maps of representative SIX structural family members. The β-strands and α-helices are represented by arrows and cylinders, respectively. The secondary structural elements are coloured in rainbow, from blue at the N-terminus to red at the C-terminus.

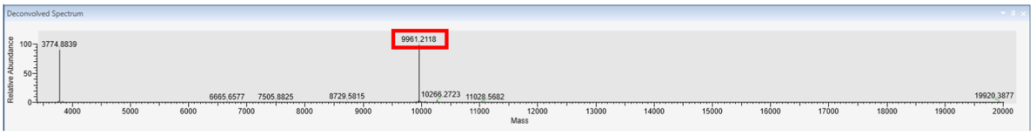
A. **PSL1**



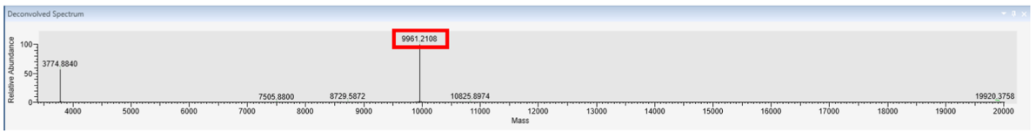
B. **PSL1 reduced**



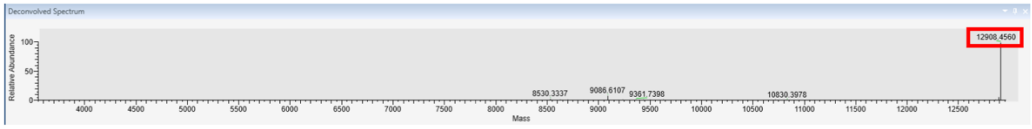
C. **SIX8**



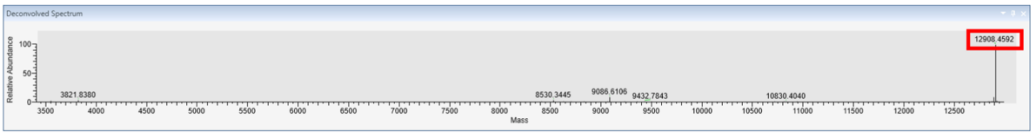
D. **SIX8 reduced**



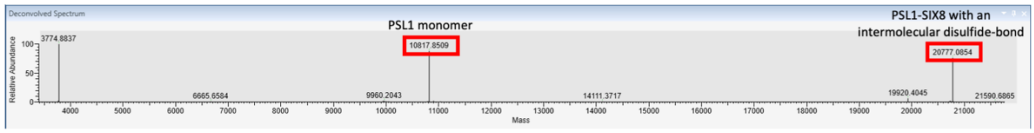
E. **AvrSr50^{RKQQC}**



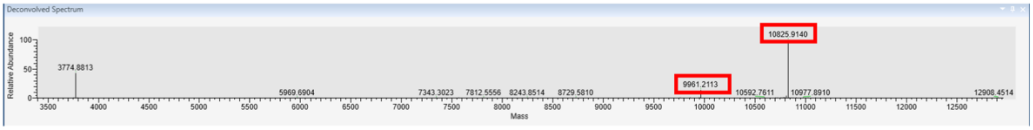
F. **AvrSr50^{RKQQC} reduced**



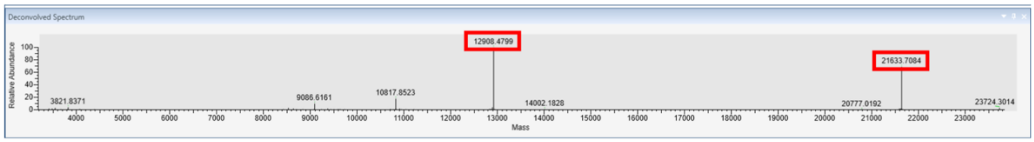
G. **PSL1 + SIX8**



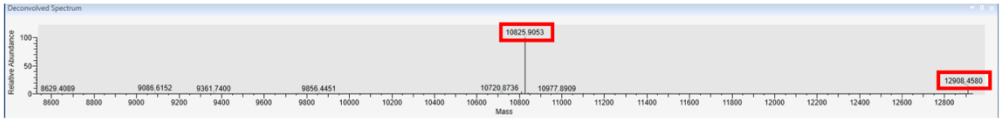
H. **PSL1 + SIX8 reduced**



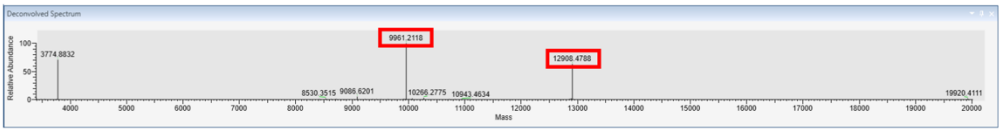
I. **PSL1 + AvrSr50^{RKQQC}**



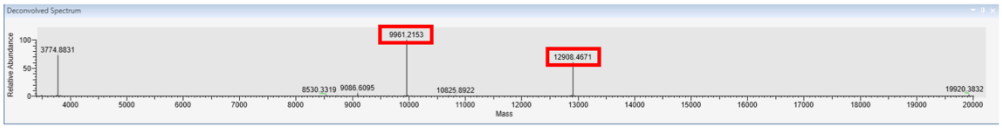
J. **PSL1 + AvrSr50^{RKQQC} reduced**



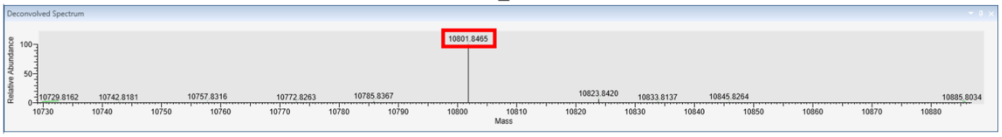
K. **SIX8 + AvrSr50^{RKQQC}**



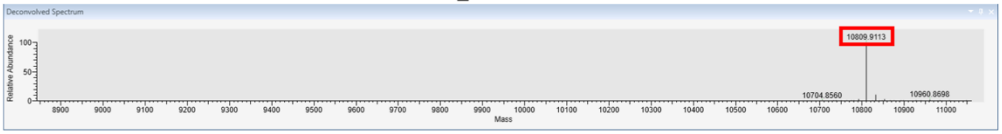
L. **SIX8 + AvrSr50^{RKQQC} reduced**



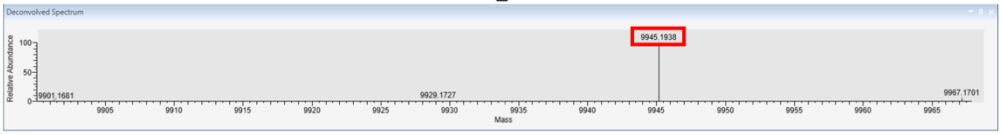
M. **PSL1_C37S**



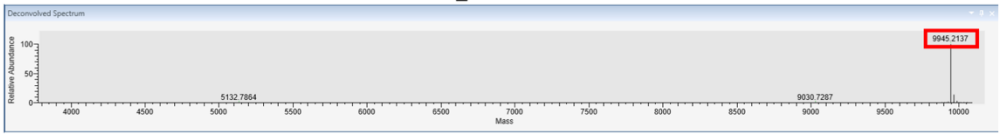
N. **PSL1_C37S reduced**



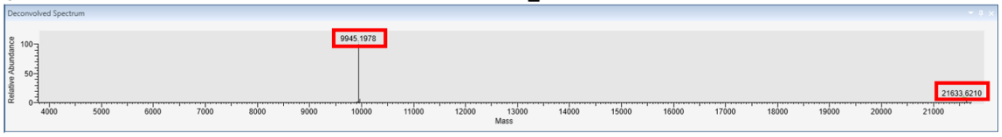
O. **SIX8_C58S**

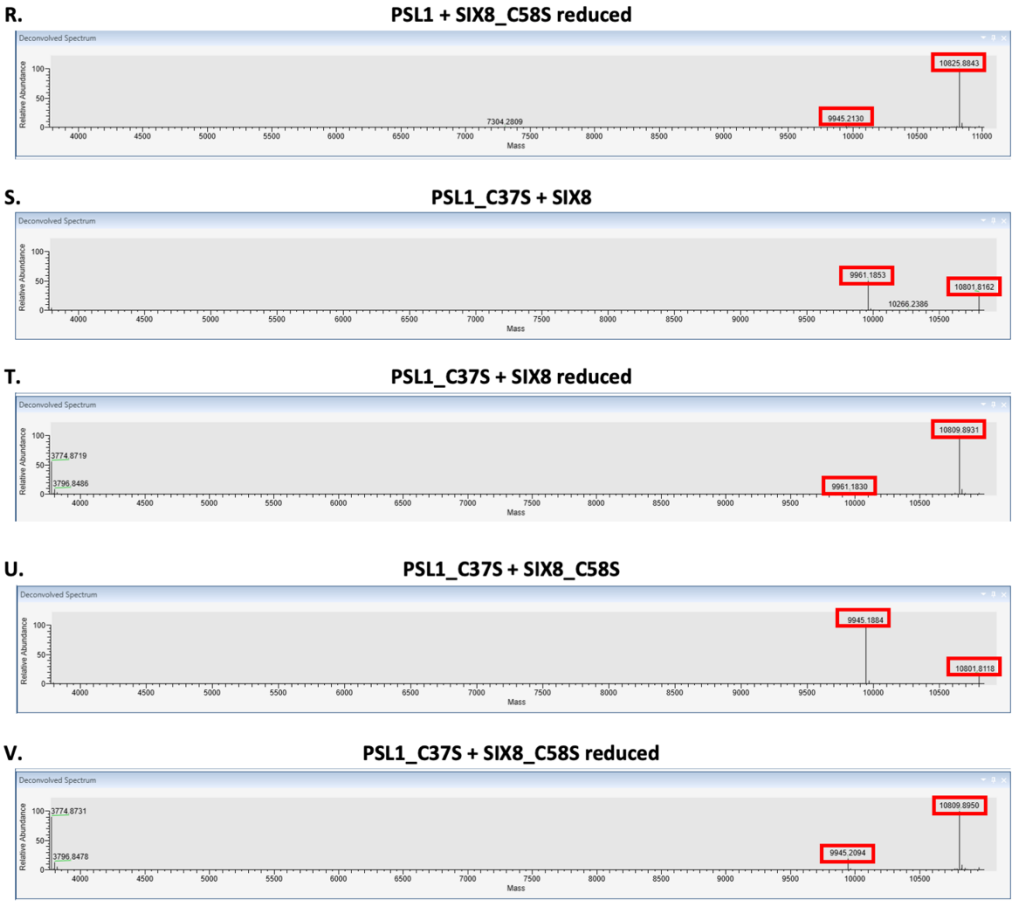


P. **SIX8_C58S reduced**

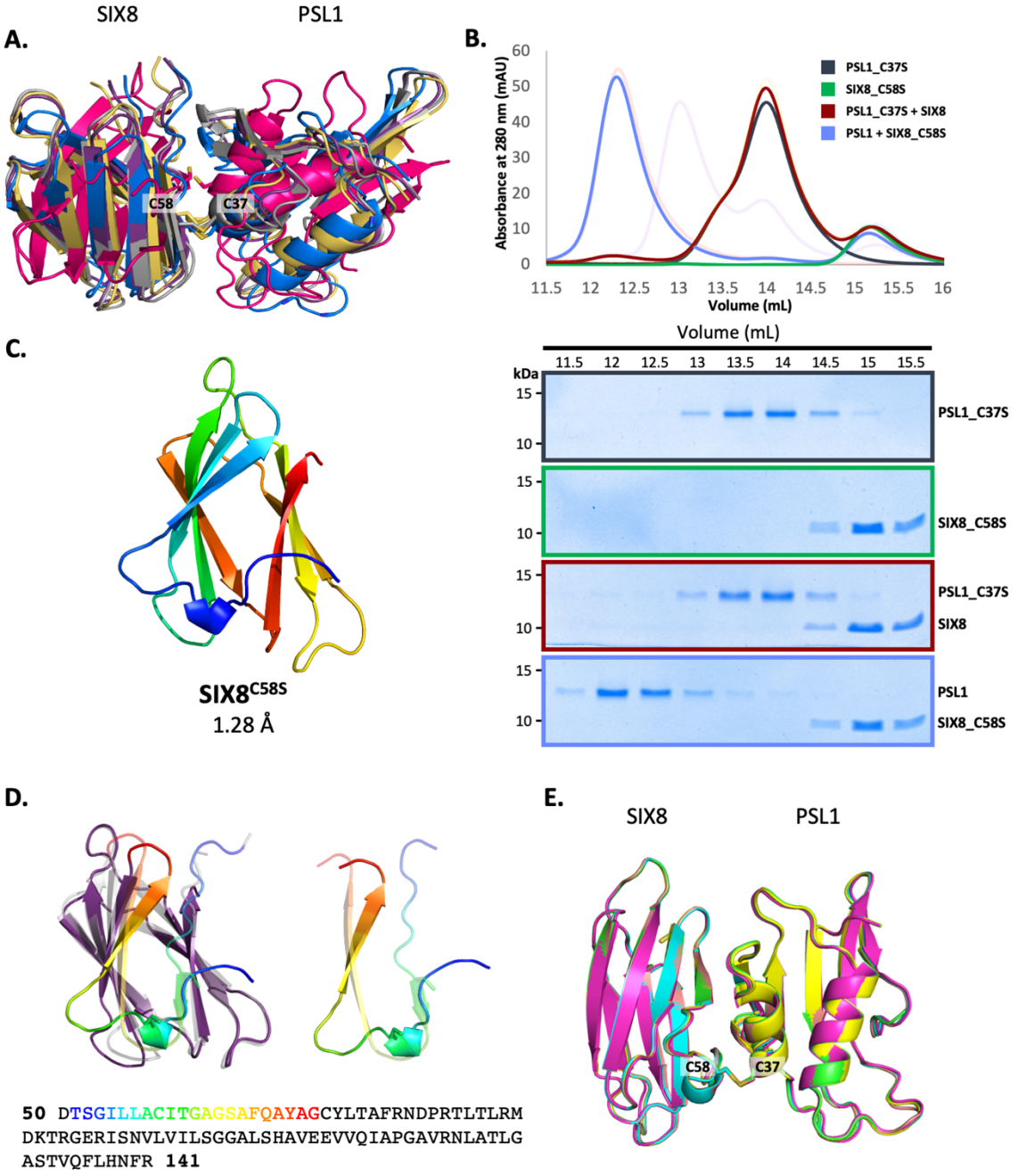


Q. **PSL1 + SIX8_C58S**



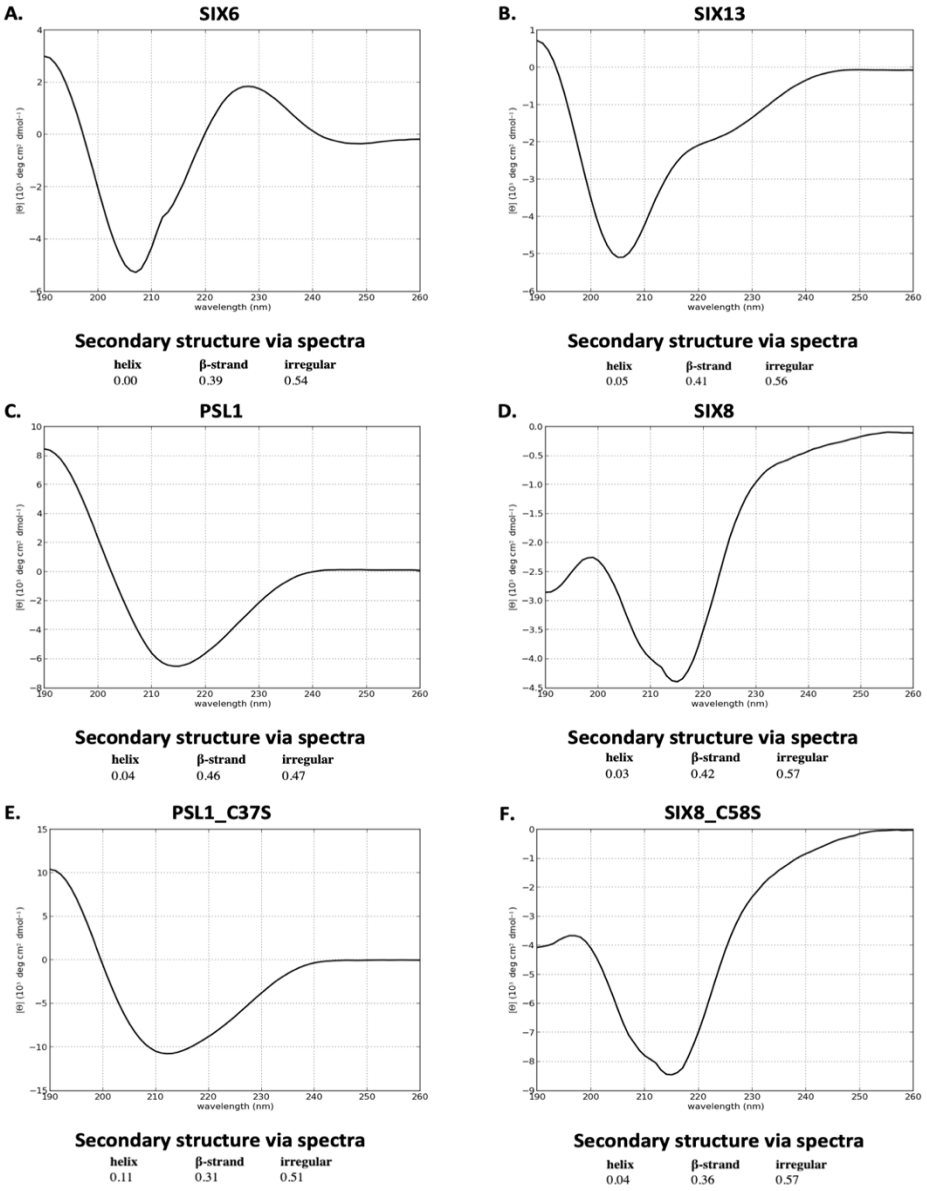


S9 Fig. Intact mass spectrometry analysis of the PSL1-SIX8 interaction. Deconvoluted mass spectra of **(A)** PSL1, **(B)** reduced PSL1, **(C)** SIX8, **(D)** reduced SIX8, **(E)** AvrSr50^{RKQQC}, **(F)** reduced AvrSr50^{RKQQC}, **(G)** PSL1 + SIX8, **(H)** reduced PSL1 + SIX8, **(I)** PSL1 + AvrSr50^{RKQQC}, **(J)** reduced PSL1 + AvrSr50^{RKQQC}, **(K)** SIX8 + AvrSr50^{RKQQC}, **(L)** reduced SIX8 + AvrSr50^{RKQQC}, **(M)** PSL1_C37S, **(N)** reduced PSL1_C37S, **(O)** SIX8_C58S, **(P)** reduced SIX8_C58S, **(Q)** PSL1 + SIX8_C58S, **(R)** reduced PSL1 + SIX8_C58S, **(S)** PSL1_C37S + SIX8, **(T)** reduced PSL1_C37S + SIX8, **(U)** PSL1_C37S + SIX8_C58S, **(V)** reduced PSL1_C37S + SIX8_C58S. Reduced samples were heated with DTT prior to running the samples.

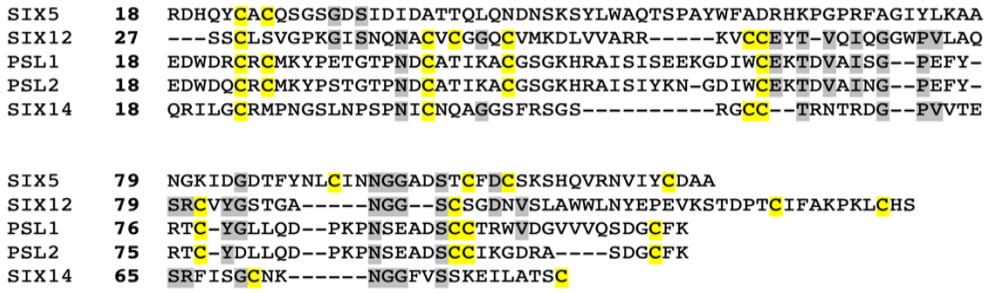


S10 Fig. Interaction between PSL1 and SIX8 mutants. (A) Model of the SIX8-PSL1 complex generated by AlphaFold2-Multimer (five models shown), co-localisation of cys 58 from SIX8 and cys 37 from PSL1 shown in stick. **(B)** Top panels: Size exclusion chromatograms of PSL1_C37S alone (black), SIX8_C58S alone (green), PSL1_C37S and SIX8 (maroon), and PSL1 and SIX8_C58S (light purple) following a 30 min incubation separated on a Superdex S75 Increase 10/300 SEC column. Equal concentrations of the protein were used (note the absorbance of SIX8 @ 280nM is ~0.3 resulting in a smaller absorbance and peak height). Indicated sizes above the chromatogram are based on protein standards run under similar conditions as presented in the manufacturer's column guidelines. Bottom panels:

Coomassie-stained SDS-PAGE gels depicting samples taken from 500 µL fractions corresponding to the volumes indicated above the gels, with molecular weights (left) and proteins (weight) annotated. **(C)** Cartoon representation of the crystal structure of SIX8^{C58S} at 1.28 Å resolution, coloured from N (blue) to C (red) terminus. **(D)** Comparison of the SIX8 structure and the AlphaFold2 model. Top panels: The SIX8 structure (purple) and AlphaFold2 model (grey) were superimposed using the DALI server [35]. The N-terminus is coloured in rainbow. The location of C58S is shown as a stick. Bottom panel: Amino acid sequence of SIX8 with residues of the N-terminus in rainbow corresponding to the structure. **(E)** Model of the SIX8-PSL1 complex generated by AlphaFold2-Multimer (five models shown), when the SIX8^{C58S} structure was used as a template. Co-localisation of cys 58 from SIX8 and cys 37 from PSL1 shown in stick.



S11 Fig. Circular dichroism analysis of purified recombinant proteins. CD spectra of the *Fusarium oxysporum* f. sp. *lycopersici* effectors **(A)** SIX6, **(B)** SIX13, **(C)** PSL1 and **(D)** SIX8 proteins and **(E)** PSL1_C37S and **(F)** SIX8_C58S mutants were plotted, and secondary structure elements analysed using the CAPITO webserver [75].



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, Avr3, SIX6 and SIX8.

	Avr1 Bromide soak - SAD	**Avr1 Native (7T6A)**	**Avr3 Bromide soak - SAD**	**Avr3 Native (7T69)**	**SIX6 (8EBB)**	**SIX8 (8EB9)**
Detector	Dectris EIGER 2 9M	Dectris EIGER 16M	Dectris EIGER 2 9M	Dectris EIGER 16M	Dectris EIGER 2 9M	Dectris EIGER 16M
Wavelength (Å)	0.91946	0.95373	0.91969	0.95373	0.95336	0.95373
Space group	P1 21 1	P1 21 1	C2 2 21	C2 2 21	P 21 21 2	P43 2 2
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	76.303 93.544 60.489 90 90 90	51.814 51.814 81.599 90 90 90
Average mosaicity (°) [b]	0.00	0.08	0.00	0.06	0.06	0.05
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)	47.40-1.88 (1.92-1.88)	36.64-1.28 (1.326-1.28)
Total no. of reflections	1247183 (94443)	295538 (14086)	1426660 (157248)	280003 (13354)	243512 (15537)	381085 (37508)
No. of unique reflections	23868 (1868)	53375 (2645)	9668 (1079)	29849 (1458)	35252 (2205)	29449 (2869)
Completeness (%)	99.8 (98.1)	99.9 (100)	99.9 (99.4)	99.8 (96.8)	97.8 (96.8)	99.98 (99.97)
Multiplicity	52.3 (50.6)	5.5 (5.3)	147.6 (145.7)	9.4 (9.2)	6.9 (7.0)	13 (13)
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)	11.8 (0.9)	23.71 (1.73)
Rmerge	0.157 (0.850)	0.074 (0.990)	0.215 (1.258)	0.072 (0.901)	0.070 (1.429)	0.048 (1.35)
Rmeas [c]	0.158 (0.858)	0.082 (1.010)	0.216 (1.262)	0.076 (0.953)	0.082 (1.668)	0.050 (1.41)
Rpim [d]	0.022 (0.117)	0.034 (0.470)	0.018 (0.102)	0.025 (0.308)	0.042 (0.857)	0.014 (0.39)
$CC_{1/2}$ [b]	0.999 (0.950)	0.999 (0.490)	1.0 (0.976)	0.999 (0.832)	0.999 (0.571)	0.999 (0.894)
Matthews coeff. ($Å^3$ Da^{-1}) [e]	2.60	2.78	2.13	2.13	2.82	2.56
Resolution range (Å)	-	39.35 - 1.65	-	45.27 - 1.68	46.77-1.88	36.64-1.28 (1.31-1.28)
R_{work} (%) [g]	-	16.85	-	16.88	19.32	17.0
R_{free} (%) [h]	-	21.10	-	21.77	22.10	19.33
No. of non-H atoms						
Total	-	3647	-	1882	2998	742
Macromolecules	-	3214	-	1713	2787	653
Ligand	-	20	-	10	20	5
Water	-	413	-	159	191	84
Average B-factor ($Å^2$)	-	26.06	-	28.27	40.73	23.79
RMSD from ideal geometry						
Bond lengths (Å)	-	0.005	-	0.012	0.008	0.011
Bond angles (°)	-	0.74	-	1.18	0.93	1.21
Ramachandran plot, residues in (%)[i]						
Favoured regions	-	97.77	-	95.57	98.58	98.85
Allowed regions	-	2.23	-	4.43	1.42	1.15
Outlier regions	-	0.00	-	0.00	0.00	0.00

[a] The values in parentheses are for the highest-resolution shell.
[b] Calculated with AIMLESS.
[c] Rmeas = $\Sigma_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\ \Sigma i|Ii(hkl)-|/\ \Sigma hkl\Sigma Ii(hkl)$, where $Ii(hkl)$ is the intensity of the ith measurement of an equivalent reflection with indices hkl.
[d] Rpim = $\Sigma_{hkl}\{1/[N(hkl)-1]\}^{1/2}\ \Sigma i|Ii(hkl)-|/\ \Sigma hkl\Sigma Ii(hkl)$.
[e] Calculated with MATTHEWS_COEF within the CCP4 suite.
[f] Generated by Crank pipeline in the CCP4 suite.
[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 (supplementary file)

S3 Table. Amino acid sequence inputs for AlphaFold2

Name	Sequence
SIX1$^{96-284}$	EPFGEESRNDRVTQDMLQALHDLCVERFGTGYRAVSGLCYTDRRATRKIECNKPSVRERDRSVTRACPKG QECTTFNAYNFRNRHHQVTFPVCGPRIEVKDRHDIGIHTEWQGTWYPESPKSPGTYDYFAQMAGTLNGYF GYDGVYSDGYKTSSHGYGHSWSCINCPRGKVTITNTYRATWAFGYTSPH
SIX2$^{98-221}$	GSCFSFPTPARGSCMIDYCWRDDNGVIYSRGITITGSNGASNPTSMRSNDPANLSLNSVFNDGYNGWFPHG HACSNSDTQIYTNHRLLQGVNGVAYVDHVRCENCNFRNVNCLSDVLKNNLIAYSNGVASQSRCT
SIX3$^{36-163}$	LPVEDADSSVGQLQGRGNPYCVFPGRPTSSTSFTTSFSTEPLGYARMLHRDPPYERAGNSGLNHRIYERSRV GGLRTVIDVAPPDGHQAIANYEIEVRRIPVATPNAAGDCFHTARLSTGSRGPATISWDADASYTYYLTISED
SIX4$^{59-242}$	SAHTESVCVHAGTATGADLHWLNAICTGKSTYTVNCAPAGNKNAGSTHTGTCPAGQDCFQLEQVGNFWG DREPDATCSPSNTVFDAVDDKEATHVNGKVVTRAGKPGIGRKLIRLKAQVYRRDGHYGQTSRMGFFRNG KEVYHIDNVASMEPTWNFDPSSDQSFSFFFTPGPNAFRIQGTLNLA
SIX5$^{18-119}$	RDHQYCACQSGSGDSIDIDATTQLQNDNSKSYLWAQTSPAYWFADRHKPGPRFAGIYLKAANGKIDGDTF YNLCINNGGADSTCFDCSKSHQVRNVIYCDAA
SIX6$^{58-225}$	DTLPVSTCPAGQKYDRSVCYKADKIRSFCVANPRSNREKITDTPCQPREICVQRNLSNGKSFAKCIPIVDLVE WKTSANGNKEGCTTTSVNPAGYHHLGTIVYDINKNPIEVDKISYFGEPGNVNEGIGGSTSYFSSDNFQFSKS RYMKTCIFSGGYGNLNAYTWSWE
SIX7$^{49-220}$	EVTFDITQNVNTFTSAASTPWTEGVGLSNIRYQWRAYYSTRQRTTFVEVRVFGTAEAQVVLLPDAPGTSRY RAIDSNVFRPNEEVTGGGLAGWGQVTTVCLQTWGRRGDITYRLRIQS
SIX8$^{50-141}$	DTSGILLACITGAGSAFQAYAGCYLTAFRNDPRTLTLRMDKTRGERISNVLVILSGGALSHAVEEVVQIAPG AVRNLATLGASTVQFLHNFR
SIX9$^{19-114}$	QTTQVGCRALDTKNDGLLTELLLNPSARGAADPDLRYGFWDAKWRKCCNKYKECDKYYTFSYNHPYPW AYRQRRGTIRGQQFDFACVNWRTGACK
SIX10$^{15-149}$	IPDSGVSTGTKDLSKRDDAYIFDVTFRVGPAGANVAPFSGSVYVQDGLTPLVRSGSGSSISDRGYNAFRGIV YFTFTHGYNQYSASTRFGVYVDTGLIVDSNGRPIYGTAPRKACIDYSPHGPTDVCSVTITRSK
SIX11$^{19-110}$	INICCSSFAGHTCTKDQYNNHRQNVILNQIIDKDGVNCVRKGAGPGRWTRKGDWSEWYDCQQWNGPEQH QIEVGECTLFCVTPSGILNRPCI
SIX12$^{27-127}$	SSCLSVGPKGISNQNACVCGGQCVMKDLVVARRKVCCEYTVQIQGGWPVLAQSRCVYGSTGANGGSCSG DNVSLAWWLNYEPEVKSTDPTCIFAKPKLCHS
SIX13$^{78-293}$	QDDEHPNGPCPRGGRLYVDSDEDSSCNAKWGTQTHNDVKTFGSTGSVCAGTFRRITCACCYTMHPITDNN VPRMDGIYCPKWEVCKQEPERWSKWGNRVSHTSCVQAKKLTEILIATKKVVKEYCTPKRWLPSAGKGKN AKFHTWAYNYSTGQLTTLKWMYLKLDGQYVKSAPGISEWGLTYSVNEHNAIELCGYPSDDMQRNSIDAE LQWEATVQ
SIX14$^{18-88}$	QRILGCRMPNGSLNPSPNICNQAGGSFRSGSRGCCTRNTRDGPVVTESRFISGCNKNGGFVSSKEILATSC

PSL1[18-111] (FOXGR_025399)	EDWDRCRCMKYPETGTPNDCATIKACGSGKHRAISISEEKGDIWCEKTDVAISGPEFYRTCYGLLQDPKPNS EADSCCTRWVDGVVVQSDGCFK
FOXG_11033[24-226]	APEAAPGYTTYGDYKGAGENLPSYPSYGSYGAKPKPKPKPAPAPKKYTNYGSYNYKKYSSYGHYKREAEP EAAPEAAPEAAPEAEAAPGYTTYGDYKGAGENLPSYPSYGSYGSKPKPKPKPAPAPKKYTNYGSYNYKKY SSYGHYKREAEPEAAPEAEAEPETYSKYGSYPKKYTHYGSYNYKKYSSYGTYKRAKEFINSLF
FOXG_05755[56-157]	NGVPDPGNFAASCHGLQVISDDLDLTGKPDFERCNDTAYEARQYFSGEYTTVEVRRTDYPDLGKEVQISAT ANYTSTNDNIVNGHLKFGDFQTKFISTPIEG
FOXG_18699[21-96]	CKRTCSASNDAGTTCSYSCTQVCSSISAKQARDTFLAALQSGGNSCSAVGTSGVSCRKTAKFGSCYDHHWS CGSGC
FOXG_04863[106-300]	MYDSSDDKGGLSDITRNAWSKFCNSPYGNNGGVTTRFILDGQWGAVGRLSGWSMRDALIHSMWQTADGI GKKNGYTVYNGCYGFTWQESKPGKANSACGGRSGKACPYNDDCPLAGMECTGLKWGTWMPSIIRMNVY NRDGSLRADAYQARISSQAVGSGGCSKAQTISAYVADFIPIVGPYFATGIRINCLYQS
FOXG_04805[39-132]	QNGQNGGRPVPSGECCVANTSLKQDACTASNGQAGRCVPGGNNCGGRLSCVAQANLQCDANVIERGKDL CRAKAANGLFDGGNIIQNLSQAKVN
FOXG_02829[18-149]	APSSPSDIQARSCVCKKVGDDWICTGTKCYDKVKRDLVPRQCSCHKIGDEWLCGGPKCPRDLPEENKLAK RQCSCHKVGDEWLCGGPKCPRSLPAEESGLEKRQCSCKKVAGEWICSGRKCPRDLSHLMGEE
FOXG_16600[17-164]	SPISKRAVFSQTTYDDLSISGGTAGNAQQEALQKLGGLPTDLSTVEKSDLDFLNSVNQIANDAEDEAFNPAI DAASGEAADALQRGKIKNKVLKLTATILKLQAQQAQGEDVADKLAEENKKLQNNISQDKDEAGKASTFLA FDATTS
FOXG_14684[38-168]	DGTCPRPMCTTPASQGPKDPPACGDSYAACKFDQFPCDEYFSPKVTDTHHCYCILANKKAMDAYCQERGF KSGTNPWKYYYAVECHGAVSNQVCNKDCRDQGRGKGRIDKAHPNGACACDKPNPPYDTCKP
FOXGR_007323[18-86 *]	SLVRRVDVNVPAMTNADGVVVPFDTAGVVQPAKKRDLEQKKRDLAQRKRHISRKRRAVSQEKQKQQQK Q
FOXGR_025639[18-61 *]	APVVRGPGGRLVQEGAGCTLVQGRSVCDDGFGNTFFEDDPFSSK
PSL2[18-106 *] (FOXGR_015322)	EDWDQCRCMKYPSTGTPNDCATIKACGSGKHRAISIYKNGDIWCEKTDVAINGPEFYRTCYDLLQDPKPNS EADSCCIKGDRASDGCFK
FOXGR_015533[19-114 *]	QTCAIAPDPQRNADAFSATSHSGNIDIAFRDHVVFARPSAGTATGVLRLSNGDSYRKIYRIAGPNNVAQFYW LDASSQCKTNLAITQMTNAAWYKE
FOXGR_015522[19-79 *] (SIX15)	TIYCRDVSPPRDTRSWCKTNTPAWQGCQRFCSEHCRSTPRDYPDGCMYHLQVGGDYDCFCK

* Effector candidates identified in the reannotation of the *Fol* genome by Sun et al. (2022) and not predicted in the original genome annotation by Ma et al. (2010).

S4 Table. Gene sequences used in this study.

Name	Sequence
SIX6	TAGGTCTCCAATGGGTCCCTTAGCCCAAACAGAATCCGAGTCGGCAGACGTCGCTGAACATACAATCAATTATATCGACATTGCC CCTGAAGAATTTGAACCGCCCAAAGCTAATTTGTCATCTCTGGTGAGTCGTGACACGCTTCCTGTCAGTACCTGTCCTGCGGGTCA GAAATACGATCGTTCCGTGTGTTACAAGGCAGACAAAATTCGTAGCTTTTGTGTCGCAAACCCTCGTAGCAACCGTGAGAAGATT ACCGACACACCGTGTCAGCCCCGTGAAATCTGTGTGCAACGCAATCTTTCCAACGGCAAGAGTTTCGCTAAGTGTATCCCCATTG TAGACCTGGTGGAATGGAAGACATCCGCAAATGGGAATAAAGAAGGCTGTACTACAACGTCCGTGAATCCGGCTGGGTACCATC ACCTTGGTACTATTGTTTACGATATCAATAAGAATCCTATCGAAGTTGATAAAATCTCGTACTTCGGCGAGCCGGGAAATGTAAA CGAGGGCATTGGTGGCAGCACAAGCTATTTTAGTAGTGACAACTTTCAATTTTCTAAGTCCCGCTACATGAAAACTTGTATTTTCA GTGGTGGGTACGGGAATCTTAACGCCTATACGTGGAGCTGGGAATCTTGGAGACCGT
SIX6-TEV	TAGGTCTCCAATGGGTCCCTTAGCCCAAACAGAATCCGAGTCGGCAGACGTCGCTGAACATACAATCAATTATATCGACATTGCC CCTGAAGAATTTGAACCGCCCAAAGCTAATTTGTCATCTCTGGTGGAAAATCTTTACTTCCAGTCCGACACGCTTCCTGTCAGTAC CTGTCCTGCGGGTCAGAAATACGATCGTTCCGTGTGTTACAAGGCAGACAAAATTCGTAGCTTTTGTGTCGCAAACCCTCGTAGC AACCGTGAGAAGATTACCGACACACCGTGTCAGCCCCGTGAAATCTGTGTGCAACGCAATCTTTCCAACGGCAAGAGTTTCGCTA AGTGTATCCCCATTGTAGACCTGGTGGAATGGAAGACATCCGCAAATGGGAATAAAGAAGGCTGTACTACAACGTCCGTGAATC CGGCTGGGTACCATCACCTTGGTACTATTGTTTACGATATCAATAAGAATCCTATCGAAGTTGATAAAATCTCGTACTTCGGCGA GCCGGGAAATGTAAACGAGGGCATTGGTGGCAGCACAAGCTATTTTAGTAGTGACAACTTTCAATTTTCTAAGTCCCGCTACATG AAAACTTGTATTTTCAGTGGTGGGTACGGGAATCTTAACGCCTATACGTGGAGCTGGGAATCTTGGAGACCGT
Avr1Thrombin	TAGGTCTCCAATGTTGCCTAAAGGAGAGGAGGGTGACATTATTGGTACTTTCAATTTCTCGTCCAGCGACAGCCAACCCCTTAAA ATCCACTGGGTCGATACGCCGGACTCATCTGGGAGCAATCTTGTTCCCCGTTCCGCTCACACGGAGAGTGTATGCGTTCACGCCG GGACCGCTACAGGTGCTGATCTGCATTGGTTGAATGCGATCTGCACCGGGAAGTCTACATACACAGTGAATTGCGCCCCGGCAGG CAACAAGAATGCTGGGTCTACGCACACAGGAACATGTCCGGCAGGTCAGGACTGTTTCCAATTAGAGCAGGTCGGAAACTTTTG GGGGGACCGTGAGCCAGATGCTACCTGTAGCCCGTCCAATACGGTATTTGACGCCGTAGATGACAAGGAAGCTACGCATGTAAA CGGCAAAGTTGTTACACGCGCGGGGAAGCCGGGCATTGGGCGCAAGCTTATTCGTCTTAAGGCTCAGGTCTATCGTCGTGATGGT CACTATGGTCAGACCTCGCGCATGGGATTCTTTCGTAACGGCAAAGAGGTTTACCATATCGACAACGTTGCCTCGATGGAACCCA CTTGGAATTTTGACCCATCGAGTGACCAATCCTTTAGCTTCTTTTTCACACCGGGACCCAACGCTTTCCGTATTCAAGGAACGCTT AATCTGGCCTCTTGGAGACCGT
SIX8Thrombin	TAGGTCTCCAATGACCCCGATTGACAAATCGTTAGATCAAGCGGCAACTATCGAAGAAACTGTCCACCAACCTCACTCCCATGAT GAGCGTGCCTTAGTTCCACGTGGCAGCGATACGAGTGGGATTTTGCTGGCGTGTATTACCGGCGCAGGATCTGCGTTTCAGGCGT ACGCTGGATGCTACTTAACAGCTTTCCGTAATGACCCTCGCACTTTAACTTTGCGTATGGATAAAACCCGTGGAGAACGTATTTCC AATGTTCTGGTTATCTTGTCAGGGGGTGCATTGAGTCACGCCGTGGAAGAAGTAGTACAGATTGCGCCTGGAGCGGTCCGCAATT TGGCAACATTAGGAGCTTCGACTGTCCAATTCCTTCATAATTTTCGTTCTTGGAGACCGT
PSL1	TAGGTCTCCAATGGAAGATTGGGATCGTTGTCGTTGCATGAAATACCCCGAAACTGGGACGCCGAATGACTGTGCCACCATTAAA GCGTGCGGTTCTGGCAAGCACCGTGCCATCTCTATTTCGGAGGAGAAAGGAGACATTTGGTGCGAAAAGACTGATGTAGCAATC AGTGGACCGGAGTTCTACCGCACTTGCTATGGGTTGTTGCAGGACCCGAAGCCTAACTCGGAGGCAGATTCTTGTTGTACTCGTT GGGTAGACGGCGTTGTGGTACAATCTGACGGGTGCTTCAAATCTTGGAGACCGT
SIX13	TAGGTCTCCAATGGAACTTGAAGTTTCCGATTTAAGTGATCAACCGCCGTCAGTGGAGAATACTTACCGCGACCAAGCGTTCAAT GAGGAGGAGTTGTTAAAGGTCGTGGACGAATTATCCGTGGAGCGTACCGAACACACCGAACGCGCGCTTGTGAGTGAAGCGGCA GTCCAAAAGCGCCAGGACGACGAACATCCAAACGGTCCTTGTCCTCGCGGAGGACGTTTGTACGTGGATTCTGACGAAGACTCTT CTTGTAATGCGAAGTGGGGAACACAGACACATAACGATGTTAAGACTTTTGGCTCCACCGGATCAGTCTGTGCCGGTACTTTTCG CCGCATTACGTGTGCTTGTTGTTACACGATGCATCCTATCACGGACAACAACGTTCCACGCATGGATGGTATTTACTGCCCTAAGT GGGAGGTGTGCAAACAAGAGCCTGAACGCTGGTCTAAGTGGGGCAATCGTGTAAGCCACACATCTTGCGTACAGGCGAAGAAGC TGACCGAAATCCTGATCGCTACTAAGAAAGTGGTGAAGGAATACTGCACCCCAAAACGTTGGTTACCATCTGCGGGAAAAGGCA AGAACGCGAAATTCCACACGTGGGCGTACAATTATTCGACGGGTCAGCTTACAACCTTGAAGTGGATGTATTTGAAGTTGGACGG TCAGTATGTCAAGAGTGCGCCTGGGATCAGCGAGTGGGGGCTTACCTATTCAGTCAATGAACATAACGCGATTGAGTTATGTGGT TACCCAAGCGATGATATGCAACGCAATTCGATCGATGCTGAGCTTCAATGGGAGGCTACCGTACAGTCTTGGAGACCGT
SIX8_C58S Thrombin	TAGGTCTCCAATGACCCCGATTGACAAATCGTTAGATCAAGCGGCAACTATCGAAGAAACTGTCCACCAACCTCACTCCCATGAT GAGCGTGCCTTAGTTCCACGTGGCAGCGATACGAGTGGGATTTTGCTGGCGTCTATTACCGGCGCAGGATCTGCGTTTCAGGCGT ACGCTGGATGCTACTTAACAGCTTTCCGTAATGACCCTCGCACTTTAACTTTGCGTATGGATAAAACCCGTGGAGAACGTATTTCC AATGTTCTGGTTATCTTGTCAGGGGGTGCATTGAGTCACGCCGTGGAAGAAGTAGTACAGATTGCGCCTGGAGCGGTCCGCAATT TGGCAACATTAGGAGCTTCGACTGTCCAATTCCTTCATAATTTTCGTTCTTGGAGACCGT
PSL1_C37S	TAGGTCTCCAATGGAAGATTGGGATCGTTGTCGTTGCATGAAATACCCCGAAACTGGGACGCCGAATGACTCTGCCACCATTAAA GCGTGCGGTTCTGGCAAGCACCGTGCCATCTCTATTTCGGAGGAGAAAGGAGACATTTGGTGCGAAAAGACTGATGTAGCAATC AGTGGACCGGAGTTCTACCGCACTTGCTATGGGTTGTTGCAGGACCCGAAGCCTAACTCGGAGGCAGATTCTTGTTGTACTCGTT GGGTAGACGGCGTTGTGGTACAATCTGACGGGTGCTTCAAATCTTGGAGACCGT

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

S5 Table. Primers used in this study

Name	Sequence
Avr1_Fw	TAGGTCTCCAATGCTTCCAAAGGGGGAGGAGGGTG
Avr1_Rv	ACGGTCTCCAAGAAGCTAAGTTAAGTGTACCTTGAATGCGA
Avr3_Fw	TAGGTCTCCAATGCAAGAGGCTGCGGTTCGGGA
Avr3_Rv	ACGGTCTCCAAGAGGCGTTTGGATATACCAGCCCACAC

S5 Table. Amino acid sequence inputs for AlphaFold2

Name	Sequence
SIX1$^{96-284}$	EPFGEESRNDRVTQDMLQALHDLCVERFGTGYRAVSGLCYTDRRATRKIECNKPSVRERDRSVTRACPKG QECTTFNAYNFRNRHHQVTFPVCGPRIEVKDRHDIGIHTEWQGTWYPESPKSPGTYDYFAQMAGTLNGYF GYDGVYSDGYKTSSHGYGHSWSCINCPRGKVTITNTYRATWAFGYTSPH
SIX2$^{98-221}$	GSCFSFPTPARGSCMIDYCWRDDNGVIYSRGITITGSNGASNPTSMRSNDPANLSLNSVFNDGYNGWFPHG HACSNSDTQIYTNHRLLQGVNGVAYVDHVRCENCNFRNVNCLSDVLKNNLIAYSNGVASQSRCT
SIX3$^{36-163}$	LPVEDADSSVGQLQGRGNPYCVFPGRPTSSTSFTTSFSTEPLGYARMLHRDPPYERAGNSGLNHRIYERSRV GGLRTVIDVAPPDGHQAIANYEIEVRRIPVATPNAAGDCFHTARLSTGSRGPATISWDADASYTYYLTISED
SIX4$^{59-242}$	SAHTESVCVHAGTATGADLHWLNAICTGKSTYTVNCAPAGNKNAGSTHTGTCPAGQDCFQLEQVGNFWG DREPDATCSPSNTVFDAVDDKEATHVNGKVVTRAGKPGIGRKLIRLKAQVYRRDGHYGQTSRMGFFRNG KEVYHIDNVASMEPTWNFDPSSDQSFSFFFTPGPNAFRIQGTLNLA
SIX5$^{18-119}$	RDHQYCACQSGSGDSIDIDATTQLQNDNSKSYLWAQTSPAYWFADRHKPGPRFAGIYLKAANGKIDGDTF YNLCINNGGADSTCFDCSKSHQVRNVIYCDAA
SIX6$^{58-225}$	DTLPVSTCPAGQKYDRSVCYKADKIRSFCVANPRSNREKITDTPCQPREICVQRNLSNGKSFAKCIPIVDLVE WKTSANGNKEGCTTTSVNPAGYHHLGTIVYDINKNPIEVDKISYFGEPGNVNEGIGGSTSYFSSDNFQFSKS RYMKTCIFSGGYGNLNAYTWSWE
SIX7$^{49-220}$	EVTFDITQNVNTFTSAASTPWTEGVGLSNIRYQWRAYYSTRQRTTFVEVRVFGTAEAQVVLLPDAPGTSRY RAIDSNVFRPNEEVTGGGLAGWGQVTTVCLQTWGRRGDITYRLRIQS
SIX8$^{50-141}$	DTSGILLACITGAGSAFQAYAGCYLTAFRNDPRTLTLRMDKTRGERISNVLVILSGGALSHAVEEVVQIAPG AVRNLATLGASTVQFLHNFR
SIX9$^{19-114}$	QTTQVGCRALDTKNDGLLTELLLNPSARGAADPDLRYGFWDAKWRKCCNKYKECDKYYTFSYNHPYPW AYRQRRGTIRGQQFDFACVNWRTGACK
SIX10$^{15-149}$	IPDSGVSTGTKDLSKRDDAYIFDVTFRVGPAGANVAPFSGSVYVQDGLTPLVRSGSGSSISDRGYNAFRGIV YFTFTHGYNQYSASTRFGVYVDTGLIVDSNGRPIYGTAPRKACIDYSPHGPTDVCSVTITRSK
SIX11$^{19-110}$	INICCSSFAGHTCTKDQYNNHRQNVILNQIIDKDGVNCVRKGAGPGRWTRKGDWSEWYDCQQWNGPEQH QIEVGECTLFCVTPSGILNRPCI
SIX12$^{27-127}$	SSCLSVGPKGISNQNACVCGGQCVMKDLVVARRKVCCEYTVQIQGGWPVLAQSRCVYGSTGANGGSCSG DNVSLAWWLNYEPEVKSTDPTCIFAKPKLCHS
SIX13$^{78-293}$	QDDEHPNGPCPRGGRLYVDSDEDSSCNAKWGTQTHNDVKTFGSTGSVCAGTFRRITCACCYTMHPITDNN VPRMDGIYCPKWEVCKQEPERWSKWGNRVSHTSCVQAKKLTEILIATKKVVKEYCTPKRWLPSAGKGKN AKFHTWAYNYSTGQLTTLKWMYLKLDGQYVKSAPGISEWGLTYSVNEHNAIELCGYPSDDMQRNSIDAE LQWEATVQ
SIX14$^{18-88}$	QRILGCRMPNGSLNPSPNICNQAGGSFRSGSRGCCTRNTRDGPVVTESRFISGCNKNGGFVSSKEILATSC

PSL1[18-111] (FOXGR_025399)	EDWDRCRCMKYPETGTPNDCATIKACGSGKHRAISISEEKGDIWCEKTDVAISGPEFYRTCYGLLQDPKPNS EADSCCTRWVDGVVVQSDGCFK
FOXG_11033[24-226]	APEAAPGYTTYGDYKGAGENLPSYPSYGSYGAKPKPKPKPAPAPKKYTNYGSYNYKKYSSYGHYKREAEP EAAPEAAPEAAPEAEAAPGYTTYGDYKGAGENLPSYPSYGSYGSKPKPKPKPAPAPKKYTNYGSYNYKKY SSYGHYKREAEPEAAPEAEAEPETYSKYGSYPKKYTHYGSYNYKKYSSYGTYKRAKEFINSLF
FOXG_05755[56-157]	NGVPDPGNFAASCHGLQVISDDLDLTGKPDFERCNDTAYEARQYFSGEYTTVEVRRTDYPDLGKEVQISAT ANYTSTNDNIVNGHLKFGDFQTKFISTPIEG
FOXG_18699[21-96]	CKRTCSASNDAGTTCSYSCTQVCSSISAKQARDTFLAALQSGGNSCSAVGTSGVSCRKTAKFGSCYDHHWS CGSGC
FOXG_04863[106-300]	MYDSSDDKGGLSDITRNAWSKFCNSPYGNNGGVTTRFILDGQWGAVGRLSGWSMRDALIHSMWQTADGI GKKNGYTVYNGCYGFTWQESKPGKANSACGGRSGKACPYNDDCPLAGMECTGLKWGTWMPSIIRMNVY NRDGSLRADAYQARISSQAVGSGGCSKAQTISAYVADFIPIVGPYFATGIRINCLYQS
FOXG_04805[39-132]	QNGQNGGRPVPSGECCVANTSLKQDACTASNGQAGRCVPGGNNCGGRLSCVAQANLQCDANVIERGKDL CRAKAANGLFDGGNIIQNLSQAKVN
FOXG_02829[18-149]	APSSPSDIQARSCVCKKVGDDWICTGTKCYDKVKRDLVPRQCSCHKIGDEWLCGGPKCPRDLPEENKLAK RQCSCHKVGDEWLCGGPKCPRSLPAEESGLEKRQCSCKKVAGEWICSGRKCPRDLSHLMGEE
FOXG_16600[17-164]	SPISKRAVFSQTTYDDLSISGGTAGNAQQEALQKLGGLPTDLSTVEKSDLDFLNSVNQIANDAEDEAFNPAI DAASGEAADALQRGKIKNKVLKLTATILKLQAQQAQGEDVADKLAEENKKLQNNISQDKDEAGKASTFLA FDATTS
FOXG_14684[38-168]	DGTCPRPMCTTPASQGPKDPPACGDSYAACKFDQFPCDEYFSPKVTDTHHCYCILANKKAMDAYCQERGF KSGTNPWKYYYAVECHGAVSNQVCNKDCRDQGRGKGRIDKAHPNGACACDKPNPPYDTCKP
FOXGR_007323[18-86 *]	SLVRRVDVNVPAMTNADGVVVPFDTAGVVQPAKKRDLEQKKRDLAQRKRHISRKRRAVSQEKQKQQQK Q
FOXGR_025639[18-61 *]	APVVRGPGGRLVQEGAGCTLVQGRSVCDDGFGNTFFEDDPFSSK
PSL2[18-106 *] (FOXGR_015322)	EDWDQCRCMKYPSTGTPNDCATIKACGSGKHRAISIYKNGDIWCEKTDVAINGPEFYRTCYDLLQDPKPNS EADSCCIKGDRASDGCFK
FOXGR_015533[19-114 *]	QTCAIAPDPQRNADAFSATSHSGNIDIAFRDHVVFARPSAGTATGVLRLSNGDSYRKIYRIAGPNNVAQFYW LDASSQCKTNLAITQMTNAAWYKE
FOXGR_015522[19-79 *] (SIX15)	TIYCRDVSPPRDTRSWCKTNTPAWQGCQRFCSEHCRSTPRDYPDGCMYHLQVGGDYDCFCK

* Effector candidates identified in the reannotation of the *Fol* genome by Sun et al. (2022) 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. 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.
31. Lazar N, Mesarich CH, Petit-Houdenot Y, Talbi N, Li de la Sierra-Gallay I, Zelie E, et al. A new family of structurally conserved fungal effectors displays epistatic interactions with plant resistance proteins. PLoS Pathog. 2022;18(7):e1010664.

32. 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.
33. Outram MA, Solomon PS, Williams SJ. Pro-domain processing of fungal effector proteins from plant pathogens. PLoS Pathog. 2021;17(10):e1010000.
34. 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.
35. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010;38(Web Server issue):W545-549.
36. 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.
37. Batson AM, Fokkens L, Rep M, du Toit LJ. Putative Effector Genes Distinguish Two Pathogenicity Groups of Fusarium oxysporum f. sp. spinaciae. Mol Plant Microbe Interact. 2021;34(2):141-156.
38. Czislowski E, Fraser-Smith S, Zander M, O'Neill WT, Meldrum RA, Tran-Nguyen LTT, et al. Investigation of the diversity of effector genes in the banana pathogen, Fusarium oxysporum f. sp. cubense, reveals evidence of horizontal gene transfer. Mol Plant Pathol. 2018;19(5):1155-1171.
39. Lievens B, Houterman PM, Rep M. Effector gene screening allows unambiguous identification of Fusarium oxysporum f. sp. lycopersici races and discrimination from other formae speciales. FEMS Microbiol Lett. 2009;300(2):201-215.
40. van Dam P, Fokkens L, Schmidt SM, Linmans JH, Kistler HC, Ma LJ, et al. Effector profiles distinguish formae speciales of Fusarium oxysporum. Environ Microbiol. 2016;18(11):4087-4102.
41. Rocafort M, Bowen JK, Hassing B, Cox MP, McGreal B, de la Rosa S, et al. The Venturia inaequalis effector repertoire is expressed in waves, and is dominated by expanded families with predicted structural similarity to avirulence proteins from other fungi. bioRxiv. 2022:2022.2003.2022.482717.
42. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50(D1):D439-D444.
43. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Gilchrist CLM, Söding J, et al. Foldseek: fast and accurate protein structure search. bioRxiv. 2022:2022.2002.2007.479398.
44. Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, et al. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat Genet. 2015;47(4):410-415.
45. 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.
46. Allen A, Chatt E, Smith TJ. The atomic structure of the virally encoded antifungal protein, KP6. J Mol Biol. 2013;425(3):609-621.
47. 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.

48. Sun X, Fang X, Wang D, Jones DA, Ma L. Transcriptome Analysis of Fusarium-Tomato Interaction Based on an Updated Genome Annotation of Fusarium oxysporum f. sp. lycopersici Identifies Novel Effector Candidates That Suppress or Induce Cell Death in Nicotiana benthamiana. J Fungi (Basel). 2022;8(7).
49. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv. 2022:2021.2010.2004.463034.
50. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19(6):679-682.
51. Stergiopoulos I, de Wit PJGM. Fungal effector proteins. Annu Rev Phytopathol. 2009;47:233-263.
52. Outram MA, Figueroa M, Sperschneider J, Williams SJ, Dodds PN. Seeing is believing: Exploiting advances in structural biology to understand and engineer plant immunity. Curr Opin Plant Biol. 2022;67:102210.
53. 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.
54. 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.
55. 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.
56. 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.
57. 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.
58. Yan X, Tang B, Ryder LS, MacLean D, Were VM, Eseola AB, et al. The transcriptional landscape of plant infection by the rice blast fungus *Magnaporthe oryzae* reveals distinct families of temporally co-regulated and structurally conserved effectors. bioRxiv. 2022:2022.2007.2018.500532.
59. Seong K, Krasileva KV. Comparative computational structural genomics highlights divergent evolution of fungal effectors. bioRxiv. 2022:2022.2005.2002.490317.
60. Ortiz D, Chen J, Outram MA, Saur IML, Upadhyaya NM, Mago R, et al. The stem rust effector protein AvrSr50 escapes Sr50 recognition by a substitution in a single surface-exposed residue. New Phytol. 2022;234(2):592-606.
61. 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.
62. 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.
63. 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.

64. 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.
65. 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.
66. Bull PC, Cox DW. Wilson disease and Menkes disease: new handles on heavy-metal transport. Trends Genet. 1994;10(7):246-252.
67. 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.
68. 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.
69. 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.
70. 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.
71. 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.
72. 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.
73. 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.
74. 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.
75. Wiedemann C, Bellstedt P, Gorlach M. CAPITO--a web server-based analysis and plotting tool for circular dichroism data. Bioinformatics. 2013;29(14):1750-1757.
76. 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.
77. Kabsch W. *XDS*. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-132.
78. 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.
79. 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.
80. Skubak P, Pannu NS. Automatic protein structure solution from weak X-ray data. Nat Commun. 2013;4:2777.

81. 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.
82. 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.
83. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *COOT*. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.
84. 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.
85. 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.
86. 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.