

RESEARCH ARTICLE

Open Access



# The effector candidate repertoire of the arbuscular mycorrhizal fungus *Rhizophagus clarus*

Kinga Sędziewska Toro\* and Andreas Brachmann

## Abstract

**Background:** Arbuscular mycorrhizal fungi (AMF) form an ecologically important symbiosis with more than two thirds of studied land plants. Recent studies of plant-pathogen interactions showed that effector proteins play a key role in host colonization by controlling the plant immune system. We hypothesise that also for symbiotic-plant interactions the secreted effectome of the fungus is a major component of communication and the conservation level of effector proteins between AMF species may be indicative whether they play a fundamental role.

**Results:** In this study, we used a bioinformatics pipeline to predict and compare the effector candidate repertoire of the two AMF species, *Rhizophagus irregularis* and *Rhizophagus clarus*. Our *in silico* pipeline revealed a list of 220 *R. irregularis* candidate effector genes that create a valuable information source to elucidate the mechanism of plant infection and colonization by fungi during AMF symbiotic interaction. While most of the candidate effectors show no homologies to known domains or proteins, the candidates with homologies point to potential roles in signal transduction, cell wall modification or transcription regulation. A remarkable aspect of our work is presence of a large portion of the effector proteins involved in symbiosis, which are not unique to each fungi or plant species, but shared along the Glomeromycota phylum. For 95 % of *R. irregularis* candidates we found homologs in a *R. clarus* genome draft generated by Illumina high-throughput sequencing. Interestingly, 9 % of the predicted effectors are at least as conserved between the two *Rhizophagus* species as proteins with housekeeping functions (similarity > 90 %). Therefore, we state that this group of highly conserved effector proteins between AMF species may play a fundamental role during fungus-plant interaction.

**Conclusions:** We hypothesise that in symbiotic interactions the secreted effectome of the fungus might be an important component of communication. Identification and functional characterization of the primary AMF effectors that regulate symbiotic development will help in understanding the mechanisms of fungus-plant interaction.

**Keywords:** AMF, Effector, *in silico* pipeline, Plant, *Rhizophagus*, Symbiosis

## Background

Arbuscular mycorrhizal fungi (AMF) live in an obligate symbiosis with the roots of around two-thirds of all studied land plant species, i.e. Angiosperms, Gymnosperms, Pteridophytes and some Bryophytes [1], and in one known case with cyanobacteria [2]. This interaction has a major impact on the entire soil ecosystem and plays a crucial role in agricultural systems by increasing plant tolerance to biotic and abiotic stresses [3]. In the

field, AMF form mycelial networks providing extensive connections between roots of different plant species and mediating interactions among plants. These networks significantly improve the rhizosphere soil stability [4]. The major benefit for plants upon mycorrhization is a more efficient nutrient uptake from the soil [5–8]. In exchange, AMF obtain photoassimilates from their hosts [9].

All AMF belong to the phylum Glomeromycota which has been divided into four orders: the *Glomerales*, the *Diversisporales*, and the two ‘ancestral’ lineages, *Paraglomerales* and *Archaeosporales* [10]. They may have started to diverge over 600 Mya. The earliest spore fossils have been found together with the first land plants (455–460

\* Correspondence: k.sedziewska@lmu.de  
Genetics, Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Straße 2-4, 82152 Planegg-Martinsried, Germany



Mya) and resemble present AMF morphological structures [11]. It has been suggested that AMF played a crucial role for the adaptation of phototrophs to the terrestrial environment [1]. The AM symbiosis persisted morphologically unchanged throughout the complete evolutionary development within the plant phylum from haploid gametophytes to diploid sporophytes [12]. Additionally, one single AMF species can often be used to inoculate dicotyledons, monocotyledons and ferns, and one plant species can be mycorrhized by several AMF species. Therefore, AMF are considered not to be host specific and there is no evidence for evolution of host specificity [12]. This finding is extremely remarkable considering the obligate biotrophic life style of AMF. Thus, the mechanism of plant infection and colonization seems to be ancient and conserved within AMF.

Especially during infection, pathogenic and symbiotic plant-microbe interactions show striking similarities, suggesting commonalities in the underlying regulation. For example, pathogenic and AM fungi develop analogous feeding structures, haustoria and arbuscules, respectively [13]. Transcriptome profiles of plant cells hosting these structures indicate activated auxin signalling and increased plant metabolism in response to both, pathogenic and symbiotic fungi [14]. The transcriptome sequencing project of the model AM fungus, *Rhizophagus irregularis* DAOM197198, showed that only a limited set of cell wall degrading enzymes is expressed during invasive growth, presumably aimed to avoid a major release of polysaccharide fragments, thereby, their detection by the plant immune system [15]. This is analogous to the gene expression patterns in obligate biotrophic pathogens such as the fungus *Blumeria graminis* and the oomycete *Albugo laibachii* [15]. Recent studies of plant-mutualistic ectomycorrhizal fungi interactions point in the same direction: genes encoding effector proteins were shown to play a key role in host colonization by controlling the plant immune system [16].

An important factor in plant-microbe interactions are microbial effector proteins released to alter plant cell structure or function, allowing successful infection by suppressing the host defence response [17]. In plant pathogenic fungi and oomycetes two classes of effectors are known. The first class comprises apoplastic effectors that are secreted into the extracellular space. They exert their function by interfering with the plant defence in the apoplast and contain small cysteine-rich proteins (SCRs) and inhibitors of plant extracellular hydrolases including proteases (e.g. *Cladosporium fulvum* avr2, avr9) [18, 19], glucanases and plant cell wall degrading enzymes [17]. The second class encompasses intracellular effectors that are translocated into the host cell and target different subcellular compartments. These effectors may be recognized inside the plant cell by intracellular NB-

LRR (nucleotide-binding site leucine-rich repeat) proteins encoded by *R* resistance genes, which results in the induction of effector-triggered immunity (ETI) [20].

Most of the identified eukaryotic pathogenic effectors do not contain domains or homologies to proteins with known function; therefore their roles remain unclear. Although in oomycete pathogens the RXLR motif has been implicated in translocation of effectors into the host cell, to date no universal host-transportation motif has been identified for fungi [21, 22]. Fungal and oomycete effector genes seem to be at the forefront of the arms race between host and pathogen: their high rate of nonsynonymous sequence substitutions is a strong signature of positive selection. This leads to a high amino acid polymorphism, so that, from the 536 predicted RXLR genes in the oomycete *Phytophthora infestans*, only 16 belong to the “core orthologs” present in other *Phytophthora* species [22].

In contrast to pathogenic systems, there is a vast gap of knowledge on effector proteins in symbiotic fungal-plant interactions [23]. There are indications that also here they may play a role in modulating the plant immune response and consequently the efficiency of fungal infection and colonization of plant cells [16, 24]. Klopffholz and colleagues [24] discovered in the AM fungus *R. irregularis* the secreted effector protein, SP7, influencing fungal development within plant roots. SP7 interacts with the pathogenesis-related transcription factor ERF19 in the *Medicago truncatula* nucleus, and its constitutive expression in roots led to increased mycorrhization and reduced expression of plant defence genes.

Since infection and colonization processes appear to be conserved within AMF, we would also expect effectors playing a role during AM symbiosis to be conserved among AMF species. The level of conservation could then be indicative of whether they play a major role or have instead a rather supplemental function during the symbiosis establishment. Therefore, this scenario would be the exact opposite situation as in pathogenic plant-microbe interactions.

In this study, we predicted and compared the effector candidate repertoire of the two AMF, *R. irregularis* and *Rhizophagus clarus*. To this end, we have used the published *R. irregularis* genome [25] and a *R. clarus* genome draft generated by Illumina high-throughput sequencing. For most of the predicted effector candidates, we were able to identify homologs in both *Rhizophagus* species. In addition, for several of them we could postulate potential roles in symbiosis by means of gene ontology.

## Methods

### *R. clarus* monoxenic culture

*R. clarus* in vitro culture was established as described by Declerck et al. [2]. Chicory Ri-T DNA transformed root

organ culture (*Cichorium intybus*, Munich, Germany) was cultivated with a single-spore isolate of *R. clarus* MUCL46238 on one side of two-compartment Petri dishes containing MRS medium with 0.5 % phytigel (Sigma-Aldrich). Fungal hyphae, but not chicory roots, were allowed to grow over the plastic barrier to the second compartment. Plates were incubated in the dark at 26 °C for up to six months until the development of secondary spores.

Extraradical mycelia and secondary spores were recovered from the root-free plate compartments by dissolving the medium in 10 mM citrate buffer [2]. The fungal material was collected under sterile conditions [26] and stored at −20 °C or −80 °C until used.

### *R. clarus* genomic DNA sequencing

Genomic DNA was extracted from 100 mg spores and hyphae using the DNeasy Plant Mini Kit (Qiagen) following manufacturer's instructions. The genomic sequencing library was constructed with the Nextera DNA Sample Preparation Kit (Illumina), which allows for low gDNA input (50 ng) and does not need prior DNA shearing. The library was quality controlled by analysis on an Agilent 2000 Bioanalyzer with Agilent High Sensitivity DNA Kit (Agilent Technologies) for fragment sizes of around 200–400 bp. Sequencing on a MiSeq sequencer (Illumina) (2x150 bp paired-end sequencing) was performed in the Genomics Service Unit (LMU Biocenter, Martinsried, Germany), yielding about 72 Mio reads and 8 Gb of primary sequence. An initial draft with approx. 53-fold coverage of the *R. clarus* genome was assembled using CLC Genomics Workbench v6.0 (Qiagen) with the following parameters: word size, 24; bubble size, 50; mismatch cost, 2; insertion cost, 3; deletion cost, 3; length fraction, 0.8; similarity fraction, 0.8; mapping mode, map reads back to contigs; minimum contig length, 1,000; mismatch cost, 2. Contaminating sequences were identified by analysis with the metagenomics server, MG-RAST [27]. 80 % of the contigs passed quality control (QC) and from these, around 25 % were recognized as bacterial contamination based on sequences available in databases. Contaminating sequences were removed by mapping against genomes of *Bacillus cereus* (NC004722.1), *Bradyrhizobium elkanii* (GCA000472865.1), *Bradyrhizobium japonicum* (BA000040.2), *Legionella pneumophila* (NC006368.1), and *Sorangium cellulosum* (AM746676.1), which were closely related to the main bacterial contaminants. In addition, all contigs with a GC content > 50 % were eliminated. 84 % of the raw sequence information could be remapped onto the final assembly. The draft assembly statistics are summarized in Additional file 1: Table S1. Completeness of the draft genome assembly was assessed using the BUSCO software tool [28] for genomes with the fungal gene set and with default parameters (Additional file 1: Table S1).

### Bioinformatic prediction of the *R. irregularis* effectome

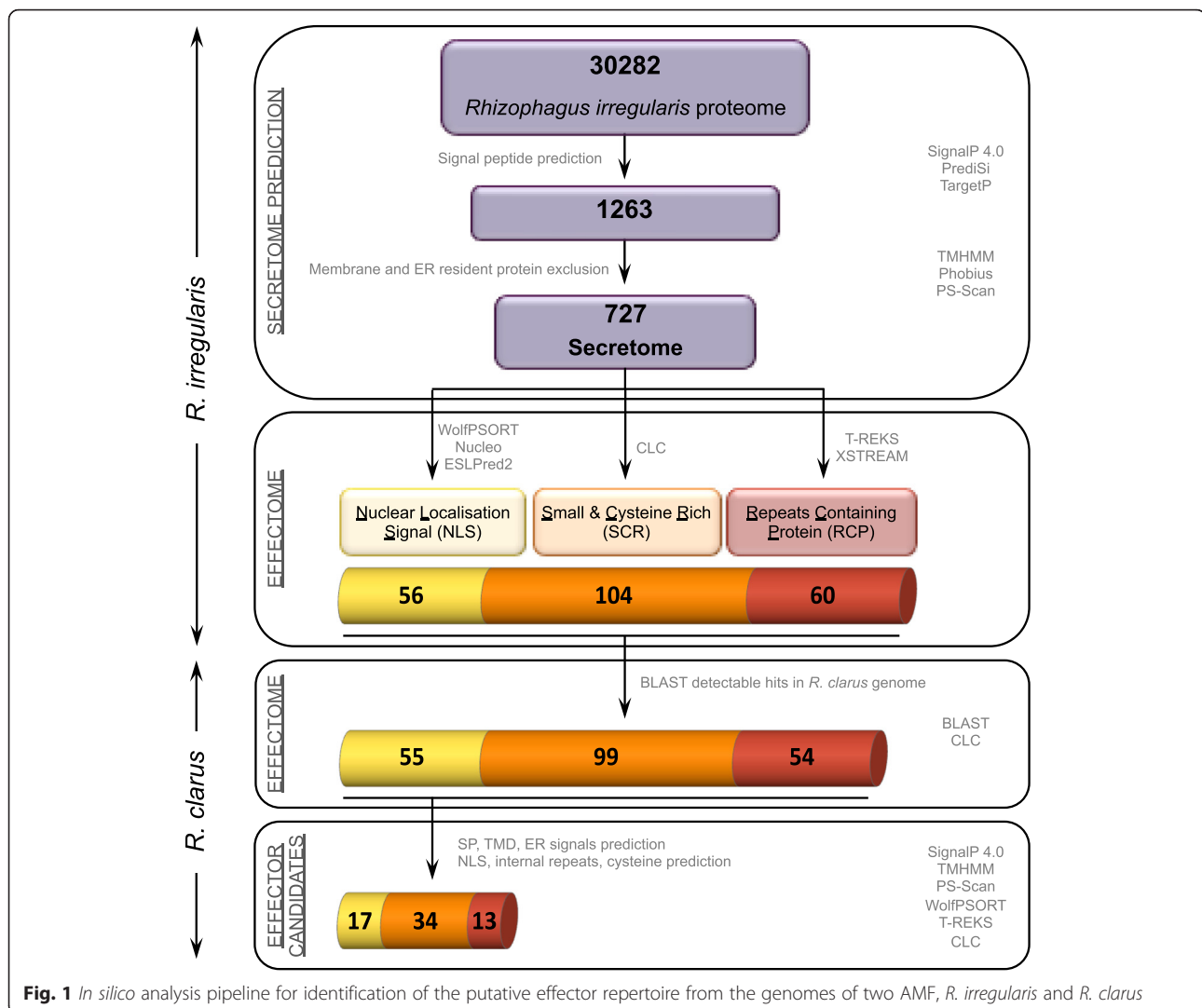
The *R. irregularis* DAOM197198 secretome was predicted from the initial putative proteome (30,282 proteins) available at the DOE Joint Genome Institute (JGI, France) database by use of a series of domain and protein structure prediction programs. The 1,263 proteins carrying N-terminal signal peptides (SPs) were predicted using the well validated SignalP 4.0 and PrediSi programs and sub-cellular targeting using TargetP [29] with default parameters. To ensure stringent standards, only proteins predicted being secreted by at least two methods were considered further. This prediction was followed by exclusion of proteins likely to be retained into the plasma membrane in which a transmembrane domain (TMD) was forecasted by TMHMM 2.0c and Phobius [29, 30] with default parameters. Proteins in which predicted TMDs overlapped with signal peptides were not excluded. A PS-Scan analysis with the Prosite motif 'PS00014' (KDEL) filtered proteins that are likely to permanently reside in the lumen of endoplasmic reticulum [31]. The final number of proteins belonging to the *R. irregularis* secretome was 727 (Fig. 1).

Effector candidates were identified from the *R. irregularis* secreted proteins that fulfil at least one of the previously mentioned criteria:

- (i) Proteins with nuclear localization signal (NLS). To find candidate effectors that target the plant nucleus, the *R. irregularis* secretome was searched for proteins containing a NLS using WoLF PSORT, Nucleo and ESLPred2 (Fungi) software with default parameters [32–34]. We identified 56 predicted proteins containing at least one NLS (Fig. 1).
- (ii) Small and cysteine rich (SCR) proteins. Several known effectors from fungal and oomycetal plant pathogen are characterized by a length of less than 150 amino acids and a cysteine content of more than 3 % [17, 35]. In order to identify candidates achieving this criterion the cysteine content and length of *R. irregularis* secreted proteins were calculated using CLC Genomics Workbench. Within the secretome, we found a total of 104 proteins fulfilling this condition (Fig. 1).
- (iii) Repeats containing proteins (RCP). Internal repeats are often found in oomycetal and fungal effectors [36]. Using the T-REKS algorithm and XSTREAM software with default parameters [37, 38] tandem repeat proteins were selected within the *R. irregularis* secretome. Sixty secreted proteins showed internal repeats (Fig. 1).

### *In silico* search for effector candidate homologs in *R. clarus*

Proteins of *R. irregularis* matching at least one effector criterion were subsequently used to predict candidate effectors in *R. clarus* (Fig. 1). Approx. 95 % (208) of *R.*



**Fig. 1** *In silico* analysis pipeline for identification of the putative effector repertoire from the genomes of two AMF, *R. irregularis* and *R. clarus*

*irregularis* effector candidates showed BLASTP (bl2seq) detectable hits with minimum similarity of 40 % within the *R. clarus* initial genome draft. Even though these proteins share sequence similarity they do not necessarily have a similar topology. Therefore, we applied the same selection criteria that were used on the *R. irregularis* secretome, filtering out proteins with TMDs and endoplasmic reticulum retention signals. Also, an additional control for the presence of SP, NLS and internal repeats was performed.

#### Effector protein characterization by BLAST2GO

To characterize predicted effector proteins the BLAST2GO software was used [39]. At first, protein sequences encoded by effector candidate genes were compared with the non-redundant sequence database. Analysis were performed using the BLASTP algorithm with expect-value lower than  $1.0 \times 10^{-5}$  recording max. 20 hits. Proteins with significant hits were classified into Gene Ontology

categories by 'Mapping'. Data from both analyses were merged into the annotation. Additionally, presence of conserved protein domains or motifs was determined using Pfam, Superfamily 1.73 software and NCBI Conserved Domain search with default parameters [40, 41].

## Results and Discussion

### Defining the *Rhizophagus* effector repertoire

To identify and classify effector candidates of the *Rhizophagus* genus we constructed a bioinformatics pipeline using current knowledge of the properties of validated fungal and oomycete plant pathogen effectors (Fig. 1). It was observed that known pathogen effector proteins from fungi and oomycetes are secreted and fulfil at least one of the following criteria: (i) they contain a nuclear localization signal (NLS), (ii) they are small and cysteine rich (SCR), (iii) they contain internal repeats (RCP), or (iv) they show similarity to haustorial (arbuscular) expressed proteins [17, 36]. For *R. irregularis* so far no

inventory of arbuscular expressed proteins exist. Therefore, we restricted the effector candidate identification to the first three criteria.

Effector candidates need to carry an N-terminal signal peptide (SP) for secretion and should not be retained in the membrane or endoplasmic reticulum of the secreting organism. The *R. irregularis* DAOM197198 secretome was predicted from the initial putative proteome (30,282 proteins) available at the DOE Joint Genome Institute (JGI, USA) database by use of a series of domain and protein structure prediction programs. The predicted *R. irregularis* secretome contains 727 proteins (Fig. 1). The final number of *R. irregularis* predicted secreted proteins that fulfil at least one of the above-mentioned effector criteria was 220, representing 23 % of all *R. irregularis* putatively secreted proteins (Additional file 1: Table S2). In addition, we found 21 proteins fulfilling two but none that matched all three of the effector criteria.

Brief Gene Ontology category analyses by ProtFun 2.2 server [42] of both, secretome and effectome, showed enrichment in hormone, stress response, growth factor, and immune response related proteins (Additional file 1: Figure S1). When comparing secretome and effectome, only three differences are apparent. The 40 % increased number of hormone-related proteins observed in the effectome correlates with the finding that in AM the root architecture is modified by essential, hormone-mediated pathways that might be controlled by AMF [43]. A reduction by 45 % of proteins involved in immune response between effectome and secretome might be indicative of the symbiotic interaction between AMF and plant. The same could be true for the relative reduction of proposed receptor proteins by 60 % in the effectome. Receptors play an important role in all kind of interactions by detecting and recognising signals from pathogenic or symbiotic organisms present in the environment [44].

### The effectome is conserved

Given the central importance of secreted proteins in determining the outcome of plant-microbe encounters in other systems, it is likely that AMF effectors also have a crucial function. However, knowledge about AMF effectors and their function during symbiosis is limited and to date only one effector has been functionally characterised [24].

We wanted to address two questions: (i) How conserved is the effectome between two closely related AMF species? And (ii), does the identity of the conserved putative effectors shed light onto their potential function during the AMF colonization process? For this purpose we generated a draft genome of *R. clarus* MUCL46238 by high-throughput sequencing resulting in 18,021 scaffolds with a total of 94 Mb, in size similar

to the 101 Mb Gloin1 assembly of *R. irregularis* DAOM197198 [25]. An analysis of genome completeness using the BUSCO software tool [28] showed comparable levels for both genomes (Additional file 1: Table S1). Therefore, we estimate the complete *R. clarus* MUCL46238 genome to be slightly smaller than the 150 Mb total genome size predicted for *R. irregularis* DAOM197198 [25, 45, 46].

To determine the level of conservation within the AMF symbiotic effectome, we compared protein sequences encoded by selected housekeeping genes and proteins involved in symbiosis from *R. irregularis* and *R. clarus* and calculated the level of identity and similarity by BLASTP (bl2seq) (Align Sequences Protein BLAST, BLASTP) (Additional file 1: Table S3). The identity of selected proteins varied between 84 % for ammonium transporter 1 to 99 % for the conserved elongation factor 1-alpha, whereas the similarity varied between 91 and 99 %, respectively. The high levels of protein identity and similarity confirm the close relationship between these two species of the same genus, which is in agreement with current knowledge [10].

Subsequently, we searched within the *R. clarus* genome for homologs of the *R. irregularis* proteins matching at least one effector criterion. Approx. 95 % (208) of *R. irregularis* effector candidates showed BLASTP (bl2seq) detectable hits within the *R. clarus* genome draft (Fig. 1). The surprisingly high number of putative effector protein homologs present in both *Rhizophagus* species clearly speaks in favour of a conserved role during the infection process and demonstrates the potential of a comparative approach.

Although the *R. clarus* draft genome is not complete and therefore incomplete ORF information in the homolog sets were to be expected, analysis of the overlapping effectome could still reveal potential roles in symbiosis. For 30 % of the effector candidates predicted in *R. irregularis* we found only truncated homolog sequences in *R. clarus*. This can only be partly explained by genome incompleteness, because the two genome assemblies were of comparable quality (Additional file 1: Table S1). A further 25 % of candidates in *R. clarus* had to be omitted because the inaccurate annotation of the *R. irregularis* proteome resulted in selection of false positive candidates in the effectome, as properly annotated *R. clarus* homologs of these candidates did not fulfil effectome criteria. Additionally, even though *R. clarus* effector protein homologs share sequence similarity with that of *R. irregularis*, they do not necessarily have a similar topology. Therefore, after applying the same selection criteria used on the *R. irregularis* secretome, where proteins predicted to hold SP, NLS, RCP but lacking TMD and ER retention signal were kept, the *R. clarus* candidate number was reduced by 13 %. The final number of full-length sequence candidates fulfilling effector selection criteria present in *R. clarus* as

well as in *R. irregularis* genome was 64, which is approx. 32 % of all candidate homologs present in both species (Fig. 1). The protein similarity between the homologs was 40–98 %.

Within the 64 *R. clarus* effector candidates we identified only five proteins that fulfilled two of the effector criteria and none that matched all three of them (Additional file 1: Table S4). Furthermore, the similarity between all effector candidates from both *Rhizophagus* species was ranged from 51 % for NLS/SCR\_9486 (Additional file 1: Table S4) to 97 % for NLS\_7749 as determined by the BLASTP (bl2seq) algorithm (Table 1). From these, 18 conserved effector candidates show more than 90 % similarity between *R. clarus* and *R. irregularis*, which falls within the same range as the similarity between housekeeping genes (Additional file 1: Table S3). More than 44 % (8) of the conserved candidates belong to the NLS group (Fig. 1, Table 1). In addition, the NLS group shows the highest conservation level (47 % of candidates are conserved) in comparison to the SCR and RCP groups (20 % and 23 % conserved, respectively). These findings suggest that the effectors translocated to the plant nucleus with possible effect on the transcriptional regulation in plant cells might be essential for AM symbiosis. The highly conserved effector candidates identified here between different AMF species are probably playing a fundamental role during fungus-plant interaction also explaining the degree of unspecific interaction with all potential hosts. In contrast, effector candidates showing high diversification may have minor host specific functions.

#### Potential functional features of *R. clarus* effector candidates

To deduct putative functions of *R. clarus* and *R. irregularis* predicted effector candidates the BLAST2GO software with automated annotation was used [39]. The obtained information was used to predict putative cellular functions of *Rhizophagus* effector candidates (Table 1).

Only 14 of *R. clarus* effector candidates (22 %) showed sequence similarity (BLASTP, cut-off  $e$ -value  $> 10^{-5}$ ) to documented proteins with defined or predicted function, and 13 (20 %) contain known domains or motifs (Fig. 2). Predicted domain and homology hits with an  $e$ -value higher than  $10^{-5}$  were ignored [47]. Current research showed that a number of fungal and oomycete effectors despite being under high diversifying or positive selective pressure still carry recognizable Pfam domains, which might be helpful in prediction of their biological function [48, 49]. Thus, for 13 (20 %) of the *R. clarus* effector candidates the putative function could be determined, hence their potential role in AM symbiosis was defined

and subdivided into three main functional groups: signalling, cell wall modification, and transcriptional regulation. However, the majority of *R. clarus* effector candidates are novel sequences showing no significant homology to known proteins in other organisms or sequence motifs (Additional file 1: Table S4), which is in agreement with previous studies highlighting the evolutionary diverse nature of fungal effector groups [36]. Based on computational approaches, we established effector candidate subcategories and speculated on putative functions after careful confirmation of presence of active sites and completeness of the predicted domains.

#### Secreted proteins likely to target plant signalling

Secreted effectors can alter plant defence by interfering with signal transduction [50]. Within the *R. clarus* effectome we detected eight effector candidates that might affect plant signalling.

Within the *R. clarus* effectome, candidate SCR\_348911 shows homology to the *Thermoanaerobacter* sp. peptidoglycan-binding LysM protein and contains a LysM domain (PF01476). Plant receptor kinases containing a LysM (lysine motif) domain are able to detect chitin-like components or chitin fragments [51]. Chitin, an N-acetylglucosamine (GlcNAc) polymer, is a main component of the fungal cell wall [52]. Chitin recognition activates signalling pathways, which lead to plant immune defence [53]. Due to this, it is suggested that many plant pathogens secrete effectors that contain LysM domains in order to inhibit the release of chitin fragments from fungal cell walls or to bind these fragments, thus, preventing their recognition by plant receptors. For instance, the rice blast fungus *Magnaporthe oryzae* secretes the LysM domain containing effector, Slp1, which binds chitin oligosaccharides. This suppresses plant defence response induced through the recognition of chitin fragments by the rice receptor CEBiP (chitin elicitor binding protein) [51, 52]. Candidate SCR\_348911 might play a similar role in AM symbiosis.

The second effector candidate NLS\_30765 contains a F-box-like domain (PF12937). Proteins containing this domain are components of the E3 ubiquitin ligases called SCF complexes (Skp, Cullin, F-box containing complex) which control many intracellular regulatory processes by catalysing the ubiquitination of proteins that targets proteins for degradation [54]. Intense analysis showed that plants defence response pathways might be under control of an SCF complex [54, 55]. Many pathogenic effectors use a strategy of ubiquitination in order to regulate plant immunity. The AvrPiz-t effector of the *M. oryzae* binds to and destabilizes the rice RING E3 ubiquitin ligase APIP6 to suppress chitin-induced plant immunity [56].

The third effector candidate NLS\_334409 shows homology to the *Ochotona princeps* nitrogen permease regulator 3-like protein (NPR3) isoform X1 and contains a NPR3

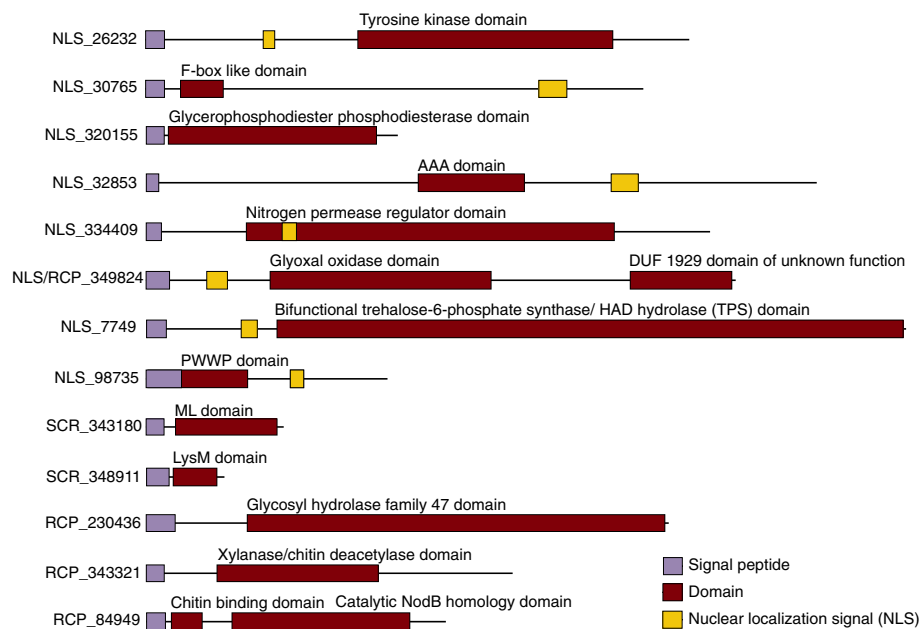
**Table 1** The *R. clarus* effector candidates with predicted function

Predicted effector	Length (aa)	Similarity %	Homology to (e-value/homology length)	Conserved domains	Cellular role	Putative function
Interference with signal transduction						
NLS_26232	592	91	Salt-inducible protein kinase (NP_001105276.1) <i>Zea mays</i> 4.43e <sup>-26</sup> /259)	Protein kinase-like, PKc (7.5e <sup>-34</sup> )	Cell envelope	Interference with signal transduction
NLS_30765	542	92	No significant hits	F-box-like domain (2.8e <sup>-07</sup> )	Cell envelope	Interference with signal transduction
NLS_334409	614	95	Nitrogen permease regulator 3-like protein isoform X1 (XP_004596640.1) <i>Ochotona princeps</i> (9e <sup>-47</sup> /595)	Nitrogen Permease regulator of amino acid transport activity 3 (9e <sup>-88</sup> )	Cell envelope	Interference with signal transduction
NLS_320155	274	89	PLC-like phosphodiesterase (EJF62321.1) <i>Dichomitus squalens</i> (2e <sup>-41</sup> /256)	Glycerophosphodiester phosphodiesterase, GDPD (3.0e <sup>-53</sup> )	Cell envelope	Interference with signal transduction
RCP_340423	175	88	Phospholipase d1 (EFX04791.1) <i>Grosmannia clavigera</i> (7e <sup>-07</sup> /98)		Cell envelope	Interference with signal transduction
SCR_348911	85	53	Peptidoglycan-binding LysM (YP_001663206.1) <i>Thermoanaerobacter sp. X514</i> (2e <sup>-07</sup> /53)	LysM domain (3.5e <sup>-12</sup> )	Cell envelope	Interference with signal transduction
NLS_7749	827	97	Trehalose phosphate synthase (EKD17284.1) <i>Marssonina brunnea</i> (0.0e/790)	Bifunctional trehalose-6-phosphate synthase/HAD hydrolase subfamily IIB domain (0.0e)	Cell envelope	Interference with signal transduction
SCR_343180	150	85	No significant hits	ML (MD-2-related lipid-recognition) (2e <sup>-07</sup> )	Cell envelope	Interference with signal transduction
Cell wall modification						
NLS/RCP_349824	610	83	Galactose oxidase (CCO36399.1) <i>Rhizoctonia solani</i> (1.8e <sup>-70</sup> /549)	Glyoxal oxidase N-terminus domain (1e <sup>-29</sup> )	Cell envelope	Plant cell wall modification
RCP_230436	569	89	class I alpha-mannosidase protein (EOD46381.1) <i>Neofusicoccum parvum</i> (1.8e <sup>-134</sup> /523)	Glycosyl hydrolase family 47 (0.0e)	Cell envelope	Fungal cell wall modification
RCP_343321	399	90	Chitin deacetylase (XP_003194854.1) <i>Cryptococcus gattii</i> (3e <sup>-64</sup> /289)	Chitin deacetylase domain (1.4e <sup>-33</sup> )	Cell envelope	Fungal cell wall modification
RCP_84949	326	70	Bifunctional xylanase/deacetylase (CCO26450) <i>Rhizoctonia solani</i> (4e <sup>-55</sup> /246)	Catalytic NodB homology domain of the carbohydrate esterase 4 superfamily (1.8e <sup>-87</sup> ) Hevein/Chitin binding domain (3.8e <sup>-10</sup> )	Cell envelope	Fungal cell wall modification
Transcription regulation						
NLS_98735	263	95	Hypothetical protein (EIE87396.1) <i>Rhizopus delemar</i> (8e <sup>-13</sup> /105)	PWWP domain (1.5e <sup>-18</sup> )	Transport and binding	Transcription regulation

**Table 1** The *R. clarus* effector candidates with predicted function (Continued)

Unknown function					
NLS_32583	730	91	DENN domain containing protein (EJY85661.1) <i>Oxytricha trifallax</i> ( $1e^{-49}/498$ )	ATPase family associated with various cellular activities (AAA) ( $1.5e^{-05}$ )	Amino acid biosynthesis
NLS_343100	654	96	P-loop containing nucleoside triphosphate hydrolase protein (EIW63580.1) <i>Trametes versicolor</i> ( $5e^{-34}/394$ )		Cell envelope
SCR_339199	125	74	Carbohydrate-binding module family 19 protein (XP_001874952.1) <i>Laccaria bicolor</i> ( $4e^{-13}/88$ )		Transport and binding
SCR_94594	137	79	Proline-rich protein (XP_001875220.1) <i>Laccaria bicolor</i> ( $8e^{-23}/78$ )		Cell envelope
RCP_349288	473	61	Carbohydrate-binding module family 19 protein (XP_001874952.1) <i>Laccaria bicolor</i> ( $9e^{-17}/81$ )		Cell envelope
RCP_335225	347	84	Hypothetical protein (EPZ34955) <i>Rozella allomyces</i> ( $2e^{-72}/315$ )		Cell envelope





**Fig. 2** Illustration of the various functional modules of *R. clarus* effector candidates

domain (PF03666) typical of this protein family. NPR3 forms a heterodimer complex with NPR2 (nitrogen permease regulator 2) protein, which mediates an amino acid starvation signal to TORC1 (target of rapamycin kinase) [57]. TORC1, an essential serine/threonine protein kinase, incorporates signals from diverse pathways in the cell. The homology of candidate NLS\_334409 to NPR3 protein suggests its interference with signal transduction most likely during amino acid deficiency conditions.

A similar role in interfering with plant signalling pathways might have the fourth candidate NLS\_26232 in which a protein kinases catalytic domain, PKc (PF07714) was detected. PKc domains can transfer a phosphate group from ATP to serine/threonine or tyrosine residues on protein substrates. PKs often function as components of signal transduction pathways in which one kinase activates another kinase. This serial action transfers a signal within a cell, which results in cellular responses [58].

Two additional candidates might also influence the plant signalling cascades by being involved in hydrolytic processes. The candidate NLS\_320155 shows homology to the *Dichomitus squalens* PLC-like (phospholipase C-like) phosphodiesterase and has a glycerophosphodiester phosphodiesterase domain, GDPD (PF03009). PLC, by hydrolysing phosphodiester bonds, integrates signalling cascades critical for a variety of cellular and physiological functions [59]. The candidate RCP\_340423 shows homology to the *Grossmannia clavigera* phospholipase d1, which catalyse the cleavage of phospholipids into fatty acids and other lipophilic substances, like

phosphatidic acid, which is an established intracellular signalling lipid [60, 61].

Candidate number seven, NLS\_7749, has homology to the glycosyltransferase family 20 protein from *Colletotrichum higginsianum* and encompasses a complex bifunctional domain, which includes a trehalose-6-phosphate synthase (TPS) domain (PF02358) and a HAD-superfamily hydrolase domain, subfamily IIB (TIGR0148). TPS catalyses the synthesis of  $\alpha,\alpha$ -1,1-trehalose-6-phosphate from glucose-6-phosphate using a UDP-glucose donor and is a key enzyme in the trehalose synthesis pathway [62]. Furthermore, trehalose was reported to have a role as a stress protectant in many organisms across kingdoms [63]. Remarkably, in pathogens trehalose is associated with several mechanisms involved in cell wall integrity, regulation of metabolism, and successful infection [64]. Absence of a trehalose metabolic apparatus is correlated with lower pathogen virulence. However, yet the underlying mechanism remains unravelled. For instance, in *M. oryzae* Tps1 regulates transcriptional effectors linked to virulence factors via NADPH-binding during appressorium-mediated rice infection [65]. Consequently, the putative effector NLS\_7749 might take part in the colonization step of AMF symbiosis by being involved in trehalose synthesis.

The last candidate, SCR\_343180, contains a ML (MD-2-related lipid-recognition, PF02221) domain, which is involved in specific lipid recognition. It is predicted that ML domain containing proteins through interacting with specific lipids can mediate diverse biological functions, like lipid metabolism, host response to pathogen components

like lipopolysaccharides, and other cellular functions [66]. This suggests that the SCR\_343180 candidate, by interacting with lipids, can interfere with the plant lipid signalling; however the direct mechanism of this effect needs to be clarified.

#### Genes encoding putative secreted enzymes engaged in cell wall modification

We identified four *R. clarus* effector candidates encoding putative secreted enzymes involved in cell wall modification. They were divided in two subgroups depending whether their putative biological functions concern plant or fungal cell wall.

**Enzymes involved in degradation of the plant cell wall** Most of the necrotrophic fungi degrade plant tissues by producing a broad set of carbohydrate-active enzymes specifically focused on plant polysaccharide degradation [67]. Transcriptomic profiling of *R. irregularis* DAOM197198 predicted 139 genes encoding carbohydrate-active enzymes [15]. This amount is relatively low in comparison to obligate biotrophic pathogens and ectomycorrhizal symbionts [68]. Also, only one putative enzyme involved in plant cell wall modification was found in the *R. clarus* effector repertoire. This suggests that during AM symbiosis the plant cell wall is not completely degraded, but instead a localized corruption of its structure might occur to make it more flexible, possibly for fungus penetration into the plant cell.

The single effector of this subgroup, candidate NLS/RCP\_349824 shows homology to the *Rhizoctonia solani* galactose oxidase and contains a glyoxal oxidase domain (PF07250). Glyoxal oxidase has a copper active site remarkably similar to that found in another fungal metalloenzyme, galactose oxidase [69]. Glyoxal oxidase is a copper metalloenzyme that catalyses the oxidation of aldehydes to carboxylic acids and is an essential component of the extracellular lignin degradation pathways [70]. Although lignin is present only in small amounts in the plant cell wall [71], it has a strong impact on the rigidity of its structure. Lignin degradation might result in a more flexible cell wall necessary for fungal penetration.

**Enzymes involved in modification of the fungal cell wall** The effectome also contains three putative enzymes involved in chitin degradation. This suggests constant remodelling of the fungal cell wall, likely an adaptation to substantial differences in environmental conditions in the soil and inside of plant tissues. Similarly, in *Cryptococcus neoformans* it was shown that chitosan, the product of chitin deacetylation, helps to sustain cell integrity by maintaining normal capsule width during vegetative growth [72].

The first effector candidate in *R. clarus*, RCP\_230436, contains a glycosyl hydrolase family 47 domain (PF01532) and shows homology to the *Neofusicoccu parvum* class I alpha-mannosidase enzyme. This enzyme is involved in the cleavage of the terminal mannose from Asn-linked oligosaccharides, which takes part at the early step of hydrolysis of N-glycan such as chitin [73].

Two other *R. clarus* effector candidates might function in fungal cell wall remodelling by chitin deacetylation [74]. Candidate RCP\_343321 shows homology to xylase/chitin deacetylase enzyme from *Cryptococcus gatti* and contains a chitin deacetylase domain (PF01522). The other candidate, RCP\_84949, shows a significant homology to *Rhizoctonia solani* bifunctional xylanase/deacetylase. The presence of hevein/chitin binding domain (PF03067) and a catalytic NodB homology domain of the carbohydrate esterase 4 superfamily (PF01074), which typically occur in chitin deacetylases [74], suggest a chitin degradation function of both candidates.

However, candidates with putative chitin deacetylase function might also have another purpose. It was shown that chitin deacetylase plays a significant role in protecting pathogenic fungal hyphae from being lysed by plant secreted chitinases, specifically in the cases of the wheat stem rust fungus *Puccinia graminis* f. sp. *tritici* and the broad bean rust fungus *Uromyces fabae* [74]. Furthermore, it was suggested that exposed fungal chitin polymers are partially de-*N*-acetylated by chitin deacetylases during the infection of the host, this with purpose to avoid plant antimicrobial hydrolases by affinity modulation [75].

#### Genes encoding proteins with putative transcriptional regulatory function

During the pathogenic or symbiotic infection thousands of genes are up- or down-regulated in the plant as revealed by several transcriptome analyses. Expression of genes encoding enzymes involved in protein degradation, cell wall modification, and secondary metabolite biosynthesis is induced in order to activate plant immune response [62]. However, pathogens are able to secrete the so-called transcription activation-like effectors, which target DNA and are involved in transcriptional regulation of a set of plant host genes, for instance to prevent defence response [76].

We found only one putative candidate within the *R. clarus* effectome that might regulate plant gene expression during establishment of the AM symbiosis. This is the nuclear localized (data not shown) candidate, NLS\_98735, which consists of a PWWP domain (PF00855), named after a conserved Pro-Trp-Trp-Pro motif. PWWP domain-containing proteins bind to chromatin and control its packing resulting in transcription and replication regulation. Moreover, PWWP proteins are also involved in DNA

damage response. While DNA damage occurs, they regulate chromatin condensation and ability of DNA repair factors, like 53BPI and CtIP to bind DNA [77]. Similarly, candidate NLS\_98735 might be involved in transcription regulation by chromatin remodelling or condensation.

#### **Effector candidates with unknown function**

Six of the *R. clarus* effector candidates show homologies to known proteins or contain identified domains or motifs. However, the similarity does not allow for prediction of a specific function. One of the candidates, NLS\_32583, has an AAA (PF00004, ATPases Associated with diverse cellular Activities) domain and shows significant homology to a protozoa protein containing a DENN (Differentially Expressed in Neoplastic vs. Normal cells) domain ( $1e^{-49}$ ). However, the homology is mostly based on the AAA domain region as a DENN domain was not detected. The AAA superfamily is a large and functionally diverse superfamily composed of members that are essential for a variety of cellular functions, like protein unfolding and degradation, DNA recombination, replication and repair, and signal transduction [78].

Candidates SCR\_94594 and RCP\_335225 show homologies to putative conserved fungal proteins with unknown function. Both of them lack of any functional domain; therefore, their putative function cannot be predicted. Candidate NLS\_343100 shows homology to the *Trametes versicolor* P-loop protein containing a nucleoside triphosphate hydrolase domain (P-loop NTPase). Nevertheless, this homology is partial and does not include the characteristics for this family of P-loop NTPase domain proteins. In addition, two other candidates, SCR\_339199 and RCP\_349288, show homology to the *Laccari bicolor* carbohydrate binding module family 19 protein (CBM19). Proteins belonging to this family contain a module of 60–70 residues (PF03427) with a chitin binding function [79]. However, once more, this module is absent in both candidates; hence, despite the homology the putative function cannot be predicted.

#### **Conclusions**

Symbiotic interaction between AMF and plant host requires a well-adjusted dialog between the two partners in order to function properly. We hypothesise that one of the key components of this communication is the fungal effectome.

In this study, we used a bioinformatic pipeline to predict effector candidates from two AMF belonging to the *Rhizophagus* genus, *R. irregularis* and *R. clarus*. Our *in silico* pipeline revealed a list of several candidate effector proteins that create a valuable source of information to elucidate the mechanism of plant infection and colonization by AM fungi.

Our *in silico* characterization of the AMF effectome focused on *R. clarus* candidates. The majority of the *R. clarus* effectors are novel proteins with no significant homology to known sequences or motifs in other organisms. This finding is in agreement to previous reports enhancing the evolution diverse nature of fungal and oomycete effectors [34]. However, for a small number of candidates putative functions during plant-fungus symbiosis might be predicted.

The set of putative effector candidates encoded by the *R. clarus* effectome appears to be suited for a fungus that wants to establish a symbiotic rather than a pathogenic relation with the host. The presence of only one candidate encoding a putative plant cell wall degrading enzyme (PCWDE) in the *R. clarus* effectome is in agreement with earlier reports, in which it was shown that several obligate biotrophic pathogens and ectomycorrhizal symbionts have a decreased repertoire of PCWDEs compared to saprophytes [25, 67, 80, 81]. In comparison to pathogenic fungi, AMF seem not to degrade the plant cells but rather grow between them or inside of them while trying to avoid recognition by the plant defence. However, localized plant cell wall modelling probably does occur to weaken and make it more flexible during fungal penetration and passage from cell to cell. The lack of the typical pathogenic effectors, such as extracellular lipases, proteases, nucleases and phytases, additionally confirms the non-pathogenic status of AMF [82].

The relationship between symbiotic fungi and plants requires a major rearrangement of the plant signalling pathways for which fungi have evolved a range of secreted proteins. For instance, fungal effectors control auxin signalling within plant cells to alter plant root architecture and favour colonization [50]. Thus, the presence of candidates in the *R. clarus* effectome that might target plant signalling in order to allow fungal entry was expected. Reprogramming of the plant defence could be used to avoid recognition. Alternatively, effectors might affect transport processes across the plant cell membrane. This was previously shown for symbiotic fungi which encode effectors to control the transfer of glucose into the apoplastic space of the root where it is easily accessible for fungus [50].

A remarkable aspect of our work is presence of a large portion of the effectors that are not species-specific, but conserved between different *Rhizophagus* species. This group of highly conserved effector proteins may play a fundamental role during fungus-plant interaction and has first priority for future functional characterization.

#### **Availability of supporting data**

The sequences of the predicted effector candidates have been deposited in GenBank under accession numbers: KU305736 - KU305799 (as presented in Additional file 1: Table S4). Other supporting data are included in an Additional file 1.

## Additional file

**Additional file 1: Figure S1.** Gene Ontology category analyses by ProtFun 2.2 server [40] of *R. irregularis* DAOM197198 secretome and effectome. **Table S1.** *R. clarus* draft genome assembly statistics in comparison to published *R. irregularis* genome assemblies. **Table S2.** The *R. irregularis* putative effectome. **Table S3.** Identity and similarity levels between *R. irregularis* and *R. clarus* based on sequences of selected proteins (Align Sequences Protein BLAST, BLASTP). **Table S4.** The *R. clarus* effector candidates. (DOCX 152 kb)

### Abbreviations

AAA: ATPases associated with diverse cellular activities; AM: Arbuscular mycorrhizal; AMF: Arbuscular mycorrhizal fungi; CEBIP: Chitin elicitor binding protein; DENN: differentially expressed in neoplastic vs. normal cells; ER: endoplasmic reticulum; ETI: effector-triggered immunity; GDPD: glycerophosphodiester phosphodiesterase domain; GlcNAc: N-acetylglucosamine; LysM: lysine motif; *M. oryzae*: *Magnaporthe oryzae*; Mio: million; ML: MD-2-related lipid-recognition; Mya: million years ago; NB-LRR: nucleotide-binding site leucine-rich repeat; NLS: nuclear localization signal; NPR2: nitrogen permease regulator 2; NPR3: nitrogen permease regulator 3-like protein; PCWDE: plant cell wall degrading enzyme; PKc: protein kinases catalytic domain; PLC-like: phospholipase C-like; P-loop NTPase: P-loop protein with a nucleoside triphosphate hydrolase domain; PWWP: Pro-Trp-Trp-Pro; QC: quality control; *R. clarus*: *Rhizophagus clarus*; *R. irregularis*: *Rhizophagus irregularis*; RCP: repeats containing proteins; SCR: small cysteine-rich; SP: signal peptide; SP7: secreted effector protein 7; TMD: transmembrane domain; TORC1: target of rapamycin kinase; TPS: trehalose-6-phosphate synthase.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

Conceived and designed the experiments: AB, KST. Genome sequencing and assembly: AB. Bioinformatics and data analysis: KST. Wrote manuscript: KST, AB. Both authors read and approved the final manuscript.

### Acknowledgements

The authors would like to thank Max Griesmann for help with the BUSCO analysis. This work was supported by the German Research Foundation grant (DFG BR 3527/1-1 / PA 493/11-1).

Received: 29 October 2015 Accepted: 1 February 2016

Published online: 09 February 2016

### References

- Wang B, Qiu YL. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*. 2006;16:299–363.
- Declerck S. *In vitro* culture of mycorrhizas. Vol. 4. Berlin Heidelberg: Springer-Verlag; 2005.
- Kiers ET, Duhamel M, Beesetty Y, Mensah JA, Franken O, Verbruggen E, et al. Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. *Science*. 2011;333:880–2.
- Rillig MC, Mummey DL. Mycorrhizas and soil structure. *New Phytol*. 2006;171:41–53.
- Bucking H, Shachar-Hill Y. Phosphate uptake, transport and transfer by the arbuscular mycorrhizal fungus *Glomus intraradices* is stimulated by increased carbohydrate availability. *New Phytol*. 2005;165:899–911.
- Cappellazzo G, Lanfranco L, Fitz M, Wipf D, Bonfante P. Characterization of an amino acid permease from the endomycorrhizal fungus *Glomus mosseae*. *Plant Physiol*. 2008;147:429–37.
- Govindarajulu M, Pfeffer PE, Jin H, Abubaker J, Douds DD, Allen JW, et al. Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature*. 2005;435:819–23.
- Lopez-Pedrosa A, Gonzalez-Guerrero M, Valderas A, Azcon-Aguilar C, Ferrol N. *GintAMT1* encodes a functional high-affinity ammonium transporter that is expressed in the extraradical mycelium of *Glomus intraradices*. *Fungal Genet Biol*. 2006;43:102–10.
- Hata S, Kobae Y, Banba M. Interactions between plants and arbuscular mycorrhizal fungi. *Internat Rev Cell Mol Biol*. 2010;281:1–48.
- Schüßler A, Walker C. The Glomeromycota: a species list with new families and new genera. Gloucester, England: The Royal Botanic Garden Edinburgh, Botanische Staatssammlung Munich, Oregon State University; 2010.
- Remy W, Taylor TN, Hass H, Kerp H. Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proc Natl Acad Sci U S A*. 1994;91:11841–3.
- Loreau M, Naeem S, Inchausti P. Biodiversity and ecosystem functioning. Oxford, Great Britain: Oxford University Press; 2002.
- Parniske M. Intracellular accommodation of microbes by plants: a common developmental program for symbiosis and disease? *Curr Opin Plant Biol*. 2000;3:320–8.
- Corradi N, Bonfante P. The arbuscular mycorrhizal symbiosis: origin and evolution of a beneficial plant infection. *PLoS Pathog*. 2012;8:e1002600.
- Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A, et al. The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. *New Phytol*. 2012;193:755–69.
- Plett JM, Kemppainen M, Kale SD, Kohler A, Legue V, Brun A, et al. A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Curr Biol*. 2011;21:1197–203.
- Kamoun S. A catalogue of the effector secretome of plant pathogenic oomycetes. *Ann Rev Phytopathol*. 2006;44:41–60.
- van der Hooft RA, Kruijt M, Roth R, Brandwaag BF, Joosten MH, de Wit PJ. Intragenic recombination generated two distinct Cf genes that mediate AVR9 recognition in the natural population of *Lycopersicon pimpinellifolium*. *Proc Natl Acad Sci U S A*. 2001;98:10493–8.
- Rooney HC, Van't Klooster JW, van der Hooft RA, Joosten MH, Jones JD, de Wit PJ. *Cladosporium Avr2* inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science*. 2005;308:1783–6.
- Jones JD, Dangl JL. The plant immune system. *Nature*. 2006;444:323–9.
- Ellis JG, Rafiqi M, Gan P, Chakrabarti A, Dodds PN. Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. *Curr Opin Plant Biol*. 2009;12:399–405.
- Raffaele S, Win J, Cano L, Kamoun S. Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*. *BMC Genomics*. 2010;11:637.
- Plett JM, Martin F. Reconsidering mutualistic plant-fungal interactions through the lens of effector biology. *Curr Opin Plant Biol*. 2015;26:45–50.
- Kloppholz S, Kuhn H, Requena N. A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr Biol*. 2011;21:1204–9.
- Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R, et al. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci U S A*. 2013;110:20117–22.
- Bago B, Pfeffer PE, Douds Jr DD, Brouillette J, Becard G, Shachar-Hill Y. Carbon metabolism in spores of the arbuscular mycorrhizal fungus *Glomus intraradices* as revealed by nuclear magnetic resonance spectroscopy. *Plant Physiol*. 1999;121:263–72.
- Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. 2008;9:386.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–2.
- Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2007;2:953–71.
- Kall L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction - the Phobius web server. *Nucl Acids Res*. 2007;35:W429–32.
- de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucl Acids Res*. 2006;34:W362–5.
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. *Nucl Acids Res*. 2007;35:W585–7.
- Garg A, Raghava GP. ESLpred2: improved method for predicting subcellular localization of eukaryotic proteins. *BMC Bioinformatics*. 2008;9:503.
- Hawkins J, Davis L, Boden M. Predicting nuclear localization. *J Proteome Res*. 2007;6:1402–9.
- Stergiopoulos I, de Wit PJ. Fungal effector proteins. *Ann Rev Phytopathol*. 2009;47:233–63.

36. Saunders DG, Win J, Cano LM, Szabo LJ, Kamoun S, Raffaele S. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PLoS ONE*. 2012;7:e29847.
37. Jorda J, Kajava AV. T-REKS: identification of Tandem REpeats in sequences with a K-meanS based algorithm. *Bioinformatics*. 2009;25:2632–8.
38. Newman AM, Cooper JB. XSTREAM: a practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. *BMC Bioinformatics*. 2007;8:382.
39. Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*. 2008;2008:619832.
40. Sonnhammer EL, Eddy SR, Birney E, Bateman A, Durbin R. Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucl Acids Res*. 1998;26:320–2.
41. Gough J, Karplus K, Hughey R, Chothia C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol*. 2001;313:903–19.
42. Jensen LJ, Gupta R, Staerfeldt HH, Brunak S. Prediction of human protein function according to Gene Ontology categories. *Bioinformatics*. 2003;19:635–42.
43. Jung JK, McCouch S. Getting to the roots of it: genetic and hormonal control of root architecture. *Frontiers Plant Sci*. 2013;4:186.
44. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat Rev*. 2010;11:539–48.
45. Sedziewska KA, Fuchs J, Temsch EM, Baronian K, Watzke R, Kunze G. Estimation of the *Glomus intraradices* nuclear DNA content. *New Phytol*. 2011;192:794–7.
46. Lin K, Limpens E, Zhang Z, Ivanov S, Saunders DG, Mu D, et al. Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genet*. 2014;10:e1004078.
47. de Jonge R. *In silico* identification and characterization of effector catalogs. In: Bolton M, Thomma B, editors. *Plant Fungal Pathogens*. New York, Heidelberg, London: Springer; 2012. p. 415–26.
48. Rafiqi M, Gan PH, Ravensdale M, Lawrence GJ, Ellis JG, Jones DA, et al. Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell*. 2010;22:2017–32.
49. van Damme M, Bozkurt TO, Cakir C, Schornack S, Sklenar J, Jones AM, et al. The Irish potato famine pathogen *Phytophthora infestans* translocates the CRN8 kinase into host plant cells. *PLoS Pathog*. 2012;8:e1002875.
50. Plett JM, Martin F. Mutualistic effectors: architects of symbiosis. In: Martin F, Kamoun S, editors. *Effectors in plant-microbe interactions*. Oxford, UK: John Wiley & Sons, Inc; 2012. p. 295–326.
51. Giraldo MC, Valent B. Filamentous plant pathogen effectors in action. *Nat Rev Microbiol*. 2013;11:800–14.
52. Chinchilla D, Boller T. Innate Immunity: Pattern Recognition in Plants. In: Martin F, Kamoun S, editors. *Effectors in plant-microbe interactions*. Oxford, UK: John Wiley & Sons, Inc; 2012. p. 3–32.
53. Buist G, Steen A, Kok J, Kuipers OP. LysM, a widely distributed protein motif for binding to (peptidoglycans. *Mol Microbiol*. 2008;68:838–47.
54. Craig KL, Tyers M. The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. *Progress Biophysics Mol Biol*. 1999;72:299–328.
55. Yu XC, Zhu SY, Gao GF, Wang XJ, Zhao R, Zou KQ, et al. Expression of a grape calcium-dependent protein kinase ACPK1 in *Arabidopsis thaliana* promotes plant growth and confers abscisic acid-hypersensitivity in germination, postgermination growth, and stomatal movement. *Plant Mol Biol*. 2007;64:531–8.
56. Park CH, Chen S, Shiresekar G, Zhou B, Khang CH, Songkumarn P, et al. The Magnaporthe oryzae effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice. *Plant Cell*. 2012;24:4748–62.
57. Neklesa TK, Davis RW. A genome-wide screen for regulators of TORC1 in response to amino acid starvation reveals a conserved Npr2/3 complex. *PLoS Genet*. 2009;5:e1000515.
58. Hubbard SR, Till JH. Protein tyrosine kinase structure and function. *Ann Rev Biochem*. 2000;69:373–98.
59. Hicks SN, Jezyk MR, Gershburg S, Seifert JP, Harden TK, Sondek J. General and versatile autoinhibition of PLC isozymes. *Mol Cell*. 2008;31:383–94.
60. Vines CM, Phospholipase C. *Adv Exp Medicine Biol*. 2012;740:235–54.
61. Corrotte M, Chasserot-Golaz S, Huang P, Du G, Ktistakis NT, Frohman MA, et al. Dynamics and function of phospholipase D and phosphatidic acid during phagocytosis. *Traffic*. 2006;7:365–77.
62. Eastmond PJ, Graham IA. Trehalose metabolism: a regulatory role for trehalose-6-phosphate? *Curr Opin Plant Biol*. 2003;6:231–5.
63. Iordachescu M, Imai R. Trehalose biosynthesis in response to abiotic stresses. *J Integrative Plant Biol*. 2008;50:1223–9.
64. Tournu H, Fiori A, Van Dijk P. Relevance of trehalose in pathogenicity: some general rules, yet many exceptions. *PLoS Pathog*. 2013;9:e1003447.
65. Wilson RA, Gibson RP, Quispe CF, Littlechild JA, Talbot NJ. An NADPH-dependent genetic switch regulates plant infection by the rice blast fungus. *Proc Natl Acad Sci U S A*. 2010;107:21902–7.
66. Inohara N, Nunez G. ML - a conserved domain involved in innate immunity and lipid metabolism. *Trends Biochem Sci*. 2002;27:219–21.
67. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiol Biotechnol*. 2011;91:1477–92.
68. Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussoy F, et al. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*. 2008;452:88–92.
69. Whittaker MM, Kersten PJ, Nakamura N, Sanders-Loehr J, Schweizer ES, Whittaker JW. Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. *J Biol Chem*. 1996;271:681–7.
70. Kersten PJ. Glyoxal oxidase of *Phanerochaete chrysosporium*: it's characterization and activation by lignin peroxidase. *Proc Natl Acad Sci U S A*. 1990;87:2936–40.
71. Gonzalez R, Reguera E, Mendoza L, Figueroa JM, Sanchez-Sinencio F. Physicochemical changes in the hull of corn grains during their alkaline cooking. *J Agricultural Food Chem*. 2004;52:3831–7.
72. Baker LG, Specht CA, Donlin MJ, Lodge JK. Chitosan, the deacetylated form of chitin, is necessary for cell wall integrity in *Cryptococcus neoformans*. *Eukaryot Cell*. 2007;6:855–67.
73. Mast SW, Moremen KW. Family 47 alpha-mannosidases in N-glycan processing. *Methods Enzymol*. 2006;415:31–46.
74. Zhao Y, Park RD, Muzzarelli RA. Chitin deacetylases: properties and applications. *Mar Drugs*. 2010;8:24–46.
75. Gueddari NE, Rauchhaus U, Moerschbacher BM, Deising HB. Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytol*. 2002;156:103–12.
76. Hann D, Boller T. Microbial effectors and their role in plant defense suppression. In: Martin F, Kamoun S, editors. *Effectors in plant-microbe interactions*. Oxford, UK: John Wiley & Sons, Inc; 2012. p. 33–52.
77. Hu J, Wang Y. p53 and the PWWP domain containing effector proteins in chromatin damage repair. *Cell & Developmental Biol*. 2013;2:1–11.
78. Snider J, Thibault G, Houry WA. The AAA+ superfamily of functionally diverse proteins. *Genome Biol*. 2008;9:216.
79. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*. 2004;382:769–81.
80. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, et al. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science*. 2010;330:1549–51.
81. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A*. 2011;108:9166–71.
82. Mueller O, Kahmann R, Aguilar G, Trejo-Aguilar B, Wu A, de Vries RP. The secretome of the maize pathogen *Ustilago maydis*. *Fungal Genet Biol*. 2008;45:563–70.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
www.biomedcentral.com/submit

