Fungal Genetics and Biology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Regular Articles

Comparative analyses of secreted proteins in plant pathogenic smut fungi and related basidiomycetes

Mariana Schuster [1], Gabriel Schweizer [1], Regine Kahmann *

Max Planck Institute for Terrestrial Microbiology, Dept. Organismic Interactions, 35043 Marburg, Germany

ARTICLE INFO

Article history:
Received 5 July 2016
Revised 13 December 2016
Accepted 15 December 2016
Available online xxxx

Keywords:
Comparative genomics
Basidiomycete fungi
Ustilago maydis
Secretome
Cas9
Multiplexing

ABSTRACT

In the ten years since the genome sequence of the basidiomycete corn smut fungus *Ustilago maydis* was published, additional genomes of smut species infecting different hosts became available. In addition, the genomes of related *Malassezia* species causing skin diseases and of *Pseudozyma* species not known to infect plants were determined. As secreted proteins are critical virulence determinants in *U. maydis* we compare here the secretomes of 12 basidiomycete species to gain information about their composition and conservation. For this we classify secreted proteins into those with and without domains using InterPro scans. Homology among proteins is inferred by building clusters based on pairwise similarities and cluster presence is then assessed in the different species. We detect in particular a strong correspondence between the secretomes of *Pseudozyma* species and plant infecting smuts. Furthermore, we identify a high proportion of secreted proteins to be part of gene families and present an advancement of the CRISPR-Cas9 technology for simultaneous disruption of multiple genes in *U. maydis* using five genes of the *eff1* family as example.

1. Introduction

Smut fungi of the order Ustilaginales represent a large group of plant pathogens which mostly infect members of the grass family including very important crop plants like maize, wheat, barley and sugarcane. The most prominent members are *U. maydis*, causing corn smut in maize, *Sporisorium reilianum*, causing maize head smut, *U. hordei* causing covered smut in barley, *U. tritici* causing loose smut of wheat and *S. scitamineum* infecting sugarcane. Except for *U. tritici* the genome sequences of these smut fungi have all been determined (Dutheil et al., 2016; Kämper et al., 2006; Laurie et al., 2012; Que et al., 2014; Schirawski et al., 2010; Taniguti et al., 2015). Recently, the genomes of two dicot-infecting species *Melanopsichium pennsylvanicum* causing gall smut of *Persicaria* species and *Ceraceosorus bombacis* infecting the cotton tree *Bombax ceiba* have also been sequenced (Sharma et al., 2014, 2015). All smut fungi are biotrophic plant pathogens which need to keep the infected host alive to cause disease. In addition, smut fungi have in common that their pathogenic form is the filamentous dikaryon. This form is generated after mating of compatible haploid, budding cells. Compatibility is determined by tetrapolar or bipolar mating type systems in which cell-cell recognition is mediated by lipopeptide pheromones and their cognate receptors. After cell fusion the switch to the filamentous dikaryotic form and its maintenance are controlled through a heterodimer of two homeodomain proteins. Again, common to all smut fungi is that their development inside the plant is intimately connected with sexual reproduction culminating in the production of huge masses of black diploid teliospores. In nature, the mating reaction occurs on the leaf surface. The resulting dikaryon then develops infection structures allowing direct penetration in a process likely aided by localized secretion of plant cell wall degrading enzymes (Lanver et al., 2014; Schirawski et al., 2010). However, smut fungi are poorly equipped with plant cell wall degrading enzymes, presumably reflecting that damage to the host has to be minimized to establish the biotrophic life style (Lo Presti et al., 2015). During the early infection stages invading fungal hyphae become encased by the host plasma membrane and this creates an extended, tight interaction zone. Later, fungal hyphae accumulate in cavities in mesophyll tissue and in and around the veins, where they are considered to be supplied with nutrients by the host. In tissue infected by *U. maydis* this stage coincides with tumor formation, in which leaf cells enlarge and resume mitotic divisions (Redkar et al., 2015). In tumor tissue sexual development of *U. maydis* is completed with the formation of spores. Similar to *U. maydis*, *Me.*

* Corresponding author at: Max Planck Institute for Terrestrial Microbiology, Dept. of Organismic Interactions, Karl-von-Frisch-Str.10, 35043 Marburg, Germany.
E-mail address: kahmann@mpi-marburg.mpg.de (R. Kahmann).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.fgb.2016.12.003
1087-1845/

pennsylvanicum infection induces tumor-like structures in vegetative tissue (Haliski, 1962). The other smut species mentioned initially cause an asymptomatic infection and symptoms, i.e. dark-colored spores, develop only late and exclusively in the male and female flowers (Vánky, 2012). So far it is unknown what links spore production in these species with flowering.

The establishment of the biotrophic interaction of smut fungi with their respective host plants is governed by secreted proteins. Such proteins actively suppress plant immune responses, facilitate nutrition and regulate the progress of the infection process. The sequencing of the genome of *U. maydis* ten years ago (Kämper et al., 2006), paved the way for the discovery of a large repertoire of novel, conventionally secreted proteins. This was aided by observing clustering of such genes in the genome, by demonstrating that the respective genes are induced during infection and by experimental verification that several clusters contribute to virulence presumably by affecting the activity of certain plant proteins (Kämper et al., 2006). Pathogen effectors are now grouped into apoplastic if they remain in the interaction zone/apoplast after secretion, and cytoplasmic if they are taken up and function inside host cells (Win et al., 2012). By now several of the secreted *U. maydis* virulence-promoting effectors have been functionally characterized. Pep1 is an inhibitor of the apoplastic plant peroxidase POX12 (Hemetsberger et al., 2012) and Pit2 inhibits a group of apoplastic cysteine proteases (Mueller et al., 2013). Both proteins play crucial roles in suppressing plant defense reactions (Doehlemann et al., 2011, 2009). The secreted *U. maydis* chorismate mutase Cmu1 is taken up by plant cells where it lowers salicylic acid levels by altering the chorismate homeostasis (Djamei et al., 2011). Other translocated effectors are Tin2 and See1. Tin2 stabilizes a maize kinase responsible for anthocyanin induction and by doing so is suggested to divert metabolites away from the lignin biosynthesis pathway that negatively affects fungal spreading in the host tissue (Tanaka et al., 2014). See1 is required for the reactivation of plant DNA synthesis after infection, which is crucial for tumor progression in leaf cells (Redkar et al., 2015). So far most functional analyses of secreted proteins were done in the *U. maydis*-maize system because this system is the most advanced and offers the most efficient toolbox for genome manipulation (Basse and Steinberg, 2004; Garcia-Pedrajas et al., 2010; Kämper et al., 2006; Khrunyk et al., 2010; Terfruchte et al., 2014), including the CRISPR-Cas9 system for genome editing (Schuster et al., 2016). In addition, in the *U. maydis*-maize interaction symptoms can be scored about a week after seedling infection and there is no need to wait until flower development. Furthermore, the system has greatly benefitted from the construction of so-called solopathogenic lines which are haploid and engineered to cause disease without having to undergo mating (Bölker et al., 1995; Kämper et al., 2006). Besides the genomes of the four grass-infecting smuts and the two species infecting dicot plants the genome of *Microbotryum lychnidis-dioicae*, a member of the Microbotryales causing anther smut in *Caryophyllaceae* species, has recently been determined (Perlin et al., 2015). In addition, the genomes of several species related to the smut fungi but not known to colonize plants have been sequenced. These are members of the genus *Malassezia* associated with dandruff (Gioti et al., 2013; Wu et al., 2015; Xu et al., 2007), and *Pseudozyma*, which are mostly used in biotechnology as well as for biocontrol purposes (Konishi et al., 2013; Lefebvre et al., 2013; Lorenz et al., 2014; Morita et al., 2014; Oliveira et al., 2013). *P. antarctica* and *P. flocculosa* were recently transferred to the genus *Moesziomyces* (Wang et al., 2015) and *Anthracocystis* (Piatek et al., 2015), respectively. For this communication we will keep their original names as these appear in publications we refer to. With respect to the phylogeny of the species compared here there are inconsistencies which may reflect whether multigene phylogenies or whole proteome phylogenies were applied (Dutheil et al., 2016; Lefebvre et al., 2013; Sharma et al., 2015; Wang et al., 2015). Based on the latter approach the grass infecting smut fungi are found in one monophyletic group in which *U. hordei* and *U. maydis* are placed ancestral to *S. reilianum* and *S. scitamineum.* (Dutheil et al., 2016; Lefebvre et al., 2013). The same group contains the dicot-infecting species *Me. pennsylvanicum* (Dutheil et al., 2016; Sharma et al., 2014). Ancestral to this group is the early branching smut fungus *C. bombacis*, a member of the Exobasidiomycetes (Sharma et al., 2015). Based on multigene phylogeny *P. antarctica* is ancestral to the four grass infecting smut fungi and *Me. pennsylvanicum*, while *P. flocculosa* is ancestral to the two *Sporisorium* species (Wang et al., 2015). *Ma. globosa* and *Ma. sympodialis* comprise an ancestral group to the clade containing plant pathogenic species (Xu et al., 2007). The anther smut fungus *Mi. lychnidis-dioicae* is a member of the Microbotryaceae family, i.e. only very distantly related to the other smut species which are all members of the Ustilaginaceae (Perlin et al., 2015).

In this communication we analyze the secreted protein repertoire of smut fungi and their relatives using the basidiomycete *Schizophyllum commune* as outgroup. We categorize secreted proteins into different groups, i.e. core secreted proteins, which are shared by all 12 species or certain subgroups compared, species-specific secreted proteins, secreted proteins with predicted domains and secreted proteins without any predicted domains. We restrict the term secreted effector to secreted proteins without domains in plant pathogens. We then look at gene families encoding secreted proteins and describe how these families are distributed. So far, only few of such gene families have been functionally analyzed (Basse et al., 2002, 2000; Doehlemann et al., 2008; Khrunyk et al., 2010; Lanver et al., 2014) due to technical limitations. To overcome this we describe modifications of the CRISPR-Cas9 technology already adapted to *U. maydis* (Schuster et al., 2016) to allow multiplexing for inactivating members of gene families simultaneously.

2. Materials and methods

2.1. Prediction of protein secretion and function

All fungal species used in this study, their number of gene models and sources of genome data are listed in Supplementary Table S1. The proteome of all species was used to predict secreted proteins. As a first step, we excluded all proteins that did not have methionine as start codon. In this way, 20 proteins in *S. reilianum*, 17 in *S. scitamineum*, 30 in *U. maydis*, 27 in *U. hordei*, 28 in *Me. pennsylvanicum*, 4 in *M. globosa*, 25 in *P. antarctica*, 942 in *S. commune* and 56 in *M. lychnidis-dioicae* are excluded from further analysis. For the remaining proteins signal peptides were predicted using SignalP 4.0 (Petersen et al., 2011). Furthermore, transmembrane regions were predicted by TMHMM 2.0c (Krogh et al., 2001) and Phobius 1.01 (Kall et al., 2004). Proteins were considered to be secreted if SignalP identified a secretion signal peptide and no transmembrane domain was predicted by SignalP 4.0, TMHMM2.0c or Phobius 1.01. Subsequently, the set of predicted secreted proteins from all species was used for functional predictions and inference of homology.

InterPro scan 5.18–57.0 (Jones et al., 2014) was used to predict functions or domains of all secreted proteins. Proteins containing domains of known function or structural domains or a combination of both were grouped together. Secreted proteins without any domain were placed together in a second group. Putative plant cell wall degrading enzymes (PCWDE) were identified by manually linking InterPro information to CAZyme entries which are considered to function as PCWDE (Lo Presti et al., 2015).

Please cite this article in press as: Schuster, M., et al. Comparative analyses of secreted proteins in plant pathogenic smut fungi and related basidiomycetes. Fungal Genet. Biol. (2017), http://dx.doi.org/10.1016/j.fgb.2016.12.003