CHAPTER SIX

Advancing Knowledge on Biology of Rust Fungi Through Genomics

Sébastien Duplessis[*,1], **Guus Bakkeren**[†], **Richard Hamelin**[‡,§]

[*]Institut National de la Recherche Agronomique (INRA), UMR 1136 INRA/Lorraine University, Interactions Arbres/Micro-organismes, Centre de Nancy, Champenoux, France
[†]Pacific Agri-Food Research Centre, Agriculture and Agri-Food Canada, Summerland, British Columbia, V0H 1Z0 Canada
[‡]Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Quebec, Canada
[§]Department of Forest and Conservation Sciences, The University of British Columbia, Vancouver, British Columbia, Canada
[1]Corresponding author: e-mail address: duplessi@nancy.inra.fr

Contents

Abstract

Pucciniales are an important group of fungal plant pathogens that cause rust diseases in a diverse group of hosts including ecologically and economically important crops and trees. Rust fungi have intriguing and complex life cycles and are obligate biotrophs. Because of their biological features, these fungi are very difficult to study under laboratory conditions. The recent advances in genomics and transcriptomics have opened great perspectives for making progress in the study of this group of fungi and more particularly to dissect the genetic determinants underlying the host infection process. In this chapter, we provide an overview of the current knowledge on rust genomics and

Advances in Botanical Research, Volume 70
ISSN 0065-2296
http://dx.doi.org/10.1016/B978-0-12-397940-7.00006-9

we particularly highlight how next-generation sequencing technologies are moving this field forward, providing new avenues in the understanding of fungal biotrophy.

1. INTRODUCTION

Rust fungi (order Pucciniales) are an important and diverse group of plant pathogens that can affect a diverse group of hosts and cause a remarkable range of symptoms such as defoliation, cankers, and witch brooms (Cummins & Hiratsuka, 2003). Rust fungi cause diseases that are responsible for some of the most severe economic losses of trees and crops, including pines, poplars, eucalyptus, wheat, coffee, and soybean. In addition to causing economic losses, rusts are responsible for impacting ecosystems.

Rusts possess unique biological features and ecology. They are strict biotrophic fungi that require a living host to complete their life cycle. Rust fungi display a diversity of life cycles. Heteroecious macrocyclic rusts possess one of the most intriguing life cycles in the fungal kingdom. They must alternate between telial and aecial hosts and produce five different spore types in order to complete their life cycle (Fig. 6.1). The symptoms caused by heteroecious rusts on telial and aecial hosts are usually different and can occur on completely unrelated hosts. For example, *Cronartium ribicola* causes branch and stem cankers on the pine aecial host but only leaf infections on the telial *Ribes* spp. hosts. *Puccinia triticina*, the wheat leaf rust fungus, produces uredospores and the survival structures, the teliospores, on wheat, whereas the sexual stage occurs on a completely unrelated plant, meadow rue (*Thalictrum speciosissimum*). Other rusts have reduced life cycles. Autoecious rusts can complete their life cycle on a single host, while demicyclic and microcyclic rusts have reduced number of spore stages.

In spite of their economic importance, there are unresolved questions about rust fungi biology, epidemiology, and host–pathogen interactions that could provide critical knowledge and impact disease management. The need for host alternation in heteroecious rusts is still not clearly understood. What are the evolutionary advantages or compromises of host alternation? Do heteroecious rusts possess alternate sets of effector genes that allow them to interact with different hosts or are the same effectors acting in different hosts? Do autoecious rusts have a reduced gene set compared with their heteroecious counterparts? Another interesting and unresolved set of questions relates to host specificity. Most rusts have a relatively narrow host range

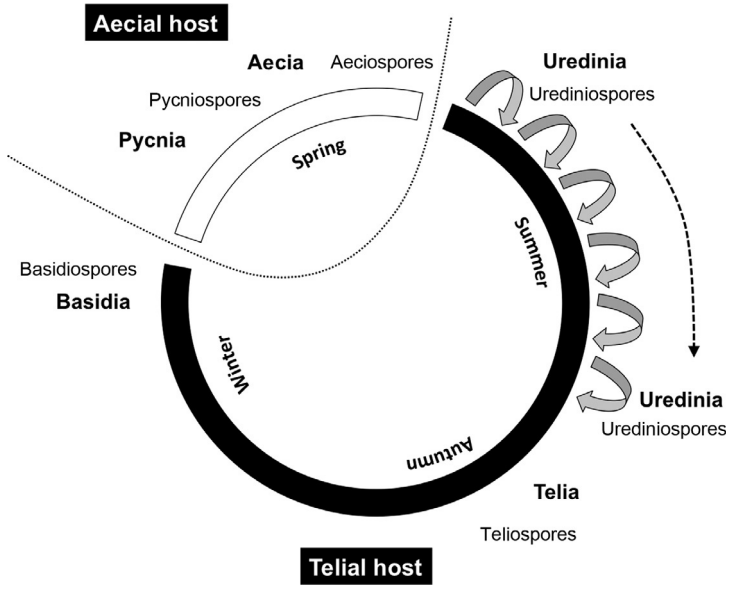

Figure 6.1 Schematic representation of the typical biological cycle of heteroecious and macrocyclic rust fungi. In spring, basidiospores infect the aecial host. Mycelia formed by the rust fungus produce pycnia, and pycniospores are released in nectar droplets in which fertilization can occur between spores and receptive hypha of compatible mating types. Spores are transmitted between nectars by insects, rain, or wind. Once dikaryons are formed, aecia are established and aeciospores are released. This leads to the telial host infection in spring. Infection of the telial host occurs throughout summer by repeated infection cycles producing uredinia, a structure that releases huge amounts of uredospores. This stage often leads to strong rust epidemics on the telial host over the growing season. At the end of summer/early autumn, black structures called telia are formed in the senescent tissues of the host. Teliospores differentiate during autumn and winter in these overwintering structures. After karyogamy and meiosis, teliospores germinate and form basidia that release basidiospores for infection of the aecial host, completing the cycle. Only the basidiospores are haploids; spores are dikaryotic at all other stages. Autoecious rust fungi proceed through their life cycle on a single host and microcyclic rust fungi only produce some of the spore types, for example, uredospores and teliospores.

on either the aecial or telial hosts. What are the determinants of this host specificity? What allows some rusts to have narrow host specificity on the aecial host but a broad host range on the telial host, or vice versa? What gives rust fungi the ability to rapidly adapt following the deployment of resistant host genotypes? At the core of the set of important questions is what drives the host–pathogen interactions. The gene-for-gene hypothesis states that each resistance gene in a host plant corresponds to an avirulent gene counterpart in the pathogen and their interaction triggers an immune response.

The experimental demonstration of this hypothesis was proven for the first time in the flax rust (Flor, 1959). These questions can be addressed using a combination of novel genomics approaches that can be used to dissect these interactions and identify genetic determinants in both the host and the pathogen.

There have been exciting new developments in the past decade in our understanding of genetics and genomics of rust fungi that is informing us on their biology, host–pathogen interactions, evolution, and epidemiology. This chapter aims at describing the latest developments in rust genomics and to take a look at future developments.

2. RUST FUNGI IN THE GENOMICS ERA

2.1. Genomics of plant-interacting fungi

Most fungal genomes exhibit a small size of 30–60 Mb compared with other eukaryotes (Raffaele & Kamoun, 2012). Fungi with smaller and compact genomes have been reported, such as Pucciniomycotina *Mixia osmundae* or members of the Ustilaginomycotina *Sporisorium reilianum*, *Ustilago maydis*, or *Ustilago hordei* (Kämper et al., 2006; Laurie et al., 2012; Schirawski et al., 2010; Toome et al., 2014). But a few are beyond this range, such as in the ascomycete *Tuber melanosporum*, the gourmet-favorite black truffle, *Blumeria* spp. causing powdery mildew diseases on plants, and the rust fungi (Duplessis, Cuomo, et al., 2011; Martin et al., 2010; Spanu et al., 2010; Wicker et al., 2013). Recent genomics reviews examined the reasons behind this wide variation in genome sizes in plant-interacting fungi and fungal-like organisms. Among the factors considered are content in repetitive elements including transposable elements (TE) and gene losses shared between different phyla of obligate biotrophic pathogens (Kemen & Jones, 2012; Martin, 2014; Raffaele & Kamoun, 2012; Spanu, 2012). The genomes of fungi interacting with plants are also marked by specific repertoires of genes related to the mutualistic or parasitic relationship established with the host plant. Genes encoding enzymes involved in the decomposition of plant cell wall components (e.g. carbohydrate-active enzymes, CAZymes), signal transduction, transport of nutrients and water, or effectors modulating the host immunity are essential for the success of colonization and acquisition of nutrients from the hosts. The content in CAZymes much likely reflects the fungal lifestyle. Mutualistic ectomycorrhizal fungi or obligate biotrophs such as rust fungi have a reduced or moderate ability to alter the plant cell wall, possibly to minimize triggering a plant response during colonization

(Duplessis, Cuomo, et al., 2011; Martin et al., 2008). In oomycetes and fungi interacting with plants, a large subset of genes code for proteins predicted to be secreted, the 'secretomes', and a subset of these consist of relatively small proteins that represent the effector repertoires necessary to achieve successful infection (Lowe & Howlett, 2012; Raffaele & Kamoun, 2012; Spanu, 2012). Those genomes are often marked by transposon invasions and eventually by the presence of multigene families, which explains their large size compared with those of other fungi and denotes a greater plasticity (Murat, Payen, Petitpierre, & Labbé, 2014; Raffaele & Kamoun, 2012). The association between candidate effector gene families and transposon-rich regions has been reported in the genomes of different plant pathogens (e.g. *Leptosphaeria maculans* and *P. infestans*) suggesting that transposon activity could have been central in the diversification of their effector repertoires (Haas et al., 2009; Rouxel et al., 2011). Indeed, effector genes found in transposon-rich regions show highly polymorphic profiles strongly supporting a possible adaptation to the host immune system (Raffaele & Kamoun, 2012).

2.2. Sequencing genomes of rust fungi

Progress has been made in the last decade, from sequencing rust fungi genomes with whole-genome shotgun strategies using the Sanger sequencing technology to NGS (Table 6.1) (Duplessis, Spanu, & Schirawski, 2014). This first generation of genome sequences has provided genome assemblies of great quality, despite a low-depth coverage due to the high cost of this technology compared with the more-recent next-generation sequencing techniques. The poplar leaf rust *M. larici-populina* and the wheat stem rust *P. graminis* f. sp. *tritici* were the first obligate biotrophic plant pathogens in the Basidiomycota phylum sequenced (Duplessis, Cuomo, et al., 2011; McDowell, 2011). The two genomes show a composition and organization illustrating not only striking similarities within the order Pucciniales but also clear differences with other obligate biotrophic pathogens in the Ascomycota and Basidiomycota phyla, such as powdery mildews and smut fungi. Indeed, both rust fungi have large genomes (80–101 Mb), a large number of repeats and TE (nearly 50% of the genome), and a large number of genes (>16,000), whereas other obligate biotrophic fungal pathogens only share one or two of these characteristics (see Duplessis et al., 2014, for details). The large genome size and the large number of genes observed for rust fungi might be related to the heteroecious and macrocyclic nature of the two species that were selected for genome sequencing. Sequencing the genomes of autoecious and microcyclic rust fungi is now required to

Table 6.1 Update on genome sequences statistics of Pucciniales

Rust species and isolate (reference)	*Melampsora larici-populina* 98AG31 (Duplessis, Cuomo, et al., 2011[a])	*Puccinia graminis* f. sp. *tritici* CDL75-36-700-3 (Duplessis, Cuomo, et al., 2011[a])	*Puccinia striiformis* f. sp. *tritici* PST-130 (Cantu et al., 2011)	*Puccinia striiformis* f. sp. *tritici* CY32 (Zheng et al., 2013)	*Puccinia triticina* 1-1 (BBBD) (C. Cuomo, J. Fellers, L. Szabo and G. Bakkeren unpublished[a])
Sequencing approach	Sanger sequencing, whole-genome shotgun strategy	Sanger sequencing, whole-genome shotgun strategy, Illumina sequencing to validate SNPs between haplotypes	Illumina Genome Analyser II sequencing (53.8 million reads paired-end 85–101 pb reads); draft genome	Fosmid-to-fosmid strategy by Illumina GA paired-end sequencing (1920 pools of 19,200 36 Kb average fosmids; validation of 10 selected fosmids by end-to-end Sanger sequencing); whole genome	Illumina whole-genome shotgun strategy, paired-end 454 pyrosequencing, paired-end Sanger sequencing of 15,000 BAC clones (see Fellers et al., 2013)
Number of contigs/ scaffolds	3264/462	4557/392	29,178 (Contigs only)	12,833/4283	14,820
Total assembly size of contigs/ scaffolds	97.7/101.1 Mb	81.5/88.6 Mb	64.8 Mb (estimated genome size 68.2–78.8 Mb)	115.5 Mb/130.6 Mb (estimated size, 110 Mb)	135.34 Mb (estimated size, 135 Mb)
Gap (% of genome)	3.4%	8%	n.a.	n.a.	n.a.
Contigs N50/ L50	27/1.1 Mb	30/0.97 Mb	n.a./5.1 Kb	18 Kb/n.a.	10.37 Kb
Scaffolds N50/ L50	265/112.3 Kb	546/39.5 Kb	n.a.	125.3 Kb/n.a.	544.26 Kb

Sequencing depth (*x*-fold)	6.9 ×	12 ×	59 ×	26 ×	31 ×
Transposable and repetitive elements (% of genome)	45%	43.7%	17.8% (partial annotation)	48.9%	42–45%
Number of predicted genes	16,399	17,773 (revised to 15,800 by the Broad Institute upon RNA-seq analysis)	20,423 (lowered to 18,149 in Cantu et al., 2013)	25,288	14,880
Resequencing of isolates	Resequencing of 14 isolates from diverse populations collected over 20 years in France, at a 20–30 × depth by Illumina paired-end sequencing; mapping onto the 98AG31 reference genome, determination of SNPs, InDels and SV (A. Persoons & S. Duplessis, unpublished)		Resequencing of four isolates from the United States and the United Kingdom by Illumina, paired-end sequencing, determination of intra- and interisolates SNPs (Cantu et al., 2013 BMC Genomics): • PST-21 (66 × coverage, 43,106 contigs, 73 Mb total length, N50	Resequencing of five isolates from diverse origin worldwide by Illumina paired-end sequencing, mapping onto CY-32, determination of SNPs, InDels, and structural variants, determination of SNPs in PST-130: • CY23 (29 × coverage)	Resequencing of 120 isolates from diverse genetic lineages and race designations across North America, and Europe, by Illumina paired-end sequencing, mapping onto *Pt* 1-1, determination of SNPs, InDels, and structural variants (J. Fellers, J.A. Kolmer, C.A. Cuomo, G. Bakkeren, B.D. McCallum, & B. Saville, et al., unpublished)

Continued

Table 6.1 Update on genome sequences statistics of Pucciniales—cont'd

Rust species and isolate (reference)	*Melampsora larici-populina* 98AG31 (Duplessis, Cuomo, et al., 2011)	*Puccinia graminis* f. sp. *tritici* CDL75-36-700-3 (Duplessis, Cuomo, et al., 2011)	*Puccinia striiformis* f. sp. *tritici* PST-130 (Cantu et al., 2011)	*Puccinia striiformis* f. sp. *tritici* CY32 (Zheng et al., 2013)	*Puccinia triticina* 1-1 (BBBD) (C. Cuomo, J. Fellers, L. Szabo and G. Bakkeren unpublished)
			3.9 Kb, 20,653 genes), • PST-43 (26× coverage, 49,784 contigs, 71 Mb total length, N50 3.3 Kb, 21,036 genes), • PST-87/7 (15× coverage, 55,502 contigs, 53 Mb total length, N50 1.3 Kb, 20,688 genes), • PST-08/21 (21× coverage, 50,898 contigs, 56 Mb total length, N50 1.6 Kb, 20,875 genes)	• PK-CDRD (27× coverage) • Pst-78 (25× coverage) • Hu09-2 (23× coverage) • 104E137A- (22× coverage)	

[a]*M. larici-populina*: http://genome.jgi.doe.gov/programs/fungi/index.jsf; *Puccinia* spp.: http://www.broadinstitute.org/annotation/genome/puccinia_group/GenomesIndex.html.

Selected genomic features from the rust fungi genomes sequenced to date. The table also includes NGS resequencing data. n.a., non available

determine how far these biological features impact the genomic landscape of rust fungi. It is likely that TE invasion played a particularly important role in genome size expansion. The genome of the flax rust *Melampsora lini* is estimated to be close to 200 Mb (Peter Dodds, personal communication), and the genome of the microcyclic rust fungus *Phakopsora pachyrhizi*, responsible for the soybean rust disease, is estimated to be >800 Mb (Igor Grigoriev, personal communication). Analysis is ongoing and the number of predicted genes is not yet revealed, but it is very likely that repetitive DNA and TE will explain such genome size suggesting that genome expansion in rust fungi is a common trend that is not related to the ability to infect multiple host or to go through multiple differentiation stages in the biological cycles.

In the wake of these rust genome sequences, the draft genome of *Puccinia striiformis* f. sp. *tritici* (isolate PST-130) was obtained by Illumina-only sequencing (Cantu et al., 2011). A partial assembly of 65 Mb was gathered with a greater sequencing depth compared with the other rust genomes, but at a lower resolution (see number and size of contigs/scaffolds in Table 6.1). The striking differences in its assembly are likely due to variable genome sizes and differing content in repetitive sequences among rust fungi. Nonetheless, this pioneer sequencing of a rust genome with a NGS-only-based approach allowed the prediction of a draft catalogue of 20,423 putative genes, which was much larger than the number reported in the sister species *P. graminis* f. sp. *tritici* (Cantu et al., 2011). By combining different sequencing technologies and pools of large DNA fragments, a different isolate of the same wheat rust species originating from China (isolate CY32) was sequenced by Illumina and partially validated by Sanger sequencing (Zheng et al., 2013). The fosmid-to-fosmid sequencing approach used in this study helped to meet the challenging profile of rust genomes and to reach greater assembly quality at higher sequencing depth compared to the previous report for this species. After assembly, the genome of *P. striiformis* f. sp. *tritici* was estimated at a much larger size of 130 Mb, and the number of predicted genes was increased to 25,288. This places the wheat stripe rust fungus among the ones having the largest gene repertoire recorded so far in fungi, ahead of the mutualistic biotrophic fungus *Laccaria bicolor* (Martin et al., 2008). The repeat and TE content of this wheat leaf rust genome was estimated at nearly 50%, similar in range to previous observations in rust fungi. The strategy established by Zheng and collaborators greatly helped improving the assembly of this complex genome. The genome size was estimated to be twice the size of the draft sequence of the PST-130 isolate. Discrepancies observed between the two genomes

could have reflected major divergence between isolates, but this was not supported by the resequencing of genomes of isolates collected worldwide (Zheng et al., 2013). The comparison more likely illustrates the limitations of a sequencing strategy based only on the sequencing of small reads from small DNA fragments as input. Although the depth of sequencing and the assembly quality of the PST-130 isolate differed to standards in genomic studies, the report confirmed both the large size of genome alluded to a similarly large genome and the gene complements of rust fungi. It also demonstrated the possibility to generate a large amount of relevant genetic information for a nonmodel species that is particularly economically relevant. Indeed, although only a portion of the genome was captured when compared with the genome of the CY-32 isolate, a closely similar number of genes was found (later reduced to 18,149 in Cantu et al., 2013; Table 6.1). Although the assembled draft genome did not properly illustrate the complexity of the genomic landscape of *P. striiformis* f. sp. *tritici*, it did succeed in capturing a significant portion of the functional genome, the gene space. On the one hand, the strategy defined by Zheng et al. (2013) could be very promising in the future for fungal genomes enriched in TE and for which difficulties in assembly are encountered (e.g. *Blumeria* spp.; Hacquard, Kracher, et al., 2013; Wicker et al., 2013); on the other hand, obtaining draft genomes at lower coverage and at lower cost based only on NGS technologies could be a faster way to gather relevant functional information in important pathogens. These studies are providing promising avenues for future sequencing of rust fungi with very large estimated genomes, that is, the soybean rust *P. pachyrhizi* or the coffee rust *Hemileia vastatrix* (Carvalho, Carvalho, Barreto, & Evans, 2014).

2.3. Major genomic features of rust fungi

The comparative analysis of gene complements and synteny in rust genomes highlights a complex evolution in the order Pucciniales with intense TE activity, rearrangements, insertions, and deletions. Specific gene families were particularly abundant compared with other basidiomycetes. Most of these families do not have functional information, and among those with annotations, helicases, leucine-rich repeat proteins, oligopeptide transporters, and different types of glycosyl hydrolases, lipases and peptidases were particularly overrepresented (Duplessis, Cuomo, et al., 2011; Zheng et al., 2013). Such expansions may correspond to specific processes related to host infection, as well as DNA repair and maintenance. Interestingly, more genes

with gene ontology (GO) terms related to DNA recombination were found between the two wheat rust fungi than with *M. larici-populina* (Zheng et al., 2013). The expansion of genes with functional annotations related to DNA metabolism could also correlate with the invasion of the genomes of rust fungi by TEs. Whereas orthology was only evidenced for half of the genes between *M. larici-populina* and *P. graminis* f. sp. *tritici*, a greater conservation of the gene complements was observed within Pucciniaceae (Zheng et al., 2013). Few synteny blocks could be detected between the genomes of the poplar rust and the genomes of the wheat stem rust (Duplessis, Cuomo, et al., 2011). On the contrary, extensive evidence of microsynteny was found between the wheat rust fungi (Cantu et al., 2011; Fellers et al., 2013). Comparisons between *M. larici-populina*, *P. graminis* f. sp. *tritici*, and *P. striiformis* f. sp. *tritici* identified an important reshuffling of the assembled genomic scaffolds, consistent with ancient transposon activity (Zheng et al., 2013). The specificity of the gene catalogues in the Pucciniaceae and in the Melampsoraceae illustrates the old divergence between these two phylogenetic lineages (Aime et al., 2006). Future data from the *M. lini* genome analysis will help in determining with more precision the level of conservation of gene content at the taxonomic level of families within the order Pucciniales. Interestingly, the species-specific genes found in the three rust genomes published to date encode a majority of proteins of unknown function that belong to large expanded families, including secreted proteins (Duplessis, Cuomo, et al., 2011; Zheng et al., 2013).

2.4. NGS to assess genome-scale polymorphism in rust fungi

The facilitated access to NGS has allowed establishing programs for sequencing genomes of selected isolates for a given species and mapping onto reference genomes, opening the door to revealing population genomic and association genetic studies. Illumina sequencing was used to study polymorphisms within and between dikaryotic isolates of *P. striiformis* f. sp. *tritici* (Cantu et al., 2013; Zheng et al., 2013) (Table 6.1). Both studies relied on paired-end sequencing of small DNA fragments mapped to a reference genome at depth coverage of 15–66× in order to determine single-nucleotide polymorphisms (SNPs) and in one case other events such as insertions/deletions (InDels) and structure variants (Table 6.1). In the study by Cantu et al. (2013), the sequenced genomes of each isolate were *de novo* assembled and compared with the reference draft genome of PST-130 in order to assess completeness of sequencing, whereas in the other study,

the reference genome was the Chinese isolate CY-32 (Zheng et al., 2013). Considering the large differences between the PST-130 and CY-32 genome assemblies, it will be particularly relevant and useful to compare the two genomic resources and to eventually refine their assemblies. All *P. striiformis* f. sp. *tritici* genomes were sequenced from dikaryotic spores, and different types of polymorphic sites were recorded, such as intra- and extra-isolate SNPs. A large portion of the SNPs identified by Zheng et al. (2013) were located in coding regions (~1/3), and half of them were non-synonymous, indicating important genetic diversity. The two studies identified different levels of polymorphisms in the sequenced genomes. In total, between 81,000 and 109,000 SNPs were detected per isolate by Zheng et al. (2013) of which ~82–84% were heterozygous. Cantu et al. (2013) identified a much larger amount of SNPs (>350,000 per genome) with a similar high rate of heterozygous SNPs. Major differences in the final numbers of SNPs probably reflect the mapping strategy and the SNP calling parameters and programs used. Such differences in SNP detection of the same rust are quite disturbing and highlight the need to refine these analyses and define parameters that will allow cross study comparisons.

In comparison, 88,083 and 129,172 SNPs were detected in *M. larici-populina*, based on Sanger sequencing only (note the low sequencing depth at 6.9 ×), and in *P. graminis* f. sp. *tritici*, based on Illumina sequencing and comparison to the Sanger assembled genome, respectively (Duplessis, Cuomo, et al., 2011). The sequencing of 14 isolates of *M. larici-populina* by Illumina at a greater depth (between 20 and 30 ×) identified a larger number of polymorphic sites when all genomes were considered (>580,000); however, overall, an average of ~150,000 SNPs per genome was found indicating polymorphic rates in the range of those reported for *Puccinia* spp. genomes (A. Persoons & S. Duplessis, unpublished data).

Although the Illumina-based strategy chosen by Cantu et al. (2013) led to a smaller *de novo* genome assembly size for isolate PST-130 than for the CY-32 isolate, the number of predicted genes was in the same range with more than 20,000 gene models per genome. The predicted gene complements have a rather good support in term of complete copies of conserved eukaryotic genes (72.1% SD ± 14.6%). This indicates once more that an Illumina-only-based sequencing strategy applied at adequate depth can capture a significant portion of the functional genome despite the complexity of the genome per se. Of course, the critical limit is the quality of the manual curation performed after the automatic gene prediction.

3. RUST TRANSCRIPTOMICS

Before the availability of NGS technologies, most of the knowledge on molecular determinants of rust fungi was based on expressed sequence tags (ESTs) and cDNA-array-based expression profiling. Pioneer studies explored the molecular biology of a variety of Pucciniales mostly focusing on the process of host infection and on purified haustorial structures from infected plant tissues (for a complete review, see Duplessis, Joly, & Dodds, 2012; Fernandez, Talhinhas, & Duplessis, 2013). These gene expression studies particularly showed that the haustorium is a crucial structure both for nutrients uptake (Mendgen & Hahn, 2002; Voegele, Hahn, & Mendgen, 2009; Voegele, Struck, Hahn, & Mendgen, 2001) and effector release (Catanzariti, Dodds, Lawrence, Ayliffe, & Ellis, 2006; Hahn & Mendgen, 1997; Joly, Feau, Tanguay, & Hamelin, 2010; Link & Voegele, 2008). A few studies explored other stages of the rust fungi biology, providing insights into relevant functions expressed at various developmental steps of the rust life cycle (Warren & Covert, 2004; Xu et al., 2011). Most expression studies were focused on resting rust uredospores collected outside the host plant and during *in planta* biotrophic growth by directly isolating RNA from infected host tissues. In less than 20 years (1990–2010), a total of 168,199 Pucciniales ESTs have been deposited at the National Center for Biotechnology Information (as of January 2014). This number has hardly changed between 2010 and 2013, whereas in the same period, a significant number of publications appeared describing the use of genome-wide oligoarrays or of NGS approaches to determine the transcriptome of rust fungi at various stages and in different types of plant–rust interactions (for a detailed list, see Table 6.2). In total, nearly 100,000 unique genes or transcripts were reported by these studies for seven rust fungi, indicating that in only 4 years, significant knowledge has been generated compared to what was accumulated in the two decades before.

3.1. Genome oligoarray-based transcriptomics

The availability of the *M. larici-populina* and *P. graminis* f. sp. *tritici* genomes allowed the design of custom whole-genome oligoarrays to perform genome-wide expression surveys during the infection process of the telial hosts of these fungi, respectively, poplar and wheat (Duplessis, Cuomo, et al., 2011; Duplessis, Haquard, et al., 2011). These pioneer

Table 6.2 Recent genome-wide expression studies of Pucciniales

Species and isolate (reference)	Interaction, biological stage	Transcriptome approach	Number of genes covered	Number of genes detected
Hemileia vastatrix 178a, CIFC collection (Fernandez et al., 2012)	Coffee rust infected leaves at 21 dpi (3 days before uredinia formation)	454 pyrosequencing GS-FLX Titanium; 352,146 reads	Unknown, no reference genome	22,774 assembled contigs, 6763 assigned to the fungus
Melampsora larici-populina 98AG31(Hacquard et al., 2010)	Laser capture microdissection of uredinia (area 1); spongy mesophyll containing infection hyphae, haustoria, and sporogenous hyphae (area 2); and infected palisade mesophyll containing infection hyphae and haustoria (area 3) from infected poplar leaves (Beaupré cv.) at 7 dpi	Custom whole-genome oligoarrays	15,388 putative genes (published ahead of the release of expert genome annotation in Duplessis, Cuomo, et al., 2011)	8145, 7786, and 7288 expressed transcripts in area 1, 2, and 3, respectively, accounting for a total of 9650 unique expressed transcript (63%)
Melampsora larici-populina 98AG31 (Duplessis, Cuomo, et al., 2011)	Infected poplar leaves (Beaupré cv.) at 4 dpi, resting and germinating (3 h) uredospores	Custom whole-genome oligoarrays	13,093 genes assayed out of 16,399 genes included in the final genome annotation of isolate 98AG31	71% of the transcripts were detected in at least one situation; 6466 were expressed in all three situations; 7582, 7541, and 7656 transcripts were expressed at 4 dpi, in resting, and in germinating uredospores, respectively

Melampsora larici-populina 98AG31 (Duplessis, Haquard, et al., 2011)	Infected poplar leaves (Beaupré cv.) at 2, 6, 12, 24, 48, 96, and 168 hpi and comparison to resting and germinating (3 h) uredospores	Custom whole-genome oligoarrays	13,093 genes assayed out of 16,399 genes included in the final genome annotation of isolate 98AG31	<500 transcripts detected at 2, 6, or 12 hpi due to fungal transcript dilution; 4279, 6216, 7856, and 8326 transcripts detected at 24, 48, 96, and 168 hpi, respectively; 7735 and 7872 in resting and germinating uredospores, respectively, when normalized separately from *in planta* situations
Melampsora larici-populina 93ID6 and 98AG31 (Petre et al., 2012)	Infected poplar leaves (Beaupré cv.) at 18, 21, 24 hpi in incompatible interaction (isolate 93ID6) and at 18, 24, 48 hpi in compatible interaction (isolate 98AG31)	454 pyrosequencing GS-FLX Titanium; 713,505 reads	16,399 genes included in the final genome annotation of isolate 98AG31	90,398 contigs, from which only 649 were assigned to 280 fungal genes (isolate 98AG31 genome annotation)
Melampsora larici-populina 98AG31 (Hacquard, Delaruelle, et al., 2013)	Telia from early autumn-collected naturally infected poplar leaves (Beaupré cv.)	Custom whole-genome oligoarrays	13,093 genes assayed out of 16,399 genes included in the final genome	9588 transcripts expressed in telia, including 395 telia-specific transcripts

Continued

Table 6.2 Recent genome-wide expression studies of Pucciniales—cont'd

Species and isolate (reference)	Interaction, biological stage	Transcriptome approach	Number of genes covered	Number of genes detected
	comparison to resting and germinating uredospores as well as *in planta* 168 hpi sample from Duplessis, Haquard, et al. (2011)		annotation of isolate 98AG31	
Phakopsora pachyrhizi Thai1 (Link et al., 2014)	Purified haustoria from 12 dpi infected *Glycine max* leaves	454 pyrosequencing GS-FLX Titanium 1,051,753 reads	Unknown, no reference genome	11,872 assembled contigs, 4483 unique *P. pachyrhizi* contigs
Phakopsora pachyrhizi MS06-1 (Tremblay, Hosseini, Li, Alkharouf, & Matthews, 2012)	10 dpi infected *Glycine max* leaves	Illumina Genome Analyser II 5.96 million 36 bp reads	Unknown, no reference genome	~2.4 million reads assigned to *P. pachyrhizi*, 32,940 assembled *P. pachyrhizi* contigs
Phakopsora pachyrhizi MS06-1 (Tremblay, Hosseini, Li, Alkharouf, & Matthews, 2013)	15 spi, 7 hpi, 48 hpi, and 10 dpi in susceptible *Glycine max* leaves	Illumina Genome Analyser II 24.6 million single reads (from 3.5 to 9 million reads per time point)	Unknown, no reference genome	23–55% Reads were assigned to *P. pachyrhizi*; 6531, 4627, 4273, and 12,284 *de novo* assembled *P. pachyrhizi* transcripts at 15 spi, 7 hpi, 48 hpi, and 10 dpi, respectively, accounting for 27,715 fungal transcripts of which 19,000 represents new transcripts not previously reported

Puccinia graminis f. sp. *tritici* CDL75-36-700-3, race SCCL (Duplessis, Cuomo, et al., 2011)	Infected wheat leaves (cv. McNair 701) and infected barley leaves (cv. Hypana) at 7 and 8 dpi, respectively, and resting and germinating (24 h) uredospores	Custom whole-genome oligoarrays	20,228 putative genes defined ahead of the final genome annotation that comprises 17,773 genes	A total of 9818 transcripts were expressed, 6570 being expressed in all four situations
Puccinia striiformis f. sp. *tritici* isolate UK PST-08/21 (Cantu et al., 2013)	Infected wheat leaves (cv. Avocet 'S') at 6 and 14 dpi and purified haustoria at 7 dpi	66.7 and 200.4 million reads from infected leaves and haustoria, respectively, by Illumina-based RNA-seq Genome Analyser II (76 bp single reads)	19,073 predicted genes based on Cantu et al. (2011, 2013)	12 and 28.8 million reads from infected leaves and haustoria; comparison between haustoria and infected leaves, focus on secreted proteins, 57-induced/31-repressed in haustoria transcripts encoding secreted proteins; 411-induced and 333-repressed transcripts encoding non secreted proteins
Puccinia striiformis f. sp. *tritici* Pst-104E137A- (Garnica, Upadhyaya, Dodds, & Rathjen, 2013)	Purified haustoria and uredospores	454 pyrosequencing GS-FLX Titanium 729,036 (413 bp) and 457,071 (420 bp) reads and Illumina Genome	Unpublished reference genome of *P. striiformis* f. sp. *tritici* local isolate Pst-104E137A-	12,846 contigs in haustorial transcriptome, 12,282 assembled transcripts for combined haustoria and

Continued

Table 6.2 Recent genome-wide expression studies of Pucciniales—cont'd

Species and isolate (reference)	Interaction, biological stage	Transcriptome approach	Number of genes covered	Number of genes detected
		Analyser GX II 500 million of 100 bp paired-end reads		uredospores transcriptomes
Puccinia triticina, isolates MHDS, MLDS, MJBJ, TDBG, THBJ, and TNRJ (Bruce et al., 2013)	Infected wheat leaves (susceptible Thatcher cv.) at 6 dpi	Illumina 165 millions reads (60 bp, paired-end) in total, 26.4, 25.5, 23.4, 27.7, 33.2, 28.4 millions reads for isolates MHDS, MLDS, MJBJ, TDBG, THBJ, and TNRJ, respectively	Mapping to the unpublished reference genome V2 of *P. triticina* isolate BBBD race 1 (J. Fellers, C. Cuomo, L. Szabo, G. Bakkeren, B. McCallum, B. Saville, unpublished)	222,571 reads assigned to the fungus, focus on 1450 secreted proteins encoding transcripts in reference genome; 543 uniquely secreted protein-encoding transcripts identified
Uromyces appendiculatus SWBR1 (Link et al., 2014)	Purified haustoria from 10 dpi infected *Phaseolus vulgaris* leaves	454 pyrosequencing GS-FLX Titanium 894,873 reads	Unknown, no reference genome	14,581 assembled contigs, 7582 unique *U. appendiculatus* contigs

The table lists recently published transcriptome analyses of rust fungi using whole-genome oligoarrays or NGS.
spi, seconds postinoculation; hpi, hours postinoculation; dpi, days postinoculation.

transcriptomics studies delivered the first snapshots of genome-wide rust fungus gene expression. They particularly showed that about a third to a half of the genes were not expressed in uredospores or during the infection process, suggesting that they may have a role at other stages of the life cycle (Table 6.2). Among the genes expressed during host colonization (including uredospore germination) were a large number of genes of unknown function. This included rust-specific gene families, among which were many of the small secreted protein-encoding genes. The induced or upregulated expression of those secreted protein genes during host colonization strengthened their profile as rust candidate effectors with important functions during infection. This is particularly true for those that are specifically expressed only during the biotrophic growth *in planta*. Some of these putative effectors are significantly overexpressed at late stages of infection when the host tissues are showing an intense colonization by infection hyphae and haustoria, which is consistent with early molecular data obtained from purified haustoria of *U. fabae* or *M. lini* (Catanzariti et al., 2006; Hahn & Mendgen, 1997) and with recent data obtained by NGS approaches for *P. striiformis* f. sp. *tritici* (Cantu et al., 2013; Garnica et al., 2013), *P. pachyrhizi*, and *U. appendiculatus* (Link et al., 2014) by sequencing RNA isolated from purified haustoria. Interestingly, late stages of infection of poplar and wheat leaves by *M. larici-populina* and *P. graminis* f. sp. *tritici*, respectively, are marked by the expression of a large panel of genes related to the biotrophic growth such as CAZymes, transporters, proteases, and lipases (Duplessis, Cuomo, et al., 2011). A time-course infection study showed that most of the genes falling in the categories mentioned in earlier text were highly and significantly differentially expressed at late stages of infection. However, distinct sets of dozens to hundreds of small secreted protein-encoding genes were expressed sequentially, that is, they appeared to be expressed in distinct waves indicating that *M. larici-populina* possesses a highly dynamic secretome, which may be important for the interplay with components of the host plant immunity system (Duplessis, Haquard, et al., 2011; Hacquard et al., 2012). Such a dynamic pattern of expression *in planta* was further confirmed in other rust–plant interaction studies (Bruce et al., 2013; Cantu et al., 2013; Fernandez et al., 2012; Tremblay et al., 2013). Transcription profiles of fungal cell types were studied when the fungus releases uredospores from the host. At this stage, only the palisade mesophyll contains infection hyphae and haustoria, whereas the spongy mesophyll is filled with huge amounts of sporogenous hyphae and spores. The rust

fungus expressed very different genetic programs in these two plant compartments and the most highly expressed fungal genes in the palisade compared to the spongy mesophyll coded for candidate effectors, suggesting that these might play a role in the maintenance of infection structures *in planta* during the later stages of biotrophy (Hacquard et al., 2010). These gene expression profiles were compared with those obtained from telia of *M. larici-populina* harvested early in autumn, allowing a direct comparison of expression profiles in another spore-forming structure in the poplar leaf (Hacquard, Delaruelle, et al., 2013). A larger number of genes specifically expressed in telia were found, most of which encode unknown functions. Similarly, a large number of specific ESTs with unknown functions were identified in *P. triticina* teliospores (Xu et al., 2011) highlighting that the biological processes associated with this type of rust spores remain mostly uncharacterized. In *M. larici-populina*, telia and uredinia have the most similar expression profiles when compared to fungal hyphae undergoing biotrophic growth in poplar leaves, indicating shared components in the genetic programs ongoing in these spore-forming structures. Among the most highly regulated genes reported in telia are several with functions possibly related to spore survival, that is, overwintering, such as those encoding thaumatin-like proteins and aquaporins that may help to prevent osmotic damage due to desiccation. Also, several meiotic-related transcripts were overexpressed in teliospores and showed temporal patterns of expression during karyogamy, an important biological process occurring in teliospores (Hacquard, Delaruelle, et al., 2013). These studies demonstrate that the use of common custom oligoarrays is an efficient approach to realize transcriptomic comparisons to gain a better understanding of the biology of the rust fungus at different life cycle stages. However, this method still has pitfalls such as the detection of pathogen transcripts at stages containing very low amounts of fungal biomass inside the host (e.g. early stages of infection) (Duplessis, Haquard, et al., 2011).

3.2. RNA-Seq-based transcriptomics

Since 2011, considerable progress has been made by applying NGS to the study of several rust fungi, including *H. vastatrix*, *M. larici-populina*, *P. pachyrhizi*, *P. graminis* f. sp. *tritici*, *P. striiformis* f. sp. *tritici*, *P. triticina*, and *U. appendiculatus* (see Table 6.2). The possibility to compare different stages of infection in the telial hosts of rust fungi can help reveal patterns of gene expression during the infection process.

One challenge when studying different infection stages of a biotrophic fungus is the effect of dilution of the pathogen/host RNA. Within 48 h of infection of poplar by *M. larici-populina*, only 649 of the one million sequences generated by 454 pyrosequencing were assigned to rust genes (Table 6.2), representing less than 1% of the total sequences (Petre et al., 2012). Still, these contigs corresponded only to 280 unique *M. larici-populina* genes, half of them encoding small secreted proteins representing early-expressed candidate effectors at a stage when the first haustoria are recorded in the host (Laurans & Pilate, 1999). The level of infection could be an important aspect of the success in obtaining fungal transcripts. At a late stage of infection of *Coffea arabica* leaves by the coffee leaf rust fungus *H. vastatrix*, but before uredinia formation, the plant tissue is heavily colonized. About 30% of the total contigs (6763) assembled from 352,146 reads were attributed to the fungus based on a predictive comparative analysis to rust fungi and *C. arabica* sequences in databases (Table 6.2) (Fernandez et al., 2012). This nonmodel rust fungus lacks genomic information support, but based on the report of gene complements in rust fungi so far, it can be estimated that about a third of the *H. vastatrix* genes have been revealed by this transcriptomics approach. The study identified different cellular categories that can relate to the fungal growth *in planta*. The presence of 382 transcripts encoding small secreted proteins, among which a small set shown to be specifically expressed in haustoria and conserved among other rust fungi, indicates that this approach is also powerful to reveal putative candidate effectors in nonmodel rust fungi.

Similarly, time-course infection of soybean leaves colonized by *P. pachyrhizi* identified more than 4000 stage-specific transcripts at four different time points and in total 27,715 expressed fungal transcripts including 19,000 unique transcripts not previously recorded for rust fungi (Tremblay et al., 2013). Those transcript numbers suggest that some rust fungi with large estimated genomes like *P. pachyrhizi* may have larger gene complements than other rust fungi, or it could mean that the assembly parameters used in the corresponding study left numerous alternate transcripts ungrouped. The comparison of transcript expression profiles at the different stages of host colonization by *P. pachyrhizi* confirms the dynamic temporal regulation of gene expression also reported for *M. larici-populina*. What we have gleaned from the various studies is that the genetic programs expressed by rust fungi are finely regulated and this likely reflects different and/or specific processes occurring in the vastly different fungal structures formed during infection (i.e. germ tubes and appressoria at the leaf surface and infection

hyphae and haustoria in the leaf mesophyll) to produce uredospores at the leaf surface (i.e. spore-forming cells).

Application of NGS technologies to precisely dissect the infection process, that is, appressorium formation, direct penetration through cuticle for the soybean rust or other rust fungi when invading their aecial hosts, development of the substomatal vesicle and further infection hyphae and haustoria in the mesophyll, and then formation of uredospores and their release, still has limitations related to the depth of sequencing in order to reach a proper coverage of the transcriptome. For example, early steps in colonization of host tissue are impossible to capture at the moment, due to the small amount of fungal biomass in the collected host tissues. It is most likely that only a small fraction of the most highly expressed genes—that is, the tip of the iceberg—are detected as illustrated in the poplar rust fungus at the onset of haustoria formation (Petre et al., 2012). This issue also poses problems in terms of normalization between colonization stages. NGS-based fold-change levels calculated between stages in time-course studies should be considered cautiously wherever saturation is not reached in the cumulative curves for transcript coverage (see chapter by Kohler & Tisserant, 2014). However, even if the transcriptome is not complete, at comparable sequencing depth, it reflects a significant expression and may help to uncover genes commonly expressed between stages to unravel the common host infection toolkit of rust fungi. A problem with experimental rust fungus systems is a near impossibility to achieve synchronous infections to obtain sufficient material from a specific infection stage for RNA extraction. Single-cell analysis using micromanipulation techniques is however becoming feasible (Lin et al., 2014), and this may be applied to rust fungal pathosystems in the near future.

An alternative approach is to focus on specialized infection structures. Haustoria are crucial infection structures that are formed by rust fungi within host cells by breaching the cell wall but only penetrating beyond by invaginating the plasmalemma leaving the membrane intact. By isolating RNA from haustoria, it is therefore possible to enrich transcripts involved in host–pathogen interactions. Isolation of haustoria and recovery of RNA led to the description of key rust determinants involved in nutrient acquisition and delivery of effectors into host cells (Catanzariti et al., 2006; Hahn & Mendgen, 1997). RNA-Seq studies in rust fungi established the transcriptome profiles of purified haustoria in *P. pachyrhizi* and *U. appendiculatus* (Link et al., 2014) and in *P. striiformis* f. sp. *tritici* (Cantu et al., 2013; Garnica et al., 2013). In these studies, a particular focus was

on transcripts coding for predicted secreted proteins in the rust infection structure. Altogether, these studies show that a plethora of putative candidate effectors are expressed in those infection structures. In *P. striiformis* f. sp. *tritici*, fungal transcripts of secreted proteins representing promising candidate effectors of the wheat stripe rust fungus were identified by comparing leaf and haustoria transcript (Table 6.2) (Cantu et al., 2013). In *P. pachyrhizi* and *U. appendiculatus*, more than 11,000 contigs were generated for each species, and among these, 4483 and 7532 contigs were assigned to the two rust fungi, respectively (Link et al., 2014). Annotation of gene families and comparison to other fungi identified conserved families and specific families of secreted proteins. The candidate effector selection pipeline applied by Cantu et al. (2013) and previously defined to analyse *M. larici-populina* and *P. graminis* f. sp. *tritici* predicted genes (Saunders et al., 2012) and also identified conserved and specific tribes of haustorially expressed secreted protein gene families in *P. striiformis* f. sp. *tritici*.

It is important to note that some of these genes, although highly expressed in the haustorium, are not specific for this structure and many are also expressed in infection hyphae, uredospores, and other cell types of rust fungi in different host tissues, illustrating very dynamic patterns of gene regulation (Duplessis et al., 2012). Also, the detailed analysis of genes, overexpressed in *P. striiformis* f. sp. *tritici* haustoria and uredospores, showed major differences in the fungus prior and during infection (Garnica et al., 2013). Particularly, the two stages are highly contrasted for many cellular categories, including cell cycle, DNA metabolic and lipid metabolic processes, signal transduction in uredospores, generation of precursor metabolites and energy, translation, and vitamin and carbohydrate metabolic processes in haustoria. Such detailed analyses are highly valuable and forthcoming cross comparison between different rust species at different stages should help to identify the core components in the infection machinery of rust fungi and to decipher more precisely the specialized mode of acquisition of nutrients from the host.

A particularly exciting prospect is that of identified race-specific determinants in rusts. The combination in a dedicated pipeline of transcriptomic and genomic data of polymorphisms in gene sequences and their comparison to corresponding genes in other rust fungus isolates differing in resistance gene interaction helped to pinpoint the most promising candidate effectors for further functional characterization; these could be candidates with a role in the interplay with the host immune system (Cantu et al., 2013). In *P. triticina*, Illumina-based RNA-Seq was used to identify potential avirulent

genes by screening expression in distinct isolates representing six rust races (Bruce et al., 2013). In this case, the sequenced reads were compared to predicted transcripts in the *P. triticina* reference genome sequenced by the Broad Institute (J. Fellers, C.A. Cuomo, L. Szabo, G. Bakkeren, et al. unpublished data) and to a large collection of wheat ESTs. A particular focus was on small secreted protein-encoding genes. Among the 543 genes identified, 15 accumulated nonsynonymous mutations representing putative effector candidates that evolved under the pressure of the host immune system, and 11 showed positive correlation with reactions to multiple resistance genes in wheat. This pioneer study is the first step in the broad identification of determinants of interest for selection breeding programs (Bruce et al., 2013). It is envisaged that similar genome-wide associations can be made when comparing isolates differing in virulence (e.g. *P. striiformis* variants infecting differently at different temperatures) or even host range.

3.3. Comparison of transcriptome in different hosts

The comparison of transcriptomes during the interaction of a rust fungus with both its telial and aecial hosts promises to elucidate questions regarding the gene sets required for interacting with different hosts. The rust is confronted with very different host challenges in these different parts of its life cycle. In pine rusts, for example, monokaryotic spores infect the pine needles and grow within the host cambium and into the woody portion of the trees where it forms a hymenium. But dikaryotic aeciospores and uredospores infect the telial hosts and germinate and penetrate via stomata, followed by intercellular growth and formation of haustoria. It is likely that different genes are required for these interactions. Comparing infection of the two hosts of heteroecious rusts is complicated by the fact that it is not possible to conduct infection with the same genetic individual. The infection of the aecial host is via meiosis-derived basidiospores that are thus segregating (Doudrick, Nance, Nelson, Snow, & Hamelin, 1993), while infection of the telial host is via postspermatization dikaryotic aeciospores. Nevertheless, by using single-spore cultures derived from the telial host, it is possible to obtain genetic material that is relatively homogeneous. *Cronartium quercuum* f. sp. *fusiforme* transcriptomic profiles were strikingly different on the telial (oak) and aecial (loblolly pine) hosts. Half of the transcriptome sampled on a microarray was differentially expressed in these two hosts (J. Davis, et al., unpublished data). Similar outcomes were observed for the poplar rust fungus, when Illumina RNA-seq was applied

to infection stages in the aecial host, the larch (S. Duplessis, et al., unpublished data), and compared with expression during infection of the telial host, poplar (Duplessis, Haquard, et al., 2011). Particularly, a core set of candidate effectors were expressed in both hosts, whereas different sets were specifically expressed in each of the two hosts. Given the divergence between angiosperms and gymnosperms, it is likely that different sets of rust effectors interact with their respective targets in the telial and the aecial hosts. Rather different gene sets are also expressed in *P. triticina* when comparing Illumina RNA-seq-generated transcriptome profiles from infected wheat samples at 5 days after infection (dikaryotic uredospore-derived) and from pycnia and aecia (haploid and mated fungal cell mixtures) from the alternate host (J. Fellers, C. Cuomo, L. Szabo, G. Bakkeren, et al., unpublished data).

4. RUST SECRETOME, EFFECTORS, AND AVIRULENT GENES

There is a long history of studying host–pathogen interactions in rust pathosystems. The gene-for-gene hypothesis was tested and verified for the first time in the flax rust pathosystem (Flor, 1959). This field has been transformed recently by the discovery of a large repertoire of 'effectors' in plant pathogens. These are mostly proteins that are secreted by pathogens into the plant apoplast or into the host cells. Effectors are believed to manipulate host cell processes and facilitate infection by suppressing host defences, or by modifying host metabolism. Several of the avirulent genes that have been studied using classical genetic approaches are now demonstrated to be effectors recognized by cognate host resistant gene products (Bozkurt, Schornack, Banfield, & Kamoun, 2012; Dodds et al., 2009; Gan et al., 2010; Kamoun, 2007). This makes effector discovery and characterization extremely promising for host resistance screening and in monitoring of pathogen avirulence.

Since effectors are secreted, several pipelines have been developed to identify and characterize the secretome of plant pathogens. Such pipelines rely on computer-assisted identification of the presence of signal peptides, which are indicators that the proteins likely are secreted. Additional criteria often include size (putative effectors are usually small proteins) and being rich in cysteine (for folding and stability). Depending on the pathogen under study, other criteria can be added to the screening pipeline. In rust fungi, for example, additional criteria that have been proposed include expression *in planta*, homology to haustorial proteins, presence of internal repeats, long

intergenic regions, and absence of PFAM domains except those associated with pathogenicity (Hacquard et al., 2012; Saunders et al., 2012). Complete genome sequencing projects for rust fungi have revealed striking numbers of lineage-specific genes at the taxonomic level of either the order Pucciniales or the Melampsoraceae and Pucciniaceae families. Among those, several expanded gene families representing unknown functions encode secreted proteins and probably candidate rust effectors required to establish a successful infection in the plant host(s). In the three rust genomes sequenced and published so far, the secretome accounts for a large portion of the gene complement (~8%) encompassing nearly 2000 secreted proteins, and a majority are encoded by multigene families (Duplessis, Cuomo, et al., 2011; Zheng et al., 2013). Although the secretomes of the three rust fungi showed similar numbers, composition, and genomic organization, they greatly differ and exhibit a high specificity, reflecting a possible adaptation to the different hosts. The sequencing of more genomes in the order Pucciniales is much awaited to determine whether this assumption is correct. Also, comparative analyses between the secretomes of rust fungi have shown that a few genes and gene families are conserved across Pucciniales or even across Basidiomycota and showed an important diversity (Duplessis, Cuomo, et al., 2011; Hacquard et al., 2012; Saunders et al., 2012). However, up to now, only a few secreted proteins have been proven 'true effectors' (i.e. *M. lini* avirulent proteins and *U. fabae* RTP1; Catanzariti et al., 2006; Duplessis et al., 2012; Kemen et al., 2005), and it remains to be seen to which extent secretomes of rust fungi contribute to the expression of virulence.

In some cases, mapped populations segregating for avirulence can be used to map markers onto whole-genome sequences and identify avirulence. By taking advantage of the fact that haploid basidiospores of *Cronartium* spp. cause infections to pines, a bulk segregant analysis was used to map *Avr1* in *C. quercuum* f. sp. *fusiforme* (Kubisiak et al., 2011). These markers are now proving extremely useful. They were used to find the scaffold that comprises the putative *Avr1* gene in the *C. quercuum* f. sp. *fusiforme* genome (J. Davis et al., unpublished data). Combining this with a whole-genome outlier F_{st} analysis contrasting the bulk segregants, pycnial pools can provide a powerful approach to identify the *Avr1* locus. Among cereal-infecting rust fungi, 57 F2 progenies of a wheat leaf rust, *P. triticina*, sexual cross between two parental isolates (race 9 with virulence formula SBDG (Kolmer, 1996) and race 161 FBDJ) were sequenced using Illumina. At least, nine seedling and two adult plant avirulent genes were shown to segregate in this F2

population. In order to generate a high-resolution genetic linkage map, genome-wide single-nucleotide polymorphisms (SNPs) were identified by comparison to the *P. triticina* 1-1 reference genome. Using this information and employing the genome sequence information from the two parents and the F1 isolate, more than 25,000 SNPs were selected and used to generate a preliminary genetic linkage map. Correlation with the virulence profiles of the segregants yielded candidate avirulent genes, which are currently being investigated. In addition, although they were obtained from different isolates, genetic linkages allowed the creation of pseudomolecules, which were used to confirm physical linkage of many unassembled contigs in the reference genome (D.L. Joly, B.D. McCallum, B.S. Mulock, & G. Bakkeren, unpublished data). A similar effort is under way for a wheat stem rust, *P. graminis* f.sp. *tritici*, mapping population (Zambino, Kubelik, & Szabo, 2000).

5. POPULATION GENOMICS: FROM GENOMES TO LANDSCAPES

5.1. The rapidly evolving rust genomes

The ability of rust fungi to evolve and adapt has been widely reported. The appearance of rust races that can overcome resistance genes in crops and trees shortly after deployment has been one of the most enduring challenges for breeders (Kinloch, 2003; Kinloch, Sniezko, & Dupper, 2003; Kolmer, 2005; Pinon & Frey, 2005). There is a long history of characterizing rusts from a phenotypic and later from a population genetics perspective. However, the identification of the genes involved in adaptation has been challenging. The identification of genetic markers to track pathogen virulence has been complicated by the lack of adequate genetic maps, at the most yielding loose linkages representing large distances between genetic markers and virulence loci (Kubisiak et al., 2011; Zambino et al., 2000).

The genomic signature of this rapid evolution has yet to be elucidated in most pathogens. In some, the ability to evade host resistance response is related to genome structure. The *L. maculans* genome has a bipartite structure with AT-rich blocks containing effector genes and transposable elements that promote rapid diversification for evasion of host recognition (Rouxel et al., 2011). In *Phytophthora infestans*, gene-sparse, repeat-rich, and effector-rich regions that can rapidly generate new alleles and promote host jumps have been implicated (Raffaele et al., 2010). The genome sequences of rusts do not display such patterns although several *P. triticina*

BAC clones harboring effector homologues were surrounded by repeats and TE-like sequences (Fellers et al., 2013). However, there is evidence that some genome compartments are under positive selection and likely candidates in the host–pathogen arms race. By resequencing and comparing the genomes of multiple rust strains, it is possible to identify and map genomic regions where among-strain differentiation points to positive selection.

The study of population genome resequencing is in its infancy, but it can be expected to generate extremely useful results. Resequencing even a few isolates, especially of the dikaryotic phase, can be informative since two different nuclei are sampled and heterozygotes can be revealed. The genome of *P. striiformis* f. sp. *tritici* is highly heterozygous and possesses a diverse gene repertoire and more genes encoding secreted proteins compared with nonobligate fungal pathogens (Cantu et al., 2013; Zheng et al., 2013). The presence of a large number of genetic variants could be part of the explanation for the rapid evolution of this rust. The discovery that a large fraction of the SNPs are nonsynonymous and vary among regions could indicate regional adaptation. This high level of diversity in Asian *P. striiformis* f. sp. *tritici* could be interpreted as a result of sexual reproduction being important in generating regional races. The aecial host of this rust, barberry, is present in western China and central Asia, and therefore, completion of the sexual stage is possible and could drive this high level of diversity. However, sampling of more isolates in regions with and without aecial hosts would be required to compare observed and expected heterozygosity and determine if sexual reproduction is frequent or a rare occurrence.

An exciting prospect is to sequence and contrast rusts with different virulence profiles or host specificity. The comparison of genomes of four UK isolates of *P. striiformis* f. sp. *tritici* that differed in virulence to two wheat varieties highlighted the enrichment of nonsynonymous polymorphisms in haustoria-enriched proteins that were predicted to be secreted (Cantu et al., 2013). This allowed the narrowing of a complement of nearly 3000 predicted secreted proteins to five candidate effectors that differed between the isolates.

5.2. Host–pathogen adaptation in coevolved pathosystems

A better understanding of host–pathogen coevolution is important to devise better deployment strategies of resistance genotypes in crops and plantations. Antagonistic coevolution is believed to be a critical force driving the

evolution of diversity, and knowledge of how this balance is achieved in nature along spatial and temporal scales could be gained via studies of natural pathosystems (Thrall et al., 2012). Some of the most insightful studies on host–pathogen coevolution have been conducted in the flax rust *M. lini* and its wild host *Linum marginale* pathosystem. Variation in host resistance and pathogen infectivity is presumed to be the hallmark of a balanced wild pathosystem and to be essential for persistence of biotrophic pathogens (Laine, Burdon, Dodds, & Thrall, 2011). Although much has been learned by studying virulence phenotypes, there are exciting new studies examining natural variation in effector genes in natural populations.

Population genomics approaches in these nonmodel pathosystems are now feasible given the low cost of genome sequencing. Processes such as adaptation and evolution and the reconstruction of introduction and colonization events can be studied at the genome level. Evolutionary changes in the adaption of *M. larici-populina* to new hosts and new environments in northeastern North America were studied by resequencing 44 strains from the source and the introduced populations. Genome-wide distribution of polymorphisms displayed the signature of a severe founder event into the North American population (N. Feau et al., unpublished data). A genome scan of the four largest scaffolds revealed 32 regions having a strong divergence between the founder and the source populations. These regions include 14 secreted protein-coding genes that are candidate for adaptation (Feau, Frey, Duplessis, & Hamelin, 2013).

An advantage of studying *M. lini* in its wild pathosystem is that effector loci have been characterized functionally (Catanzariti et al., 2006, 2010; Dodds, Lawrence, Catanzariti, Ayliffe, & Ellis, 2004). Extensive sequence variation and an excess of nonsynonymous to synonymous mutations at two effector loci (AvrP123 and AvrP4) were found in the flax rust in its native range in Australia (Barrett et al., 2009). These loci therefore display the hallmark of positive selection. Alleles at these loci were demonstrated functionally to be differentially recognized by *R*-genes in the host. This result is consistent with the presence of variants of both effector genes that evolve to evade recognition by *R*-genes in the host in a classic antagonistic coevolutionary scenario (Barrett et al., 2009; Thrall et al., 2012).

It may also be informative to study evolution of the secretome in closely related rust species that differ in host specificity. The hosts are likely imposing evolutionary constraints acting more strongly on secreted proteins or effectors compared to the rest of the genome (Joly et al., 2010). More than half of the predicted proteins discovered in EST libraries of poplar rust

species specialized on different poplar hosts were lineage-specific. Secreted proteins had accelerated divergence and increased ratios of nonsynonymous to synonymous mutations when compared with nonsecreted proteins. In comparison of flax rusts attacking different hosts, the *avrP4* locus displayed 14 codons under significant positive selection. These mutations were present in the C-terminal 28-amino-acid region, suggesting that this region could be important in the interactions with host recognition proteins (Van der Merwe et al., 2009). Assessing allelic variation in putative effectors between distinct rust isolates will be important to identify and monitor avirulence alleles in populations and provide risk prediction and early warning for appearance of novel alleles. The rust *Melampsora allii-populina* shares poplar trees as the same telial hosts with *M. larici-populina*. The scheduled sequencing of this poplar rust fungus within the 1000 fungal genomes project at the JGI (http://genome.jgi.doe.gov/programs/fungi/index.jsf) will bring a great opportunity to compare at the genome scale the gene complements in the two genomes in order to determine the extent of specific genes related to poplar and to the distinct telial hosts (*Larix* spp. and *Allium* spp.). Whole-genome resequencing in those pathosystems promises to highlight additional genomic features associated with host specificity and host-pathogen equilibrium.

6. COMING UP NEXT IN RUST GENOMICS

The latest developments in rust fungal genomics have been providing exciting insights and tremendous amounts of new information increasing our understanding of the biology of these ecologically and economically relevant fungal pathogens. Since 2011, three genomic analyses and 14 genome-wide or NGS-based transcriptome studies have been published, and many more are expected in the next years. Pucciniales remains one of the orders with the lowest genome coverage in the fungal kingdom, and efforts are currently under way by the JGI and other international sequencing centres to address this paucity and add more species to this particularly important clade of pathogens (Grigoriev, 2014; Grigoriev et al., 2011, 2014). It is clear that the large sizes of rust genomes are challenging for such genome sequencing and assembly efforts. However, the most recent examples that were based on NGS technology to obtain draft or high-quality genome sequences are paving the way to future research in rust genomics to better understand the biology and ecology of these unique plant pathogens.

The different studies reported in this chapter have now provided a considerable level of information about the genomes, predicted secretomes, and correlated transcriptomes of different rust fungi. A detailed cross-species comparative analysis using these combined data is now the next challenge to precisely determine the extent of the core repertoire of rust effectors and the specificity of the secretome at different taxonomic levels. Correlating expression profiles in the various specific infection structures and comparing a variety of rust species should help to define a minimal set of functional effectors required for infection among the thousands of predicted secreted proteins. Although somewhat out of the scope of this review chapter, the generation of the genomic and transcriptomic data for rust fungi has made feasible interesting and complementary proteomic research. In particular, with respect to revealing potential effector suites produced in *P. triticina* haustoria, protein identification using generated proteomic spectra was greatly improved because of available comprehensive gene model predictions (Song et al., 2011). In this pathosystem, protein spectra derived from purified haustoria from various virulent and avirulent isolates are currently analysed against generated genomic resources from the corresponding isolates to identify specific effector variants (C. Rampitsch, B.D. McCallum, & G. Bakkeren, unpublished data). Similar approaches have led to successful effector identification in the barley powdery mildew fungus (Bindschedler et al., 2009; Godfrey, Zhang, Saalbach, & Thordal-Christensen, 2009) and in general, proteomic research greatly complements genomic and transcriptomic approaches, assists with genome annotations, and helps answer the various questions raised in this chapter (Tan, Ipcho, Trengove, Oliver, & Solomon, 2009).

We have highlighted some of the limitations that still hinder the analysis of important steps of the infection process by transcriptomics, such as during early stages of colonization *in planta*. Other stages are also lacking detailed information such as meiosis during teliospore germination, the colonization of the aecial host by basidiospores, and the subsequent mating during the sexual stage. NGS may help to capture relevant information about these obscure and sometimes ephemeral stages. However, this requires a very high sequencing depth, which at the moment still comes at a high cost. Micromanipulation and single-cell transcriptome analysis may soon be commonplace. Other future developments in sequencing technologies, particularly the possibility to sequence longer DNA fragments and improvements of the computer programs used to assemble short sequences, may help to cope with the richness in repetitive elements in the rust genomes. Further than the

rust transcriptome per se, another advantage of using a NGS RNA-seq approach is the possibility to realize dual transcriptome sequencing, capturing both the host and the pathogen transcript expression profiles. This approach has so far been poorly explored, probably due to the lack of assembled and annotated plant host genomes. The availability of the poplar and wheat genome sequences will allow for such kind of analyses with *M. larici-populina* and the *Puccinia* cereal rust fungi (Petre et al., 2012). But even in the absence of a known host genome, it is possible to collect such information on the host based on the availability of extensive plant and fungal genome databases to ensure proper transcript designation in mixed infections (Fernandez et al., 2012). Beyond the identification of rust disease determinants during host infection, this will help to unravel the functions and the pathways that are triggered by the rust effectors to achieve proper biotrophic growth in the case of compatible plant–rust interactions and components of rust resistance and defence in incompatible interactions.

With the actual sequencing depth of NGS, it is now possible to pool DNA of various isolates and gather very large amounts of genome sequences that can be compared to a reference genome. This opens great perspectives for sampling and resequencing rust isolates from different landscapes and different hosts, across altitude and latitude gradients to identify adaptive genome regions. Population genomics approaches will allow the identification of genomic regions subjected to positive and purifying selection to target loci responsible for major traits such as virulence. The inclusion of very large sampling in populations will allow the comparison of rusts with different demographic histories (bottleneck and expansions) and the identification of genome regions under selective sweeps. Genome resequencing is still in its infancy for rust fungi, but the first reports already indicate a great genomic variability, which is a good basis to conduct population genomics of rust fungi.

ACKNOWLEDGEMENTS

S. D. would like to acknowledge his colleagues Pascal Frey, Fabien Halkett, and Stéphane De Mita at INRA Nancy for fruitful discussions on the *Populus–Melampsora* pathosystem; the ANR for supporting rust genomics projects in the frame of the young scientist grant POPRUST (ANR-2010-JCJC-1709-01) and the 'Investissements d'Avenir' program (ANR-11-LABX-0002-01, Lab of Excellence ARBRE); and the Joint Genome Institute (Office of Science of the U.S. Department of Energy under contract no. DE-Ac02-05cH11231) for the sequencing of the genome of the poplar rust fungus *Melampsora larici-populina*.

G. B. acknowledges the collaborations with Christina Cuomo, John Fellers, Les Szabo, Jim Kolmer, Brent McCallum, and Barry Saville and the sequencing work performed by the Broad Institute, Cambridge, MA (funded through grants from NSF-CSREES/USDA), and the Michael Smith Genome Sciences Centre, Vancouver, BC (funded through grants from Genome BC and the Ontario Research Fund), and the research funds from AAFC.

R.H. acknowledges the collaborations with John Davis, Nicolas Feau, Braham Dhillon, Amanda Pendleton and Igor Gregoriev and Genome BC's support through Strategic Opportunity Funds #131 and Genome Canada through the Large Scale Applied Genomic Program project #164.

REFERENCES

Aime, M. C., Matheny, P. B., Henk, D. A., Frieders, E. M., Nilsson, R. H., Piepenbring, M., et al. (2006). An overview of the higher level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. *Mycologia*, *98*, 896–905.

Barrett, L. G., Thrall, P. H., Dodds, P. N., van der Merwe, M., Linde, C. C., Lawrence, G. J., et al. (2009). Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. *Molecular Biology and Evolution*, *26*, 2499–2513.

Bindschedler, L. V., Burgis, T. A., Mills, D. J., Ho, J. T., Cramer, R., & Spanu, P. D. (2009). In planta proteomics and proteogenomics of the biotrophic barley fungal pathogen *Blumeria graminis* f. sp. *hordei*. *Molecular and Cellular Proteomics*, *8*, 2368–2381.

Bozkurt, T. O., Schornack, S., Banfield, M. J., & Kamoun, S. (2012). Oomycetes, effectors, and all that jazz. *Current Opinion in Plant Biology*, *15*, 483–492.

Bruce, M., Neugebauer, K. A., Joly, D. L., Migeon, P., Cuomo, C. A., Wang, S., et al. (2013). Using transcription of six *Puccinia triticina* races to identify the effective secretome during infection of wheat. *Frontiers in Plant Science*, *4*, 520.

Cantu, D., Govindarajulu, M., Kozik, A., Wang, M., Chen, X., Kojima, K. K., et al. (2011). Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. *PLoS One*, *6*, e24230.

Cantu, D., Segovia, V., MacLean, D., Bayles, R., Chen, X., Kamoun, S., et al. (2013). Genome analyses of the wheat yellow (stripe) rust pathogen *Puccinia striiformis* f. sp. *tritici* reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. *BMC Genomics*, *14*, 270.

Carvalho, G. M. A., Carvalho, C. R., Barreto, R. W., & Evans, H. C. (2014). Coffee rust genome measured using flow cytometry: Does size matter? *Plant Pathology*, http://dx.doi.org/10.1111/ppa.12175.

Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A., & Ellis, J. G. (2006). Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell*, *18*, 243–256.

Catanzariti, A. M., Dodds, P. N., Ve, T., Kobe, B., Ellis, J. G., & Staskawicz, B. J. (2010). The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein. *Molecular Plant-Microbe Interactions*, *23*, 49–57.

Cummins, G. B., & Hiratsuka, Y. (2003). *Illustrated genera of rust fungi* (3rd ed.). St Paul, MN: American Phytopathological Society.

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Ayliffe, M. A., & Ellis, J. G. (2004). The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell*, *16*, 755–768.

Dodds, P. N., Rafiqi, M., Gan, P. H., Hardham, A. R., Jones, D. A., & Ellis, J. G. (2009). Effectors of biotrophic fungi and oomycetes: Pathogenicity factors and triggers of host resistance. *New Phytologist*, *183*, 993–1000.

Doudrick, R., Nance, W., Nelson, C., Snow, G., & Hamelin, R. (1993). Detection of DNA polymorphisms in a single urediniospore-derived culture of *Cronartium quercuum* f. sp. *fusiforme*. *Phytopathology*, *83*, 388–392.

Duplessis, S., Cuomo, C. A., Lin, Y. C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., et al. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 9166–9171.

Duplessis, S., Haquard, S., Delaruelle, C., Tisserant, E., Frey, P., Martin, F., et al. (2011). *Melampsora larici-populina* transcript profiling during germination and time-course infection of poplar leaves reveals dynamic expression patterns associated with virulence and biotrophy. *Molecular Plant-Microbe Interactions*, *24*, 808–818.

Duplessis, S., Joly, D. L., & Dodds, P. N. (2012). Rust effectors. In F. Martin & S. Kamoun (Eds.), *Effectors in plant–microbes interactions* (pp. 155–193). Oxford, UK: Wiley-Blackwell, Chapter 7.

Duplessis, S., Spanu, P. D., & Schirawski, J. (2014). Biotrophic fungi (powdery mildews, Rusts and Smuts). In F. Martin (Ed.), *Ecological genomics of the fungi* (pp. 149–168). Oxford, UK: Wiley-Blackwell.

Feau, N., Frey, P., Duplessis, S., & Hamelin, R. C. (2013). Genome-wide signature of adaptation in a recently introduced pathogen. *Phytopathology*, *103*, 42.

Fellers, J., Soltani, B., Bruce, M., Linning, R., Cuomo, C., Szabo, L., et al. (2013). Conserved loci of leaf and stem rust fungi of wheat share synteny interrupted by lineage-specific influx of repeat elements. *BMC Genomics*, *14*, 60.

Fernandez, D., Talhinhas, P., & Duplessis, S. (2013). Rust fungi achievements and future challenges on genomics and host-parasite interactions. In F. Kempken (Ed.), *The Mycota: Vol XI, Agricultural applications* (pp. 315–342). (2nd ed.). Heidelberg, Germany: Springer.

Fernandez, D., Tisserant, E., Talhinhas, P., Azinheira, H., Vieira, A., Petitot, A. S., et al. (2012). 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta expressed pathogen-secreted proteins and plant functions expressed in a late compatible plant-rust interaction. *Molecular Plant Pathology*, *13*, 17–37.

Flor, H. H. (1959). Differential host range of the monocaryon and the dicaryons of a eu-autoecious rust. *Phytopathology*, *49*, 794–795.

Gan, P. H., Rafiqi, M., Ellis, J. G., Jones, D. A., Hardham, A. R., & Dodds, P. N. (2010). Lipid binding activities of flax rust AvrM and AvrL567 effectors. *Plant Signaling and Behaviour*, *5*, 1272–1275.

Garnica, D. P., Upadhyaya, N. M., Dodds, P. N., & Rathjen, J. P. (2013). Strategies for wheat stripe rust pathogenicity identified by transcriptome sequencing. *PLoS One*, *8*, e67150.

Godfrey, D., Zhang, Z., Saalbach, G., & Thordal-Christensen, H. (2009). A proteomics study of barley powdery mildew haustoria. *Proteomics*, *9*, 3222–3232.

Grigoriev, I. V. (2014). A changing landscape of fungal genomics. In F. Martin (Ed.), *Ecological genomics of the fungi* (pp. 3–20). Oxford, UK: Wiley-Blackwell.

Grigoriev, I. V., Cullen, D., Goodwin, S. B., Hibbett, D., Jeffries, T. W., Kubicek, C. P., et al. (2011). Fueling the future with fungal genomics. *Mycology*, *2*, 192–209.

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., et al. (2014). MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Research*, *42*, D699–D704.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, *461*, 393–398.

Hacquard, S., Delaruelle, C., Frey, P., Tisserant, E., Kohler, A., & Duplessis, S. (2013). Transcriptome analysis of poplar rust telia reveals overwintering adaptation and tightly coordinated karyogamy and meiosis processes. *Frontiers in Plant Science*, *4*, 456.

Hacquard, S., Delaruelle, C., Legue, V., Tisserant, E., Kohler, A., Frey, P., et al. (2010). Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Molecular Plant-Microbe Interactions*, *23*, 1275–1286.

Hacquard, S., Joly, D. L., Lin, Y. C., Tisserant, E., Feau, N., Delaruelle, C., et al. (2012). A comprehensive analysis of genes encoding small secreted proteins identifies candidate effectors in *Melampsora larici-populina* (poplar leaf rust). *Molecular Plant-Microbe Interactions*, *25*, 279–293.

Hacquard, S., Kracher, B., Maekawa, T., Vernaldi, S., Schulze-Lefert, P., Loren, Ver, et al. (2013). Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, E2219–E2228.

Hahn, M., & Mendgen, K. (1997). Characterization of in planta-induced rust genes isolated from a haustorium-specific cDNA library. *Molecular Plant-Microbe Interactions*, *10*, 427–437.

Joly, D. L., Feau, N., Tanguay, P., & Hamelin, R. C. (2010). Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (*Melampsora* spp.). *BMC Genomics*, *11*, 422.

Kamoun, S. (2007). Groovy times: Filamentous pathogen effectors revealed. *Current Opinion in Plant Biology*, *10*, 358–365.

Kämper, J., Kahmann, R., Bölker, M., Ma, L. J., Brefort, T., Saville, B. J., et al. (2006). Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, *444*, 97–101.

Kemen, E., & Jones, J. D. (2012). Obligate biotroph parasitism: Can we link genomes to lifestyles? *Trends in Plant Science*, *17*, 448–457.

Kemen, E., Kemen, A. C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., et al. (2005). Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Molecular Plant-Microbe Interactions*, *18*, 1130–1139.

Kinloch, B. B. (2003). White pine blister rust in north america: Past and prognosis. *Phytopathology*, *93*, 1044–1047.

Kinloch, B., Jr., Sniezko, R., & Dupper, G. (2003). Origin and distribution of Cr2, a gene for resistance to white pine blister rust in natural populations of western white pine. *Phytopathology*, *93*, 691–694.

Kohler, A., & Tisserant, E. (2014). Exploring the transcriptome of mycorrhizal interactions. *Advances in Botanical Research*, *70*.

Kolmer, J. A. (1996). Genetics of resistance to wheat leaf rust. *Annual Review of Phytopathology*, *34*, 435–455.

Kolmer, J. A. (2005). Tracking wheat rust on a continental scale. *Current Opinion in Plant Biology*, *8*, 441–449.

Kubisiak, T. L., Anderson, C. L., Amerson, H. V., Smith, J. A., Davis, J. M., & Nelson, C. D. (2011). A genomic map enriched for markers linked to Avr1 in *Cronartium quercuum* f. sp. *fusiforme*. *Fungal Genetics and Biology*, *48*, 266–274.

Laine, A. L., Burdon, J. J., Dodds, P. N., & Thrall, P. H. (2011). Spatial variation in disease resistance: From molecules to metapopulations. *Journal of Ecology*, *99*, 96–112.

Laurans, F., & Pilate, G. (1999). Histological aspects of a hypersensitive response in poplar to *Melampsora larici-populina*. *Phytopathology*, *89*, 233–238.

Laurie, J. D., Ali, S., Linning, R., Mannhaupt, G., Wong, P., Güldener, U., et al. (2012). Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. *Plant Cell*, *24*, 1733–1745.

Lin, K., Limpens, E., Zhang, Z., Ivanov, S., Saunders, D. G., Mu, D., et al. (2014). Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genetics*, *10*, e1004078.

Link, T. I., Lang, P., Scheffler, B. E., Duke, M. V., Graham, M. A., Cooper, B., et al. (2014). The haustorial transcriptomes of Uromyces appendiculatus and Phakopsora pachyrhizi and their candidate effector families. *Molecular Plant Pathology*, http://dx.doi.org/10.1111/mpp.12099.

Link, T. I., & Voegele, R. T. (2008). Secreted proteins of *Uromyces fabae*: Similarities and stage specificity. *Molecular Plant Pathology*, *9*, 59–66.

Lowe, R. G., & Howlett, B. J. (2012). Indifferent, affectionate, or deceitful: Lifestyles and secretomes of fungi. *PLoS Pathogens*, *8*, e1002515.

Martin, F. (2014). *The ecological genomics of fungi*. Oxford, UK: Wiley-Blackwell.

Martin, F., Aerts, A., Ahrén, D., Brun, A., Danchin, E. G., Duchaussoy, F., et al. (2008). The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*, *452*, 88–92.

Martin, F., Kohler, A., Murat, C., Balestrini, R., Coutinho, P. M., Jaillon, O., et al. (2010). Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature*, *464*, 1033–1038.

McDowell, J. M. (2011). Genomes of obligate plant pathogens reveal adaptations for obligate parasitism. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 8921–8922.

Mendgen, K. W., & Hahn, M. (2002). Plant infection and the establishment of fungal biotrophy. *Trends in Plant Science*, *7*, 352–356.

Murat, C., Payen, T., Petitpierre, D., & Labbé, J. (2014). Repeated elements in filamentous fungi with a focus on wood-decay fungi. In F. Martin (Ed.), *Ecological genomics of the fungi* (pp. 21–40). Oxford, UK: Wiley-Blackwell.

Petre, B., Morin, E., Tisserant, E., Hacquard, S., Da Silva, C., Poulain, J., et al. (2012). RNA-Seq of early-infected poplar leaves by the rust pathogen *Melampsora larici-populina* uncovers PtSultr3;5, a fungal-induced host sulfate transporter. *PLoS One*, *7*, e44408.

Pinon, J., & Frey, P. (2005). Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In M. H. Pei, & A. R. McCracken (Eds.), *Rust diseases of willow and poplar* (pp. 139–154). Wallingford, UK: CAB International.

Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., et al. (2010). Genome evolution following host jumps in the irish potato famine pathogen lineage. *Science*, *330*, 1540–1543.

Raffaele, S., & Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: Why bigger can be better. *Nature Reviews. Microbiology*, *10*, 417–430.

Rouxel, T., Grandaubert, J., Hane, J. K., Hoede, C., van de Wouw, A. P., Couloux, A., et al. (2011). Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature Communications*, *2*, 202.

Saunders, D. G. O., Win, J., Cano, L. M., Szabo, L. J., Kamoun, S., & Raffaele, S. (2012). Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PLoS One*, *7*, e29847.

Schirawski, J., Mannhaupt, G., Münch, K., Brefort, T., Schipper, K., Doehlemann, G., et al. (2010). Pathogenicity determinants in smut fungi revealed by genome comparison. *Science*, *330*, 1546–1548.

Song, X., Rampitsch, C., Soltani, B., Mauthe, W., Linning, R., Banks, T., et al. (2011). Proteome analysis of wheat leaf rust fungus, *Puccinia triticina*, infection structures enriched for haustoria. *Proteomics*, *11*, 944–963.

Spanu, P. D. (2012). The genomics of obligate (and nonobligate) biotrophs. *Annual Review of Phytopathology*, *50*, 91–109.

Spanu, P. D., Abbott, J. C., Amselem, J., Burgis, T. A., Soanes, D. M., Stüber, K., et al. (2010). Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science*, *330*, 1543–1546.

Tan, K. C., Ipcho, S. V., Trengove, R. D., Oliver, R. P., & Solomon, P. S. (2009). Assessing the impact of transcriptomics, proteomics and metabolomics on fungal phytopathology. *Molecular Plant Pathology*, *10*, 703–715.

Thrall, P. H., Laine, A. L., Ravensdale, M., Nemri, A., Dodds, P. N., Barrett, L. G., et al. (2012). Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*, *15*, 425–435.

Toome, M., Ohm, R. A., Riley, R. W., James, T. Y., Lazarus, K. L., Henrissat, B., et al. (2014). Genome sequencing provides insight into the reproductive biology, nutritional mode and ploidy of the fern pathogen Mixia osmundae. *New Phytologist,* (in press).

Tremblay, A., Hosseini, P., Li, S., Alkharouf, N. W., & Matthews, B. F. (2012). Identification of genes expressed by *Phakopsora pachyrhizi*, the pathogen causing soybean rust, at a late stage of infection of susceptible soybean leaves. *Plant Pathology*, *61*, 773–786.

Tremblay, A., Hosseini, P., Li, S., Alkharouf, N. W., & Matthews, B. F. (2013). Analysis of *Phakopsora pachyrhizi* transcript abundance in critical pathways at four time-points during infection of a susceptible soybean cultivar using deep sequencing. *BMC Genomics*, *14*, 614.

Van der Merwe, M. M., Kinnear, M. W., Barrett, L. G., Dodds, P. N., Ericson, L., Thrall, P. H., et al. (2009). Positive selection in AvrP4 avirulence gene homologues across the genus *Melampsora*. *Proceedings. Biological sciences/The Royal Society*, *276*, 2913–2922.

Voegele, R. T., Hahn, M., & Mendgen, K. (2009). The uredinales: Cytology, biochemistry, and molecular biology. In K. Esser, & H. B. Deising (Eds.), *The Mycota. A comprehensive treatise on fungi as experimental systems for basic and applied research* (pp. 69–98). Heidelberg, Germany: Springer.

Voegele, R. T., Struck, C., Hahn, M., & Mendgen, K. (2001). The role of haustoria in sugar supply during infection of broad bean by the rust fungus Uromyces fabae. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 8133–8138.

Warren, J. M., & Covert, S. F. (2004). Differential expression of pine and *Cronartium quercuum* f. sp. *fusiforme* genes in fusiform rust galls. *Applied and Environmental Microbiology*, *70*, 441–451.

Wicker, T., Oberhaensli, S., Parlange, F., Buchmann, J. P., Shatalina, M., Roffler, S., et al. (2013). The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nature Genetics*, *45*, 1092–1096.

Xu, J., Linning, R., Fellers, J., Dickinson, M., Zhu, W., Antonov, I., et al. (2011). Gene discovery in EST sequences from the wheat leaf rust fungus *Puccinia triticina* sexual spores, asexual spores and haustoria, compared to other rust and corn smut fungi. *BMC Genomics*, *12*, 161.

Zambino, P. J., Kubelik, A. R., & Szabo, L. J. (2000). Gene action and linkage of avirulence genes to DNA markers in the rust fungus Puccinia graminis. *Phytopathology*, *90*, 819–826.

Zheng, W., Huang, L., Huang, J., Wang, X., Chen, X., Zhao, J., et al. (2013). High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. *Nature Communications*, *4*, 1–10.