Title: Genomics and evolution of Pneumocystis species.

Authors: Cissé OH, Hauser PM

Journal: Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases

Year: 2018 Nov

Issue: 65

Pages: 308-320

DOI: 10.1016/j.meegid.2018.08.015

HIGHLIGHTS

- Advances in genomics of *Pneumocystis* species unlocked new areas of research
- Slow genome decay and limited expansions of specific gene families and introns
- Adaptation influenced by self-fertility, host specificity, and transmission mode
- Establishment of culture *in vitro* needed to unravel the forces driving evolution

Revised version: cleaned.

Genomics and evolution of *Pneumocystis* species

Ousmane H. Cissé [1] and Philippe M. Hauser [2]

Correspondence: ousmane.cisse@nih.gov and Philippe.Hauser@chuv.ch

[1] Critical Care Medicine Department, NIH Clinical Center, National Institutes of Health, Bethesda, Maryland 20892, USA

[2] Institute of Microbiology, Lausanne University Hospital, Lausanne, Switzerland

TABLE OF CONTENTS

ABSTRACT

The genus *Pneumocystis* comprises highly diversified fungal species that cause severe pneumonia in individuals with a deficient immune system. These fungi infect exclusively mammals and present a strict host species specificity. These species have co-diverged with their hosts for long periods of time (> 100 MYA). Details of their biology and evolution are fragmentary mainly because of a lack of an established long-term culture system. Recent genomic advances have unlocked new areas of research and allow new hypotheses to be tested. We review here new findings of the genomic studies in relation with the evolutionary trajectory of these fungi and discuss the impact of genomic data analysis in the context of the population genetics. The combination of slow genome decay and limited expansion of specific gene families and introns reflect intimate interactions of these species with their hosts. The evolutionary adaptation of these organisms is profoundly influenced by their population structure, which in turn is determined by intrinsic features such as their self-fertilizing mating system, high host specificity, long generation times, and transmission mode. Essential key questions concerning their adaptation and speciation remain to be answered. The next cornerstone will consist in the establishment of a long-term culture system and genetic manipulation that should allow unravelling the driving forces of *Pneumocystis* species evolution.

Main text : 7482 words

BACKGROUND

History

Pneumocystis species form a group of opportunistic fungi that cause severe pulmonary infections in mammals with a deficient immune system. These organisms infect exclusively mammals. They were first described by Chagas (Chagas, 1909), and wrongly classified as special forms of trypanosomes. They were later identified as a *bona fide* separate species by the Delanoë couple at the Pasteur Institute in Paris (Delanoë and Delanoë, 1912). Their taxonomic classification remained then elusive because of a phenotypic resemblance with the protists. The issue was resolved using molecular phylogeny based on sequencing ribosomal DNA, which clearly indicated their fungal nature (Edman et al., 1988).

Phylogeny and taxonomy

Pneumocystis species belong to the subphylum of Taphrinomycotina within the Ascomycota (Eriksson, 1997; Sugiyama et al., 2006). The Taphrinomycotina subphylum is monophyletic and encompasses mostly plant-associated or soil-dwelling fungi (Liu et al., 2009). *Pneumocystis* closest relatives are *Schizosaccharomyces pombe* and *Taphrina deformans*, their common ancestor having diverged from the other Taphrinomycota members ca. 467 million years ago (MYA) (Beimforde et al., 2014).

Although all *Pneumocystis* species are ubiquitous, each mammal species can be infected with only one or two of them. Five species have been formally described so far based on the requirements of the International Code of Botanical Nomenclature (ICBN): *Pneumocystis*

jirovecii in *Homo sapiens* (Frenkel, 1999), *Pneumocystis carinii* in *Rattus norvegicus* (Frenkel, 1999), *Pneumocystis wakefieldiae* also in *Rattus norvegicus* (Cushion et al., 1993; Cushion et al., 2004), *Pneumocystis murina* in *Mus musculus* (Keely et al., 2004a), and *Pneumocystis oryctologi* in Old World rabbits (*Oryctolagus cuniculus*; Dei-cas et al., 2006). Antigenic and DNA based studies suggest the presence of distinct species also in macaques, ferrets, bats, shrews, horses, pigs, and dogs (Banerji et al., 1994; Peters et al., 1994; Christensen et al., 1996; English et al., 2001; Guillot et al., 2004).

P. jirovecii is the only species known to infect humans and has never been detected in any other animals. *P. carinii* is the best studied species because of the availability of protocols for experimental or natural infections in laboratory rats. *P. wakefieldiae* was reported either mixed with *P. carinii* (Cushion et al., 1993; Cushion, 1998; Cushion et al., 2004; Chabé et al., 2010), or alone (Palmer et al., 2000). The two latter species are different in terms of electrophoretic karyotypes, gene localization on the chromosomes, sequence identity (4-7% nucleotide divergence in seven orthologs; Cushion, 1998; Cushion et al., 2004), antigenic profiles (Vasquez et al., 1996), and major surface glycoproteins (MSG) expression (Schaffzin and Stringer, 2000). They might be competing against each other for resources when present together within the same rat (Icenhour et al., 2006a).

Species divergence

According to the evolutionary rates of several genomic loci, the radiation of the *Pneumocystis* genus occurred ca.100 MYA (Keely et al., 2003a; Keely et al., 2004a), which roughly overlaps with the radiation of the mammalian species (Holmes, 1991; dos Reis et al., 2015). *P. murina* would have diverged from *P. carinii* between 51 and 71 MYA (Keely et al., 2003a), while *P.*

carinii and *P. wakefieldiae* diverged between 15 and 22 MYA (Cushion et al., 2004; Fischer et al., 2006). The neat superposition of multiple *Pneumocystis* species phylogenetic trees with those of their respective hosts supports a co-evolution of these organisms (Guillot et al., 2001). Therefore, a plausible co-speciation scenario is that each species became physically separated from the other species, the hosts acting as barriers that led to the accumulation of genetic differences and the gradual reproductive isolation over time. The absence of gene flow or mating among the different species has been inferred based on linkage disequilibrium analysis consistent with an ancient reproductive isolation (Mazars et al., 1997; Keely et al., 2004a; Keely and Stringer, 2009). Furthermore, no evidence of hybridization was detected between *P. carinii* and *P. wakefieldiae*, even during co-infection of the same rat (Cushion, 1998; Cushion et al., 2004). However, caution is warranted because the absence of gene flow was inferred from a small set of conserved markers, which may have not allowed detecting all genetic events. Consequently, whole genome sequencing studies are necessary to validate these findings.

Life cycle

The life cycle of *Pneumocystis* organisms is still hypothetical and mostly derived from microscopic and molecular studies on *P. carinii* (Figure 1). As fungal organisms with an obligate parasitic behavior, the cycle would occur only inside host's lungs, and begin with the inhalation of infectious asci. Once inhaled, each ascus would release first eight ascospores which will evolve to what is known as trophic forms that bind to the type I pneumocytes of the alveolar epithelium. The cycle would then alternate between asexual multiplication of metabolically active trophic cells by binary fission, and sexual reproduction upon mating of two trophic cells that would culminate by the production of asci containing eight ascospores (Figure 2). Trophic

cells are amoeboid in shape and represent generally 90-98% of the populations in the infected lungs (Aliouat-Denis et al., 2009). These forms are mononuclear, 2-8 μm in diameter (Dei-Cas et al., 2004), and mostly haploid (Stringer and Cushion, 1998; Wyder et al., 1998; Martinez et al., 2011). Multiploid forms are rare and possibly caused by asymmetrical or post-mating divisions (Martinez et al., 2011). Trophic cell surface is composed of a single layer of electron dense material containing glycoproteins, but possibly no β-glucans. Indeed, the enzymes responsible for the synthesis of β-glucans and the associated endo-1,3-β glucanase are expressed almost exclusively in asci (Nollstadt et al., 1994; Kottom and Limper, 2000; Kutty et al., 2015). The presence of structural carbohydrate polymers of glucans in asci increases the physical strength of the cell wall, which might facilitate the survival outside the host. The doubling times are relatively long compared to free-living yeasts (~2 hours) and range from 1.5 to 10.5 days depending on the species (Aliouat et al., 1999; Keely et al., 2003b). The presence of a sexual cycle was initially supported by the ultrastructural observations of synaptonemal complexes (Matsumoto and Yoshida, 1984) and the expression of one pheromone receptor at the surface of *P. carinii* trophic cells (Vohra et al., 2004). Recent comparative genomic studies suggest that *Pneumocystis* species use primary homothallism (self-fertility) based on the genes number and arrangement on the chromosomes as a fusion of Plus and Minus mating type loci (Almeida et al., 2015). Thus, each strain would be able to produce asci on its own, without the need to find a compatible partner. Asci would be expelled by infected hosts and be the infectious stages because their specific inactivation or removal blocks the transmission chain (Cushion et al., 2010; Martinez et al., 2013). Consistently, recent analyses suggested that *Pneumocystis* sexuality is obligatory within host's lungs in order to complete the cell cycle and produce asci that are necessary for airborne transmission to new hosts (Richard et al., 2018). Furthermore, the

necessity of asci for transmission has been demonstrated by inhibition of the sexual cycle using echinocandins (Cushion et al., 2010), and by the fact that only purified asci could transmit the disease (Martinez et al., 2013). Recently, activation of sex-related genes upon treatment with echinocandins in RNA-seq analyses also suggested that sexuality is obligate (Cushion et al., 2018).

Transmission

Pneumocystis jirovecii pneumonia is a major public health problem with >400,000 cases per year worldwide and a mortality rate possibly as high as 80% if untreated (Brown et al., 2012). Epidemiological data for *Pneumocystis* species in animal populations are scarce, but investigations in shrews and rats suggest a pervasive low level of infections (Laakkonen, 1998; Chabé et al., 2010).

Pneumocystis organisms are transmitted via the air from infected individuals to new hosts (Hughes, 1982), including between individuals within hospitals (de Boer et al., 2011), but also possibly via the transplacental route (Ceré et al., 1997; Sanchez et al., 2007; Montes-Cano et al., 2009). The current hypothesis is that infections occur over short distance among infected and susceptible individuals (Chabé et al., 2011). The transfer of parasites from animals to humans is no longer considered as a valid hypothesis based on the strict host species specificity (Chabé et al., 2011). Consistently, no convincing evidence of an environmental source of *Pneumocystis* has been found so far, which strongly suggests that mammals constitute the only reservoir of these fungi. Furthermore, the erosion of metabolic capabilities evidenced by the genome sequencing studies suggests that these organisms are unable to live outside their hosts (see below, losses of metabolic machinery section). Finally, they apparently complete their whole cell cycle within host's lungs since sexuality occurs therein. Healthy infected hosts colonized by the organism are believed to contribute greatly to the transmission and circulation process (Chabé et al., 2004; Peterson and Cushion, 2005; Le Gal et al., 2012; Alanio and Bretagne, 2017).

Host specificity and biotrophy

The strict host species specificity of the *Pneumocystis* species means that the fungal cells can only infect or survive in the host in which they were isolated in the first place. This view is mainly supported by the systematic failure of cross-infection experiments involving severely combined immuno-deficient animals and nude rats (Aliouat et al., 1993; Furuta et al., 1993; Gigliotti et al., 1993; Aliouat et al., 1994; Atzori et al., 1999; Durand-Joly et al., 2002). The selective activation of trophic cells by their host seems to trigger the formation of cytoplasmic projections by *Pneumocystis* cells, the filopodia (Aliouat-Denis et al., 2008). Accordingly, *P. carinii*, the species naturally infecting rats, is unable to form filopodia and infect when inoculated in mice, whereas *P. murina*, the natural parasite of mice, produce filopodia and high parasite loads under the same conditions (Aliouat-Denis et al., 2008). The function of the filopodia remains elusive but these structures display ultrastructural differences that are species specific, and that might account for some aspects of the host specificity.

Another aspect of this host specificity is that *Pneumocystis* species are most probably obligate biotrophs (Cushion et al., 2007; Cushion and Stringer, 2010; Hauser, 2014; Ma et al., 2016a). The way fungal parasites scavenge nutrients from their host is an active research field and three modes are broadly recognized: (i) biotrophy, where the parasite acquires nutrients from a living cell, (ii) necrotrophy, where host cells are killed to release nutrients, and (iii) saprotrophy, where the organism feeds on dead or decaying organic material. Biotrophs do little damage to host cells and lack virulence factors (van der Does and Rep, 2007). *Pneumocystis* perfectly fits to the biotrophy definition because they cause no apparent cell death and lack any experimentally verified fungal virulence factors such as glyoxylate cycle, secondary metabolism, and secreted effectors (Cushion et al., 2007; Cissé et al., 2012; Cissé et al., 2014; Ma et al.,

2016a). This implies that they rely entirely on their host for their survival and thus have evolved close relationships that rendered them host species specific.

The physiological characteristics of the hosts are key determinants of parasite adaptation (Poulin et al., 2006). For example, micromammals are small bodied with short lifespans, high reproduction rates, and high population densities, whereas these distinctive features are reversed in large mammals such as humans. The co-evolution theory predicts that parasitic species infecting micromammals exhibit a weaker host specificity compared those adapted to long-lived hosts with more stable population densities (Poulin et al., 2006). This prediction has been validated in fish parasites, among which strong host specificity is favored in stable resources found in hosts with a large body size (Sasal et al., 1999; Desdevises et al., 2002). As far as *Pneumocystis* is concerned, humans are infected by only one species whereas rats can be co-infected by two (Cushion et al., 1993 and 2004; Icenhour et al., 2006a; Golab, 2009). The number of *Pneumocystis* species able to infect rodents might even be more important, as shown by the recent discovery of multiple lineages shared among species and genera of the Southeast Asian murid species (Latinne et al., 2017). These findings might indicate a relaxation of the strict host specificity in small mammals harboring *Pneumocystis*, although additional supporting data are needed to fully challenge the concept of widespread strict host specificity.

Co-phylogenetic studies of *Pneumocystis* species and their hosts suggest that the host specificity evolved as a continuous trait resulting from a long-lasting co-evolution (Demanche et al. 2001; Guillot et al., 2001; Hugot et al., 2003). Strict host specificity is rare in animal pathogens but widespread in plant fungal pathogens (Parker and Gilbert, 2004; Restrepo et al., 2014). In the latter, the ecological adaptation often results in a pronounced specialization to particular hosts (Clay and Kover, 1996). In these systems, host specificity acts as a reproductive

isolating mechanism because it favors higher rate of mating between individuals on the same host and reduced gene flow among populations from different hosts (Giraud, 2006). A rapid divergence of the virulence factors, the pathogen "effector repertoire", is often associated with the emergence of host specificity (Schulze-Lefert and Panstruga, 2011). The hypothesis of the latter authors states that changes in pathogen host range is driven by variation in the pathogen effector repertoire. This description fits the lineage specific expansion of the MSG superfamily in *Pneumocystis* species (Ma et al., 2016a; Ma et al., 2016b; Schmid-Siegert et al., 2017), which suggests that these proteins might account for some aspects of the host specificity.

GENOME ORGANIZATION

Genomic data acquisition

The quest for genome sequence data began with the successful cloning of *P. carinii* genomic fragments (Tanabe et al., 1988). Pulse field gradient gel electrophoreses have been then instrumental for karyotypic characterization of *Pneumocystis* genomes and evidenced 12 to 20 chromosomes according to the species totaling ca. 8 Mb (Hong et al., 1990; Yoganathan et al., 1989; Stringer and Cushion, 1998). Differences in karyotype profiles determined that the species infecting humans and rats are genetically distinct (Stringer et al., 1993). Significant genome size variations among species have been reported, *e.g.* that of ferret *Pneumocystis* would be ca. 1.7 times bigger than that of *P. carinii* (Stringer and Cushion, 1998). A draft of *P. carinii* genome covering ca. 70% of genome was generated in 2006 using a clone-based Sanger sequencing approach from infected laboratory rats (Slaven et al., 2006). In 2012, the first draft of *P. jirovecii* genome was obtained from a single bronchoalveolar lavage of a patient with pneumonia (Cissé et al., 2012). This assembly encompasses 358 contigs capturing 90 to 95% of the genome, but the repetitive subtelomeric and centromeric regions could not be resolved. The centromeres have not been discovered yet in *Pneumocystis*, whereas the subtelomeric regions were resolved using Sanger sequencing of cosmids (Keely et al. 2005), and more recently assembled using sequencing generating long reads (Ma et al., 2016a; Ma et al., 2016b; Schmid-Siegert et al., 2017) (see below, chromosomal ends section). Chromosomal level assemblies of *P. jirovecii*, *P. carinii*, and *P. murina* were recently published revealing genome sizes ranging from 7.4 to 8.3 Mb (Table 1; Ma et al., 2016a). It became evident that the genomes of *Pneumocystis* species had undergone an important reduction relative to *S. pombe* (7.5 to 8.3 Mb versus 12.5 Mb).

Nuclear Genome content

The analysis of the *Pneumocystis* genome assemblies validated the presence of single copy ribosomal DNA reported previously for *P. jirovecii* (Giuntoli et al. 1994; Stringer, 1996; Tang et al., 1998; Nahimana et al., 2000a). This is similar to *Taphrina deformans* (Cissé et al., 2013), but contrasts with most fungi which harbors commonly tens or hundreds of copies of the locus. Figure 3 shows the genome compositions of *Pneumocystis* species compared to related fungi. These data highlight the contraction of the protein coding regions as compared to free-living yeasts, which reflect massive gene losses. Figure 3 also evidences the expansions of the MSG superfamily, of introns, as well as of the cumulative length of the intergenic regions (IGR). We previously reported that IGR in *P. jirovecii* occupy a larger genome fraction as compared to free living yeasts *Saccharomyces cerevisiae* and *S. pombe* despite a significantly smaller genome (Cissé et al., 2014). This observation holds when we re-evaluate here IGRs in the newly published full-length genomes of *P. jirovecii*, *P. carinii*, and *P. murina* (Ma et al. 2016a). This strongly suggests that genome streamlining in *Pneumocystis* species is driven by gene deletions rather than reduction of IGRs. This observation seems counterintuitive because genome reduction is almost always associated to a reduction of introns and IGRs in parasites (Keeling and Slamovits, 2005). Alternatively, large IGRs might favor chromosomal re-arrangements by increasing the number of possible breakpoints, as hypothesized in fungal microsporidian parasites (Slamovits et al., 2004; Keeling and Slamovits, 2005). The other characteristics of these genomes are discussed in the following sections.

Chromosomal ends

Subtelomeres in microbial parasites are often enriched with multi-copy surface glycoprotein gene families (Deitsch et al. 2009). These genomic regions are prone to (i) gene silencing that can be used for mutually exclusive expression, (ii) enhanced mutagenesis, and (iii) ectopic recombinations facilitated by the formation of clusters of telomeres at the nuclear periphery (Barry et al., 2003). These regions correspond to an important proportion of the *Pneumocystis* genomes (ca. 5%), and harbor a superfamily including five to six families of highly polymorphic multi-copy proteins called major surface glycoproteins (MSG) that are believed to be crucial for the fungus' lifestyle (Ma et al., 2016a; Schmid-Siegert et al., 2017). These *msg* genes exist only in *Pneumocystis* species and all species of *Pneumocystis* have their own repertoire, which suggest they have been acquired in a common ancestor, although their origin is not known. The absence of homology of these MSGs outside *Pneumocystis* lineages might indicate a transfer from an unknown species or a gene co-option. *msg* families have been first described and studied in *P. carinii* (Kovacs et al., 1993; Sunkin et al., 1994; Sunkin et al., 1996; Keely et al., 2005; Keely and Stringer, 2009), and subsequently analyzed in *P. jirovecii* and *P. murina* (Haidaris et al., 1998; Kutty et al., 2008; Ma et al., 2016a; Schmid-Siegert et al., 2017). Important differences exist among *Pneumocystis* species in terms of *msg* gene copy numbers, 60 to 140 copies per cell, and protein divergence (Ma et al., 2016a).Moreover, one MSG family is present only in *P. jirovecii* (*msg*-IV or -B), whereas another one is present only in *P. carinii* and *P. murina* (MSR family, *i.e.* MSG-related). MSGs are believed to be involved in antigenic variation (Stringer, 2007). MSGs would also mask glucans at the asci surface from the immune recognition (Kutty et al., 2016). The antigenic diversity seems to be created via intra-family recombination of *msg* genes encoding different isoforms, creating mosaic genes, as well as through increased

mutagenesis (Kutty et al., 2008; Keely and Stringer, 2009; Schmid-Siegert et al., 2017). The expression of the most abundant MSG family (*msg*-I or –A1) that is present in all species is subject to mutually exclusive expression of a single isoform in each cell by using a single copy transcription promoter (the upstream conserved sequence, UCS) (Edman et al., 1996; Kutty et al., 2001; Sunkin et al., 1996; Wada et al., 1995). The UCS ends by the conserved recombination joint element (CRJE) which is also present at the beginning of each *msg*-I gene and may serve as recombination breakpoint (Stringer, 2007). The CRJE would be larger in *P. wakefieldiae* (ca. 330 bps) than in *P. murina* (132), which in turn is larger than in *P. carinii* and *P. jirovecii* (28 and 33, respectively) (Keely et al., 2007). On the other hand, at least in *P. jirovecii,* members of the other five families possess each their own promoter (Schmid-Siegert et al., 2017), but their expression patterns remain to be characterized. Recently, one family has been shown in *P. murina* to be expressed only in ascospores within asci and young trophic forms (Bishop et al., 2018).

Introns

Introns are extremely abundant in *Pneumocystis* genes and are as many as several tens per gene with a mean of five, and more than 40% of genes are interrupted by at least four introns (Stringer and Cushion, 1998; Ma et al., 2016a). Their presence can be equally explained by massive gains in *Pneumocystis* most recent common ancestry, or retention of ancestral elements that would have been lost in some Taphrinomycotina lineages such as *Schizosaccharomyces*. The introns are short (average length of 48 nucleotides), have a strong adenine and thymine bias, and present typical donor, acceptor and branch site patterns (Slaven et al., 2006). *Pneumocystis* introns cannot be processed by *S. pombe* and *S. cerevisiae* spliceosomes because of the divergence in intron-exon boundaries and branching sites within the introns (Thomas et al., 1999). RNA-seq

data indicate that intron retention affects ca. 45% of all introns (Ma et al., 2016a). *Pneumocystis* species contain self-splicing group I introns that are absent in higher eukaryotes such as humans (Liu et al., 1994), which renders them a prime target for the development of new drugs. These latter introns catalyze their own excision from RNA transcripts, a reaction that is inhibited by the drug pentamidine that is used against *Pneumocystis* (Liu and Leibowitz, 1993).

Given the important genome reduction at the *Pneumocystis* genus level, the presence of a high intron density per gene suggests a selective constraint to conserve them. Intron loss is dominant in fungi (Stajich et al., 2007), and this tendency is even more pronounced in some parasites such as microsporidia (Keeling et al., 2010). The intron history is highly flexible within the Taphrinomycota, with the plant-associated *Neolecta* having a high intron density similar to *Pneumocystis* (Nguyen et al., 2017), and the intron-poor free-living yeast *S. pombe* (Wood et al., 2002). The non-sense-mediated mRNA decay machinery is conserved in *Pneumocystis* species (Ma et al., 2016a). Under neutral scenario (no advantage) and widespread intron retention, most of the introns would produce non-functional transcripts tagged for destruction. This would be an incredible waste of resources in absence of another function. The latter could consist in alternative splicing increasing transcript diversity and regulating gene transcription or mRNA stability. Consistently, the *P. carinii* inosine 5'-monophosphate dehydrogenase pre-mRNA is differentially spliced, which was suggested to reflect changes in environmental stresses (Ye et al., 2001). These considerations suggest that introns might be neutral elements involved in many cellular processes via a greater proteome diversity, possibly including acting as a favorable substrate to facilitate shifts in lifestyle (*i.e.* parasite transition from one host species to another, or from plant to animal).

Mitogenomes

The mitochondrial genomes of *P. carinii* (Sesterhenn et al., 2010; Ma et al., 2013), *P. jirovecii* (Cissé et al., 2012; Ma et al., 2013), and *P. murina* (Ma et al., 2013) have been sequenced. The mitogenome sizes range from 24 to 35-kb with a substantial size variability among isolates in all species. *P. carinii* and *P. murina* mitogenomes end with single-stranded loop sequences that would allow forming linear concatemers and protecting the ends of the molecule. The presence of these repeats might account for the variable size of *P. carinii* mitogenomes. *P. jirovecii* mitochondrial genome is circular since it lacks inverted terminal repeat allowing circulation. The significance of circularity versus linearity is unknown. Related Taphrinomycota of the genera *Schizosaccharomyces, Taphrina,* and *Neolecta* have circular genomes (Bullerwell et al., 2003; Cissé et al., 2013; Tsai et al., 2014; Nguyen et al., 2017), which might indicate that the circular form is ancestral. Interestingly, *P. carinii* and *P. murina* mitogenomes are highly co-linear whereas *P. jirovecii* mitogenome presents some re-arrangements, similarly to the nuclear genomes (see below Chromosomal re-arrangement section). The gene content is highly conserved among the three *Pneumocystis* species, although there is a substantial nucleotide divergence among species (27 to 31%) (Ma et al., 2013). These mitogenomes encode ca. 17 genes commonly found in mitochondrial fungal genomes such as ATP synthases, cytochrome *c* oxidases, NADH dehydrogenases, and the full repertoire of at least 20 transfer RNAs. Reports investigating the dynamics of the mitochondrial genes during infection have revealed that mitogenomes would be very plastic in terms of copy number variations (Valero et al., 2016), and of genetic diversity including heteroplasmy (Alanio et al., 2016). The subsequent sections of this review focus on nuclear genomes.

EVOLUTION

Comparison of the gene families and pathways present in *Pneumocystis* genomes to those in selected fungi has revealed numerous losses / contractions and relatively few expansions (Table 2). The hypothetical evolutionary history of *Pneumocystis* species derived from these observations is represented in Figure 4 and discussed in the following sections.

Losses in the metabolic and cellular machineries

Massive gene losses suggest that *Pneumocystis* species are auxotroph for essential nutrients, which might explain the recurrent failures of *in vitro* culturing attempts. The lost pathways include basic components of metabolic machinery such as the synthesis of amino acids or carbohydrates (Table 2). The loss of purines catabolism seems unique to *Pneumocystis* (Chitty and Fraser, 2017). *Pneumocystis* species are able to synthesize fecosterol and episterol but lack enzymes to convert them into ergosterol. Consequently, their membranes contain cholesterol instead of ergosterol, which probably explains their resilience to azole treatment. *Pneumocystis* organisms are also able to synthetize a unique class of sterols, the "pneumocysterols" (Kaneshiro et al. 1994; Kaneshiro et al. 1999; Florin-Christensen et al. 1994; Giner et al., 2002). It interesting to note the early steps of the sterol biosynthetic pathway leading to the formation of pneumocysterol and episterol are conserved in *Pneumocystis* species, and only the final steps toward ergosterol/cholesterol production are missing (Joffrion et al., 2010). This is exemplified by the fact that key enzymes for the formation of ergosterol (*i.e. erg3*, *erg4* and *erg5*) are not identifiable within the genomes. Analysis of the sterol biosynthesis machinery suggest that these species may be able to synthetize ergosterol/cholesterol precursors such as zymosterol, fecosterol

and episterol. Thus, the sterol pathway may have been re-routed and branch to form unique sterols, the pneumocystisterols.

Overall, these observations are consistent with the idea that losses of metabolic genes correlate with an increased dependency of the parasite on its host. Therefore, nutrients need to be scavenged from the host, which often mechanistically involves large batteries expanded transporters (*e.g.* as observed in microsporidia [Cuomo et al., 2012]). This is not the case in *Pneumocystis* since transporters families are also greatly reduced (Cissé et al., 2014; Ma et al., 2016a). For instance, the amino acid permeases and transporters that can respectively carry amino acids and oligopeptides are greatly reduced relatively to other Taphrinomycota (one copy of general amino acid permease versus 21 copies in *S. pombe*). Transmembrane proteins such as those of the major facilitator superfamily, sugar transporters, or more specific transporters (*e.g.* efflux pumps) are significantly reduced in *Pneumocystis*. The reduction of the transporters battery might be compensated by the use of highly selective transporters for critical compounds. The recent discovery of the import of *myo*-inositol in *Pneumocystis* cells via a low affinity but highly selective system supports this idea (Cushion et al., 2016). Unfortunately, high affinity transporters cannot be identified solely by computational means. Alternatively, simple diffusion across the membrane may occur, as evidenced in *P. carinii* for amino acids uptake using *in vitro* experiments (Basselin et al, 2001a; Basselin et al, 2001b). Basic cellular machinery is also affected by the loss of several fungal specific transcription factor families and the RNA interference machinery (Table 2).

Evolutionary basis of gene loss

Gene loss is a common trend in parasitic and symbiotic species, which often harbor a small sized genome (Keeling and Slamovits, 2005; Wolf and Koonin, 2013). The driving factors are often unknown or specific to the lifestyle of the species under study. A central question in evolutionary biology is whether gene loss is neutral or adaptive. In *Pneumocystis* species, there are footprints of both processes and we discuss here a few examples.

The neutral theory is usually sufficient to explain gene loss in parasites (O'Malley et al., 2016). Organisms with narrow host niche such as *Pneumocystis* are predicted to have small sized populations with increased **genetic drift** (bold: see glossary) (Papkou et al., 2016). The main mechanisms for gene loss are pseudogenization and sudden DNA deletions. Pseudogenization consists in the accumulation of deleterious mutations in non-essential genes ultimately leading to the loss (Kuo and Ochman, 2009; Wernegreen, 2015). The proportion of pseudogenes in *P. jirovecii* is low and equivalent to that present in free-living yeasts (0.02 pseudogene per protein-coding gene [Cissé et al., 2014]). This observation might indicate that pseudogenization is not the main driver of gene loss in this species. The following considerations do not undermine this observation but suggest that caution must be exercised: (i) this rate of pseudogenization is valid only for *P. jirovecii* and for the single isolate which genome was sequenced (Cissé et al., 2012), and (ii) only genes including stop codons were considered, that is, other types of gene inactivation were not considered (*e.g.* untranslated RNA genes or unfixed mutations). Gene loss can also be result of deletions independent of selection such as the movement of transposable or integrated viruses (reviewed by Albalat and Canestro, 2016).

The adaptive theory of gene loss implies a selective advantage and has been demonstrated to have occurred in many pathogenic lineages, for example for the Allergen 1 in

Cryptococcus neoformans (Jain et al., 2009), and for the *de novo* biosynthesis of nicotinic acid genes in *Candida glabrata* (Domergue et al., 2005). In *Pneumocystis*, the loss of chitin might have been lost to allow avoiding recognition from the host immune system (Ma et al., 2016a). The gene families and pathways cited in Table 2 are missing in the three *Pneumocystis* genomes available (Ma et al 2016a), which suggests that these losses occurred before the radiation of the genus. An unexpected consequence is that the observed gene losses might not reflect the current selective forces, and therefore might not be relevant for the host specificity.

Chromosomal re-arrangement

The chromosome level assemblies revealed that an important chromosomal re-arrangement occurred among *Pneumocystis* species (Ma et al., 2016a). The re-arrangement, however, followed the species tree, that is, the macrosynteny is broken between rodents infecting *Pneumocystis* (*P. carinii* and *P. murina*) and the humans infecting species (*P. jirovecii*), whereas *P. carinii* and *P. murina* genomes are highly collinear. Nevertheless, the gene order is conserved in syntenic regions among the three species (>92% of the genes), and ca. 83% of gene families are orthologous, with 4 to 30% of divergence at the nucleotide sequence level. The high gene conservation among the three species suggest that re-arrangements occurred mostly in the intergenic regions (IGR). In fungi, IGRs are often enriched in regulatory functions such as signal transduction or binding sites of transcription factors (Noble and Andrianopoulos, 2013). Chromosomal translocations impact gene expression as well as long-distance gene-to-gene contact via chromatin interactions, and thus might be involved in speciation (Rieseberg, 2001; Bakloushinskaya, 2016). Protein evolution is also faster in re-arranged chromosomes than collinear chromosomes because re-arrangements reduce homologous recombination and

facilitate positive selection (Rieseberg, 2001). A key question here is whether chromosomal re-arrangements are involved in the adaptation of each *Pneumocystis* species to its host. Future studies are required to probe an eventual role of these re-arrangements in *Pneumocystis* evolution.

Loss of complex multicellularity

The recent sequencing of the *Neolecta irregularis* genome revealed that the Taphrinomycotina last common ancestor was probably multicellular (Nguyen et al., 2017). These findings suggest that *Pneumocystis* organisms evolved from a plant-associated or soil-adapted multicellular organism. The shift in cell morphology to single celled organisms is associated with the deletion of an ancestral morphogenic kit that included many cell differentiation and cell-to-cell signaling genes. These losses are not specific to *Pneumocystis* and were observed in a wide range of unrelated yeasts (Nguyen et al., 2017; Nagy et al., 2014; Nagy, 2017), which suggests a convergent evolution. The transition from a hyphal to yeast form takes place in many fungal lineages and is often triggered by a thermal stimulus (Köhler et al., 2017), CO_2 levels (Hall et al., 2010), or pH (Davis, 2009), and is directly linked to the ability to invade hosts. Notable examples include the dimorphic human pathogenic fungi *Histoplasma*, *Blastomyces*, *Coccidioides*, and *Paracoccidioides* (Beaman et al., 1981; Medoff et al., 1987; Inglis et al., 2013).

The ancestral morphogenic kit for complex multicellularity (fruiting bodies) is lost in *Pneumocystis*. However, *Pneumocystis* species are able to produce biofilms (Cushion et al., 2009), which is an undifferentiated form of aggregative multicellularity often seen in bacteria (Claessen et al., 2014). Inversely, the yeast *Saitoella complicata* grows primarily by budding

(Goto et al., 1987), despite having the cellular machinery for the production of fruiting bodies (Nguyen et al., 2017). Comparative genomics and epigenomics would be extremely valuable to explore the molecular process underlying the loss of the multicellular phenotype. These considerations highlight the fact that phenotypes cannot be explained solely by gene loss and gain balance, and that other subtle mechanisms need to be considered.

POPULATION GENETICS

Strain typing

Given the high homogeneity of genomic sequences at the nucleotide sequence level among *P. carinii* isolates, strain typing for this species relied on chromosomes' size analyses which allowed identifying numerous different karyotypic forms (Lundgren et al., 1990; Cushion, 1998; Wakefield, 1998a; Nahimana et al., 2001). On the other hand, the low but significant heterogeneity in many genomic loci among *P. jirovecii* isolates allowed using multilocus sequence typing (Wakefield, 1998b). The latter method represents nowadays the most used technique for *P. jirovecii* strains identification. The discrimination power of eight distinct loci has been validated and extensively used for epidemiological studies of *P. jirovecii* pneumonia (Maitte et al., 2013). Genotypes identification is performed by PCR of multiple loci followed by direct DNA sequencing (Sanger), restriction fragment length polymorphism, single-strand conformation polymorphism, type-specific oligonucleotide hybridization, tandem repeats number analysis, or high-throughput amplicon sequencing (Hauser et al., 1997; Hauser et al., 1998; Lee et al, 1993; Lu et al., 1995; Ma et al., 2002; Alanio et al., 2016; Esteves et al., 2016).

Genetic diversity

The conclusions drawn from the studies concerning *Pneumocystis* genetic diversity were often contradictory. Low levels of genetic diversity as defined by Shannon diversity and Simpson indexes (Shannon, 1948; Simpson 1949) have been reported at the *P. jirovecii* and *P. carinii* internal transcribed spacers of the nuclear rDNA operon using PCR-based Sanger sequencing (Palmer et al., 2000; Beser et al., 2011). On the other hand, moderate to important levels of diversity measured in term of DNA polymorphisms in *P. jirovecii* using multilocus sequence

typing have been reported (Matos and Esteves, 2010; Jarboui et al., 2013; Sun et al., 2015; Alanio et al., 2017). The lack of whole genome sequence data, differences in sampling strategies, differences in interpretation, as well as the likely frequent *in vitro* formation of PCR chimeras (Beser et al., 2007), make difficult the reconciliation of these conclusions.

Moreover, sexual recombination could explain partly these conflicting conclusions. Indeed, sexual reproduction is one of the main mechanisms to generate genetic diversity in fungi. It is believed to favor adaptation in fluctuating conditions while purging deleterious alleles (Heitman, 2010). *Pneumocystis* are probably homothallic species (Almeida et al., 2015; see life cycle section), and self-fertilization favors mating by avoiding the search of a compatible partner, a strategy thought to be favorable to and adopted by several human pathogens such as *Cryptococcus* and *Candida* species (Heitman, 2010). Sexual reproduction is based on classical Mendelian segregation, which supports both cross- and self-fertilization (Buscaglia et al., 2015). *Pneumocystis* would be able to perform both clonal and sexual propagation with various degrees of inbreeding or outcrossing. These variations in the multiplication process could explain the conflicting patterns of genetic diversity reported.

Polymorphism rates change substantially across loci and chromosomes in various species, including fungi, plants, and animals (Ellegren and Galtier, 2016). Genetic diversity is influenced by three main forces: mutation, demography (migration and **bottlenecks**), and selection (**selective sweeps** or **clonal interference**). Demography and selection create differences in the effective population size, whilst variations in mutation rate may create differences in the level of genetic diversity according to the geographical location. Neutral mutation rates in eurotiomycetes are typically between 1×10^{-8} and 1×10^{-9} substitutions per site per year (Kasuga et al., 2002), and a rate of 1.2×10^{-10} for the 18S rDNA has been used to

estimate *Pneumocystis* species divergence (Keely et al., 2003a). However, the genome-wide mutation rates for these species are unknown and expected to fluctuate greatly among genomic regions. For example, subtelomeric regions harboring MSGs have high substitution rates (Keely and Stringer, 2009; Schmid-Siegert et al., 2017), whereas ribosomal regions display a normal rate (Fischer et al., 2006). Moreover, given their likely variations according to the host, the mutation rates for each species must be determined independently. Care must be taken inferring these rates because recombination can be mutagenic and its impact as well as other confounder effects need to be addressed.

The size of the populations of *Pneumocystis* species are not known, but they are expected to be small because of their narrow host ranges. *P. jirovecii* would have a small population size relative to the species infecting micro-mammals, thus reflecting the small size of human populations relative to those of rodents. Variations in population size over time affect the genetic diversity, *e.g.* a strong population **bottleneck** creates a loss of allele diversity due to increased **genetic drift**. Using non-recombining neutral loci, realistic mutation rates, and appropriate molecular clock models, past population history can be traced back using coalescent theory based applications such as skyline plot methods (Drummond et al., 2005; Heled and Drummond, 2008). These demographic reconstructions would provide key metrics such as ancestral population sizes and evolutionary rates.

Interestingly, the strongest prediction of genetic diversity in many species is the life history, not the population history (Ellegren and Galtier, 2016). This means that there is a strong correlation between phenotypic traits (*e.g.* mating system, generation times) and the genetic diversity. For example, homothallism is expected to have long term evolutionary cost fitness because selfing populations experience reduced recombination rates and size, which ultimately

reduce the strength of purifying selection and increase genetic drift (Charlesworth and Wright, 2001; Hill and Robertson, 1966; Otto and Lenormand, 2002; Pollak, 1987). The homothallism used by *Pneumocystis* species is also often associated to higher probability to experience population **bottlenecks** via founder effects and linked selection (Jarne, 1995; Charlesworth and Wright, 2001). There is a complex interplay between demographic, selective factors, and genetic diversity. Alternative scenarios, such as purifying selection purging deleterious alleles, which is known as "background selection" (Charlesworth, 1994), need also to be considered. In conclusion, many factors may have influenced genetic diversity of *Pneumocystis* species, which remains unclear.

Population structure

The population structure of *Pneumocystis* species is also controversial. Indeed, data support an absence of strong subdivision in *P. jirovecii* (Parobek et al., 2014) and *P. carinii* (Palmer et al., 2000), whereas other data support possible geographical clusters in *P. jirovecii* (Esteves et al., 2016; Alanio et al., 2017). Importantly, Matos and Esteves (2010) noted that the infections are not necessarily clonal and recombination between multi-locus genotypes is possible. All these inferences are based on a relatively small number of markers (*e.g.* ITS, mitochondrial large subunit rDNA), and need to be validated at the genome scale using appropriate Bayesian methods based on unlinked multi-allelic genotypes, such as STRUCTURE (Pritchard et al., 2000). In the meantime, interesting clues can be extracted from the biological cycle. The question is whether the fluctuation of the population structure is caused by variations in spore dispersal or in sexual recombination. The asci are 4-6 μm in size, which is small enough to be airborne dispersed efficiently over long distances. The asci cell wall is enriched with

glycoproteins, melanin, β-glucans, and mannans without outer chain (Kottom and Limper, 2000; Icenhour et al., 2003; Icenhour et al., 2006b; Ma et al., 2016a), which might allow them to resist desiccation and UV irradiation usually fatal to many fungal spores (Golan and Pringle, 2017; Latgé, 2007). The viability of spores for extended periods of time is supported by the detection of *P. jirovecii* mRNA in hospital air samples (Latouche et al., 2001; Maher et al., 2001). Their resistance to physical assaults is suggested by their detection in air spora trapped in rural locations (Wakefield, 1996).

Dispersal of fungi can occur in two modes: (i) multiple sequential short-distance dispersal, and (ii) a single successful long-distance move of spores ultimately coinciding with optimal conditions for the growth of the fungus (Golan and Pringle, 2017). The former option produces a strong population subdivision, while the latter ends up with no or weak population structure because the same genotype(s) will be spread over large geographical distance. If long distance dispersal occurs on a global scale, it will result in a global population structure (Pringle et al., 2005). Rare long-distance dispersal would involve stochastic founding events, which can be revealed by population structures with an excess of rare alleles. Future studies combining genetic and geography are needed to fully access the population structures of *Pneumocystis* species.

Clonal evolution or predominant sex/recombination?

P. jirovecii infections are most often caused by multiple populations co-infecting the same individual (Hauser et al., 1997; Nahimana et al., 2000b; Palmer et al., 2000; Ma et al., 2002; Alanio et al., 2016). Multilocus genotypes (MLGs), which refer to a unique combination of alleles, can persist over long periods of time (4 to 9 years), and be observed across different

countries (Wakefield et al., 1994; Esteves et al., 2010). Recombination was also detected among MLGs (Esteves et al., 2010), which would explain the reported lack of strong population subdivision, at least in *P. jirovecii* (Parobek et al., 2014). Under **panmictic** population conditions, MLGs should not persist in the population because they will be disturbed consistently by recombination.

Buscaglia et al. (2015) proposed that "a highly structured (*i.e.* clonal) population indicates that the main mode of reproduction for such a species lacks genetic exchange (*i.e.* is primarily asexual) or sex occurs only rarely". In *Pneumocystis*, MLGs do recombine which would indicate a limited global population structure (Esteves et al., 2010; Parobek et al., 2014). Thus, the definition proposed by Buscaglia et al. would suggest a widespread sexual reproduction in *Pneumocystis* species. However, some MLGs persist over time, which suggests that these species might be mostly clonal and only rarely engage to sexual events. This latter scenario would be consistent with the theory of predominant clonal evolution (Tibayrenc and Ayala, 2012; Tibayrenc and Ayala, 2014), which proposes that restrained recombination is not strong enough to disturb the pattern of clonal structure. The frequency of recombination events at the genome level is unknown in *Pneumocystis*, which currently prevents reaching definitive conclusions.

Intra individual short-term evolution

Infections are usually caused by multiple *P. jirovecii* strains acquired from different origins (infections *de novo* but also possibly re-activation of organisms). The balance between different strains will likely change over the course of the disease because of either drug treatment, pressures from the host immune system, and/or varying metabolism and fitness among the strains

present. Other pathogens such as *Candida* and *Cryptococcus* species evolve rapidly within their hosts by acquiring new mutations or changes in genomic heterozygosity associated with drug resistance (Ford et al., 2015; Chen et al., 2017). It is unclear if theses mutations result from positive selection or DNA repair errors in *Cryptococcus* (Rhodes et al., 2017), although it is also possible that DNA repair errors are selected by positive selection. Competition among multiclonal parasite populations within the same host can, in theory, promotes parasite diversity (Bashey, 2015). The full extent of *Pneumocystis* short-term evolution within their host is unknown. Interestingly, Alanio and colleagues used a set of markers to evidence changes in population composition during *P. jirovecii* infections (Alanio et al. 2016). Multiple strains infections are frequently found in pathogens and may have clinically relevant consequences (Balmer and Tanner, 2011). Different strains might have different susceptibility to treatment or evolve differently so that they may escape detection by the immune system or diagnostics tools. We anticipate that the characterization of multiclonal infections will have serious implications for the treatment and the management of *P. jirovecii* pneumonia. Experimental setups will become realistic when long-term *in vitro* culture method will become widely reproducible.

PERSPECTIVES AND CONCLUSIONS

The research on *Pneumocystis* is still in its infancy mainly because of the lack of culture *in vitro*, but the availability of genomic data will help exploring the mysteries of their evolution. The next cornerstone will be the establishment of a long-term culture system and genetic manipulation. The upcoming expectation goes far beyond the *Pneumocystis* research community and will allow exploring key questions in evolutionary cell biology such as the evolution of parasitism and multicellularity. The study of *Pneumocystis* organisms has the unique interest that they are the only strictly mammalian-adapted fungal pathogens. Thus, determining the molecular basis of their adaptation and speciation are of uttermost importance. The key questions are: what are the determinants of the genome reduction? What are the molecular determinants of the host specificity and speciation? Why introns are so abundant and what are their function(s)? What are the impact of multiclonal infections and short-term evolution within host in the context of drug resistance and development of vaccines? How do natural populations of *Pneumocystis* evolve in different hosts?

ACKNOWLEDGMENTS

OHC is currently supported by an US National Institutes of Health Fogarty fellowship. PHM's research is presently funded by the Swiss National Science Foundation (grant 310030_165825). This research was supported by the Intramural Research Program of the NIH, Clinical Center. We thank Joe Kovacs and Liang Ma for critical reading of the manuscript and fruitful comments.

GLOSSARY

Background selection

Reduction of genetic diversity at linked loci owing to selection against deleterious mutations.

Bottleneck

A sharp and rapid reduction in the size of a population.

Clonal interference

Phenomenon in population genetics of organisms with significant linkage disequilibrium (*i.e.* absence of recombination), especially in asexual organisms. It occurs when two (or more) different beneficial mutations arise independently in different individuals.

Effective population size

The size that a theoretical population evolving under a Wright-Fisher model would need to be in order to match aspects of the observed genetic data.

Genetic drift

Fluctuation of allele frequency among generations in a population owing to the randomness of survival and reproduction of individuals, irrespective of selective pressures.

Haploid selfing

Refers to true homothallic species. A species able to accomplish their entire sexual reproduction without the need of a partner.

Heterozygosity

Measure of the genetic diversity, which represents the presence of different alleles at one or more loci on homologous chromosomes. Often presented as a probability that two randomly sampled gene copies in a population carry distinct alleles.

Linkage disequilibrium

Nonrandom association of alleles at two loci often but not always due physical linkage. Such association is broken over time by recombination.

Panmictic population

Random mating among individuals in an idealized population.

Selective sweep

Elimination or reduction of genetic diversity in the neighborhood of a beneficial allele that increases in frequency in the population, typically after an environmental change.

Selective sweeps

Elimination or reduction of genetic diversity in the neighborhood of a beneficial allele that increases in frequency in the population, typically after an environmental change.

REFERENCES

Alanio, A., Gits-Muselli, M., Mercier-Delarue, S., Dromer, F., Bretagne, S., 2016. Diversity of *Pneumocystis jirovecii* during infection revealed by ultra-deep pyrosequencing. Front. Microbiol. 7, 733.

Alanio, A., Bretagne, S., 2017. *Pneumocystis jirovecii* detection in asymptomatic patients: what does its natural history tell us? F1000 Faculty Rev. 6, 739.

Alanio, A., Gits-Muselli, M., Guigue, N., Desnos-Ollivier, M., Calderon, E. J., Di Cave, D., Dupont, D., Hamprecht, A., Hauser, P.M., Helweg-Larsen, J., Kicia, M., Lagrou, K., Lengerova, M., Matos, O., Melchers, W.J.G., Morio, F., Nevez, G., Totet, A., White, L. P., Bretagne, S., 2017. Diversity of *Pneumocystis jirovecii* across Europe: a multicentre observational study. EBioMedicine 22, 155-163.

Albalat, R., Canestro, C., 2016. Evolution by gene loss. Nat. Rev. Genet. 17, 379-391.

Aliouat, E.M., Mazars, E., Dei-Ca, E., Cesbron, J., Camus, D., 1993. Intranasal inoculation of mouse, rat or rabbit-derived *Pneumocystis* in SCID mice. J. Protozool. Res. 3, 94-98.

Aliouat, E.M., Mazars, E., Dei-Cas, E., Delcourt, P., Billaut, P., Camus, D., 1994. *Pneumocystis* cross infection experiments using SCIDmice and Nude rats as recipient host, showed strong hostspecies specificity. J. Eukaryot. Microbiol. 41, S71.

Aliouat, E.M., Dujardin, L., Martinez, A., Duriez, T., Ricard, I., Dei-Cas, E., 1999. *Pneumocystis carinii* growth kinetics in culture systems and in hosts: involvement of each life cycle parasite stage. J. Eukaryot. Microbiol. 46, S116-S117.

Aliouat-Denis, C.M., Chabé, M., Demanche, C., Aliouat, E.M., Viscogliosi, E., Guillot, J., Delhaes, L., Dei-Cas, E., 2008. *Pneumocystis* species, co-evolution and pathogenic power. Infect. Genet. Evol. 8, 708-726.

Aliouat-Denis, C.M., Martinez, A., E. Aliouat, M. Pottier, N. Gantois, Dei-Cas, E., 2009. The *Pneumocystis* life cycle. Memor. Instit. Oswaldo Cruz 104, 419-426.

Almeida, J.M.G.C.F., Cissé, O.H., Fonseca, A., Pagni, M., Hauser, P.M., 2015. Comparative genomics suggests primary homothallism of *Pneumocystis* species. mbio 6, e02250-14.

Atzori C, A.F., Angeli, E., Mainini, A., Micheli, V., Cargnel, A., 1999. *P. carinii* host specificity: attempt of cross infections with human derived strains in rats. J. Eukaryot. Microbiol. 46, S112.

Bakloushinskaya, I.Y., 2016. Chromosomal rearrangements, genome reorganization, and speciation. Biol. Bull. 43, 759-775.

Balmer, O., Tanner, M., 2011. Prevalence and implications of multiple-strain infections. Lanc. Infect. Dis. 11, 868-878.

Banerji, S., Lugli, E.B., Wakefield, A.E., 1994. Identification of two genetically distinct strains of *Pneumocystis carinii* in infected ferret lungs. J. Eukaryot. Microbiol. 41, S73.

Bao, W., Kojima, K.K., Kohany, O., 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA 6, 11.

Barry, J.D., Ginger, M.L., Burton, P., McCulloch, R., 2003. Why are parasite contingency genes often associated with telomeres? Int. J. Parasitol. 33, 29-45.

Bashey, F., 2015. Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. Phil. Trans. R. Soc. B 370, 20140301.

Basselin, M., Lipscomb, K.J., Qiu, Y.H., Kaneshiro, E.S., 2001a. Transport of aspartic acid, arginine, and tyrosine by the opportunistic protist *Pneumocystis carinii*. Biochim. Biophys. Acta 1511, 349-359.

Basselin, M., Qiu, Y.H., Lipscomb, K.J., Kaneshiro, E.S., 2001b. Uptake of the neutral amino acids glutamine, leucine, and serine by *Pneumocystis carinii*. Arch. Biochem. Biophys. 391, 90-98.

Beaman, L., Benjamini, E., Pappagianis, D., 1981. Role of lymphocytes in macrophage-induced killing of *Coccidioides immitis* in vitro. *Infect. Immun.* 34, 347-353.

Beimforde, C., Feldberg, K., Nylinder, S., Rikkinen, J., Tuovila, H., Dorfelt, H., Gube, M., Jackson, D.J., Reitner, J., Seyfullah, L.J., Schmidt, A.R., 2014. Estimating the Phanerozoic history of the Ascomycota lineages: combining fossil and molecular data. Mol. Phylogenet. Evol.78, 386-398.

Beser, J., Hagblom, P., Fernandez, V., 2007. Frequent in vitro recombination in internal transcribed spacers 1 and 2 during genotyping of *Pneumocystis jirovecii*. J. Clin. Microbiol. 45, 881-886.

Beser, J., Botero-Kleiven, S., Lebbad, M., Hagblom, P., Fernandez, V., 2011. A limited number of ITS haplotypes defines the diversity of *Pneumocystis jirovecii* strains in Sweden. Infect. Genet. Evol. 11, 948-954.

Bishop, L.R., Davis, A.S., Bradshaw, K., Gamez, M., Cissé, O.H., Wang, H., Ma, L., Kovacs, J.A., 2018. Characterization of p57, a stage-specific antigen of *Pneumocystis murina*. J. Infect. Dis. doi: 10.1093/infdis/jiy099. [Epub ahead of print].

Brown, G.D., Denning, D.W., Gow, N.A., Levitz, S.M., Netea, M.G., White, T.C., 2012. Hidden killers: human fungal infections. Sci. Transl. Med. 4, 165rv113.

Bullerwell, C.E., Leigh, J., Forget, L., Lang, B.F., 2003. A comparison of three fission yeast mitochondrial genomes. Nucl. Ac. Res., 31, 759-768.

Buscaglia, C.A., Kissinger, J.C., Aguero, F., 2015. Neglected tropical diseases in the post-genomic era. Trends Genet. 31, 539-555.

Ceré , N., Drouet-Viard, F., Dei-Cas, E., Chanteloup, N., Coudert, P., 1997. *In utero* transmission of *Pneumocystis carinii* sp. f. *oryctolagi*. Parasite 4, 325–330.

Chabé, M,. Dei-Cas, E., Creusy, C., Fleurisse, L., Respaldiza, N., Camus, D., Durand-Joly, I., 2004. Immunocompetent hosts as a reservoir of *Pneumocystis* organisms: histological and RT-PCR data demonstrate active replication. Eur. J. Clin.. Microbiol. 23, 89-97.

Chabé, M., Herbreteau, V., Hugot, J.P., Bouzard, N., Deruyter, L., Morand, S., Dei-Cas, E., 2010. *Pneumocystis carinii* and *Pneumocystis wakefieldiae* in wild *Rattus norvegicus* trapped in Thailand. J. Eukaryot. Microbiol. 57, 213-217.

Chabé, M., Aliouat-Denis, C. M., Delhaes, L., Aliouat, E.M., Viscogliosi, E., Dei-Cas, E., 2011. *Pneumocystis*: from a doubtful unique entity to a group of highly diversified fungal species. FEMS Yeast Res. 11, 2-17.

Chagas, C., 1909. Nova tripanozomiaze humana: estudos sobre a morfolojia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., ajente etiolojico de nova entidade morbida do homem. Mem. Inst. Oswaldo Cruz 1, 159-218.

Charlesworth, B., 1994. The effect of background selection against deleterious mutations on weakly selected linked variants. Genet. Res. 63, 213-227.

Charlesworth, D., Wright, S.I., 2001. Breeding systems and genome evolution. . Curr. Opin. Genet. Dev. 11, 685-690.

Chen Y, F. R., Giamberardino C, Sakthikumar S, Jones A, and e. a. Yang T (2017). Microevolution of serial clinical isolates of *Cryptococcus neoformans* var. *grubii* and *C. gattii*. MBio **8**: e00166–00117.

Cherry, J.M., Adler, C., Ball, C., Chervitz, S.A., Dwight, S.S., Hester, E.T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., Weng, S., Botstein, D., 1998. SGD: *Saccharomyces* Genome Database. *Nucl. Ac. Res.* 26, 73-79.

Chitty, J.L., Fraser, J.A., 2017. Purine acquisition and synthesis by human fungal pathogens. Microorg. 5, E33. doi: 10.3390/microorganisms5020033.

Christensen, C.B.V., Settnes, O.P., Bille-Hansen, V., Jorsal, S.E., Henriksen, S.A., Lundgren, B., 1996. *Pneumocystis carinii* from pigs and humans are antigenically distinct. J. Med. Vet. Mycol. 34, 431-433.

Churukian, C.J., Schenk, E.A., 1977. Rapid Grocott's methenamine silver nitrate method for fungi and *Pneumocystis carinii*. Am. J. Clin. Pathol. 68, 427-428

Cissé, O.H., Pagni, M., Hauser, P.M., 2012. *De novo* assembly of the *Pneumocystis jirovecii* genome from a single bronchoalveolar lavage fluid specimen from a patient. mBio 4, e00428-00412.

Cissé, O.H., Almeida, J.M., Fonseca, A., Kumar, A.A., Salojarvi, J., Overmyer, K., Hauser, P.M., Pagni, M., 2013. Genome sequencing of the plant pathogen *Taphrina deformans*, the causal agent of peach leaf curl. mBio 4, e00055-00013.

Cissé, O.H., Pagni, M., Hauser, P.M., 2014. Comparative genomics suggests that the human pathogenic fungus *Pneumocystis jirovecii* acquired obligate biotrophy through gene loss. Gen. Biol. Evol. 6, 1938-1948.

Claessen, D., Rozen, D.E., Kuipers, O.P., Sogaard-Andersen, L., van Wezel, G.P., 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies" Nat. Rev. Microbiol. 12, 115-124.

Clay, K., Kover, P. X., 1996. The Red Queen Hypothesis and plant-pathogen interactions. Ann. Rev. Phytopathol. 34, 29-50.

Cuomo, C.A., Desjardins, C.A., Bakowski, M.A., Goldberg, J., Ma, A.T., Becnel, J.J., Didier, E.S., Fan, L., Heiman, D.I., Levin, J.Z., Young, S., Zeng, Q., Troemel, E.R. 2012. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Gen. Res. 22, 2478-2488.

Cushion, M.T., Zhang, J., Kaselis, M., Giuntoli, D., Stringer, S.L., Stringer, J.R., 1993. Evidence for two genetic variants of *Pneumocystis carinii* coinfecting laboratory rats. J. Mycol. Med. 8, 1-6.

Cushion, M.T., 1998. Genetic heterogeneity of rat-derived *Pneumocystis*. FEMS Immunol. Med. Microbiol. 22, 51-58.

Cushion, M.T., Keely, S.P., Stringer, J.R., 2004. Molecular and phenotypic description of *Pneumocystis wakefieldiae* sp. nov., a new species in rats. Mycologia 96, 429-438.

Cushion, M.T., Smulian, A.G., Slaven, B.E., Sesterhenn, T., Arnold, J., Staben, C., Porollo, A., Adamczak, R., Meller, J., 2007. Transcriptome of *Pneumocystis carinii* during fulminate infection: carbohydrate metabolism and the concept of a compatible parasite. Plos One 2, e423.

Cushion, M.T., Collins, M.S., Linke, M.J., 2009. Biofilm formation by *Pneumocystis* spp. Eukaryot. Cell 8, 197-206.

Cushion, M.T., Linke, M.J., Ashbaugh, A., Sesterhenn, T., Collins, M.S., Lynch, K., Brubaker, R., Walzer, P D., 2010. Echinocandin treatment of *Pneumocystis* pneumonia in rodent models depletes cysts leaving trophic burdens that cannot transmit the infection. PLoS One 5, e8524.

Cushion, M.T., Stringer, J.R., 2010. Stealth and opportunism: alternative lifestyles of species in the fungal genus *Pneumocystis*. Annu. Rev. Microbiol. 64, 431-452.

Cushion, M.T., Collins, M.S., Sesterhenn, T., Porollo, A., Vadukoot, A.K., Merino, E.J., 2016. Functional Characterization of *Pneumocystis carinii* Inositol Transporter 1. mBio 7, e01851-16.

Cushion, M.T., Ashbaugh, A., Hendrix, K., Linke, M.J., Tisdale, N., Sayson, S.G., Porollo, A., 2018. Gene expression of *Pneumocystis murina* after treatment with anidulafungin results in strong signals for sexual reproduction, cell wall integrity, and cell cycle arrest, indicating a requirement for ascus formation for proliferation. Antimicrob. Agents Chemother. 62, e02513-17
Davis, D.A., 2009. How human pathogenic fungi sense and adapt to pH: the link to virulence. Curr. Opin. Microbiol. 12, 365–70

de Boer, M.G., de Fijter, J.W., Kroon, F.P., 2011. Outbreaks and clustering of *Pneumocystis* pneumonia in kidney transplant recipients: a systematic review.
Med. Mycol. 49, 673-80.

Dei-Cas, E., Aliouat, E.M., Cailliez, J.C., 2004. Cellular structure. In: Walzer, P.D., Cushion, M.T. (Eds.). *Pneumocystis* Pneumonia, Thrid ed. Marcel Dekker Inc., New York, pp. 61-94.

Dei-Cas, E., Chabé, M., Moukhlis, R., Durand-Joly, I., Aliouat, E.M., Stringer, J.R., Cushion, M.T., Noël, C., de Hoog, G.S., Guillot, J., Viscogliosi, E., 2006. *Pneumocystis oryctolagi* sp. nov., an uncultured fungus causing pneumonia in rabbits at weaning: review of current knowledge, and description of a new taxon on genotypic, phylogenetic and phenotypic bases. FEMS Microbiol. Rev. 80, 853-871.

Deitsch, K.W., Lukehart, S.A., Stringer, J.R., 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat. Rev. Microbiol. 7, 493-503.

Delanoë, P., Delanoë, E., 1912. Sur les rapports des kystes de carinii du poumon des rats avec le *Trypanosoma lewisi*. CR Acad. Sci. 155, 658-660.

Demanche, C., Berthelemy, M., Petit, T., Polack, B., Wakefield, A. E., Dei-Cas, E., Guillot, J., 2001. Phylogeny of *Pneumocystis carinii* from 18 primate species confirms host specificity and suggests coevolution. J. Clin .Microbiol. 39, 2126-2133.

Desdevises, Y., Morand, S., Legendre, P., 2002. Evolution and determinants of host specificity in the genus Lamellodiscus (Monogenea). Biol. J. . Linnean Soc. 77, 431-443.

Domergue, R., Castano, I., De Las Penas, A., Zupancic, M., Lockatell, V., Hebel, J.R., Johnson, D., Cormack, B.P., 2005. Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTI. Science 308, 866-870.

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M.J., Donoghue, P.C.J., Yang, Z.H., 2015. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. Curr. Biol. 25, 2939-2950.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 2, 1185-1192.

Durand-Joly, I., Aliouat, E.M., Recourt, C., Guyot, K., Francois, N., Wauquier, M., Camus, D., Dei-Cas, E., 2002. *Pneumocystis carinii* f. sp. *hominis* is not infectious for SCID mice. J. Clin. Microbiol. 40, 1862-1865.

Edman, J.C., Kovacs, J.A., Masur, H., Santi, D.V., Elwood, H.J., Sogin, M.L., 1988. Ribosomal RNA sequence shows *Pneumocystis carinii* to be a member of the fungi. Nature 334, 519-522.

Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus,C.W., Kovacs, J.A., 1996. A single expression site with a conserved leader sequence regulates variation of expression of the *Pneumocystis carinii* family of major surface glycoprotein genes. DNA Cell Biol 15, 989 –999

Ellegren, H., Galtier, N., 2016. Determinants of genetic diversity. Nat. Rev. Genet. 17, 422.

Eriksson, O.E., Winka, K., 1997. Supraordinal taxa of Ascomycota. Myconet 1, 1-16.

Esteves, F., Gaspar, J., Tavares, A., Moser, I., Antunes, F., Mansinho, K., Matos, O., 2010. Population structure of *Pneumocystis jirovecii* isolated from immunodeficiency virus-positive patients. Infect. Genet. Evol. 10, 192-199.

Esteves, F., de Sousa, B., Calderon, E.J., Huang, L., Badura, R., Maltez, F., Bassat, Q., de Armas, Y., Antunes, F., Matos, O., 2016. Multicentre study highlighting clinical relevance of new high-throughput methodologies in molecular epidemiology of *Pneumocystis jirovecii* pneumonia. Clin. Microbiol. Infect. 22, 566.e9-566.e19.

English, K., Peters, S.E., Maskell, D.J., Collins, M.E., 2001. DNA analysis of *Pneumocystis* infecting a Cavalier King Charles spaniel. J. Eukaryot. Microbiol. 48, S106.

Fischer, J.M., Keely, S.P., Stringer, J.R., 2006. Evolutionary rate of ribosomal DNA in *Pneumocystis* species is normal despite the extraordinarily low copy-number of rDNA genes. J. Eukaryot. Microbiol. 53, S156-S158.

Florin-Christensen, M., Florin-Christensen, J., Wu, Y.P., Zhou, L., Gupta, A., Rudney, H.,Kaneshiro, E.S., 1994. Occurrence of specific sterols in *Pneumocystis carinii*. Biochem. Bioph. Res. Co. 198, 236–242.

Ford, C.B., Funt, J., Abbey, D., Issi, L., Guiducci, C., Martinez, D.A., Delorey, T., White, T.C., Cuomo, C., Rao, R.P., Berman, J., Thompson, D.A., Regey, A., 2015. The evolution of drug resistance in clinical isolates of *Candida albicans*. elife 4, e00662.

Frenkel, J., 1999. *Pneumocystis* pneumonia, an immunodeficiency-dependent disease (IDD): a critical historical overview. J. Eukaryot. Microbiol. 46, S89–S92.

Furuta, T., Fujita, M., Mukai, R., Sakakibara, I., Sata, T., Miki, K., Hayami, M., Kojima, S., Yoshikawa, Y., 1993. Severe pulmonary pneumocystosis in simian acquired immunodeficiency syndrome induced by simian immunodeficiency virus: its characterization by the polymerase-chain-reaction method and failure of experimental transmission to immunodeficient animals. Parasitol. Res. 69, 624-628.

Gigliotti, F., Harmsen, A.G., Haidaris, C.G., Haidaris, P.J., 1993. *Pneumocystis carinii* is not universally transmissible between mammalian species. Infect. Immun. 61, 2886-2890.

Giner, J-L., H. Zhao, D. H. Beach, E. J. Parish, K. Jayasimhulu, Kaneshiro, E. S., 2002. Comprehensive and definitive structural identities of *Pneumocystis carinii* sterols. J. Lipid Res. 43, 1114–1124.

Giraud, T., 2006. Speciation: selection against migrant pathogens: the immigrant inviability barrier in pathogens. Heredity 97, 316-318.

Giuntoli, D., Stringer, S.L., Stringer, J.R., 1994. Extraordinarily low number of ribosomal RNA genes in P. carinii. J. Eukaryot. Microbiol. 41, 88S.

Gołab, E., 2009. Presence of *Pneumocystis carinii* and *Pneumocystis wakefieldiae* DNA in the extrapulmonary tissues of immunocompromised laboratory rats. Wiad. Parazytol. 55,167-71.

Golan, J.J., Pringle, A., 2017. Long-Distance Dispersal of Fungi. Microbiol. Spectr. 5, doi: 10.1128/microbiolspec.FUNK-0047-2016.

Goto, S., Sugiyama, J., Hamamoto, M., Komagata, K., 1987. *Saitoella*, a new anamorph genus in the Cryptococcaceae to accommodate two himalayan yeast isolates formerly identified as *Rhodotorula glutinis*. J. Gen. Appl. Microbiol. 33, 75-85.

Guillot, J., Demanche, C., Hugot, J.P., Berthelemy, M., Wakefield, A.E., Dei-Cas, E., Chermette, R., 2001. Parallel phylogenies of *Pneumocystis* species and their mammalian hosts. J. Eukaryot. Microbiol. suppl, 113s-115s.

Guillot, J., Demanche, C., Norris, K., Wildschutte, H., Wanert, F., Berthelemy, M., Tataine, S., Dei-Cas, E., Chermette, R., 2004. Phylogenetic relationships among *Pneumocystis* from Asian macaques inferred from mitochondrial rRNA sequences. Mol. Phylogenet. Evol. 31, 988-996.

Haidaris, C.G., Medzihradsky, O., Gigliotti, F., Simpson-Haidaris, P.J., 1998. Molecular characterization of mouse *Pneumocystis carinii* surface glycoprotein A. DNA Res. 5, 77-85.

Hall, R.A., De Sordi, L., MacCallum, D.M., Topal, H., Eaton, R., et al., 2010. CO2 Acts as a Signalling Molecule in Populations of the Fungal Pathogen *Candida albicans*. PLOS Pathogens 6, e1001193.

Hauser, P.M., Francioli, P., Bille, J., Telenti, A., Blanc, D.S., 1997. Typing of *Pneumocystis carinii* f. sp. *hominis* by single-strand conformation polymorphism of four genomic regions. J. Clin. Microbiol. 35, 3086-3091.

Hauser, P.M. , Blanc, D.S., Bille, J., Francioli, P., 1998. Typing methods to approach *Pneumocystis carinii* genetic heterogeneity. FEMS Immunol. Med. Microbiol. 22, 27-35

Hauser, P.M., Burdet, F.X., Cissé, O.H., Keller, L., Taffe, P., Sanglard, D., Pagni, M., 2010. Comparative genomics suggests that the fungal pathogen *Pneumocystis* is an obligate parasite scavenging amino acids from its host's lungs. Plos One 5, e15152.

Hauser, P.M., 2014. Genomic insights into the fungal pathogens of the genus *Pneumocystis*: obligate biotrophs of humans and other mammals. PLoS Pathog. 10, e1004425.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 289.

Heitman, J., 2010. Evolution of eukaryotic microbial pathogens via covert sexual reproduction. Cell Host Microbe 8, 86-99.

Hill, W.G., Robertson, A., 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8, 269-294.

Holmes, E.C., 1991. Different Rates of Substitution May Produce Different Phylogenies of the Eutherian Mammals. J. Mol. Evol. 33, 209-215.

Hong, S.T., Steele, P.E., Cushion, M.T., Walzer, P.D., Stringer, S.L., Stringer, J.R., 1990. *Pneumocystis carinii* karyotypes. J. Clin. Microbiol. 28, 1785-1795.

Hughes, W.T., 1982. Natural mode of acquisition for *de novo* infection with *Pneumocystis carinii*. J. Infect. Dis. 145, 842-848.

Hugot, J. P., Demanche, C., Barriel, V., Dei-Cas, E., Guillot, J., 2003. Phylogenetic systematics and evolution of primate-derived *Pneumocystis* based on mitochondrial or nuclear DNA sequence comparison. Syst. Biol. 52, 735-744.

Icenhour, C.R., Kottom, T.J., Limper, A.H., 2003. Evidence for a melanin cell wall component in *Pneumocystis carinii*. Infect. Immun. 71, 5360-5363.

Icenhour, C.R., Arnold, J., Medvedovic, M., Cushion, M.T., 2006a. Competitive coexistence of two *Pneumocystis* species. Infect. Genet. Evol. 6, 177-186.

Icenhour, C.R., Kottom, T.J., Limper, A.H., 2006b. *Pneumocystis* melanins confer enhanced organism viability. Eukaryot. Cell 5, 916-923.

Inglis, D.O., Voorhies, M., Hocking Murray, D.R., Sil, A., 2013. Comparative transcriptomics of infectious spores from the fungal pathogen *Histoplasma capsulatum* reveals a core set of transcripts that specify infectious and pathogenic states. *Eukaryot. Cel*l 12, 828-852.

Jain, N., Li, L., Hsueh, Y.P., Guerrero, A., Heitman, J., Goldman, D.L., Fries, B.C., 2009. Loss of allergen 1 confers a hypervirulent phenotype that resembles mucoid switch variants of *Cryptococcus neoformans*. Infect. Immun. 77, 128-140.

Jarboui, M.A., Mseddi, F., Sellami, H., Sellami, A., Makni, F., Ayadi, A., 2013. Genetic diversity of *Pneumocystis jirovecii* strains based on sequence variation of different DNA region. Med. Mycol. 51, 561-567.

Jarne, P., 1995. Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. Genet. Res. 65, 193-207.

Kaneshiro, E. S., Ellis, J. E., Jayasimhulu, K. and Beach, D. H., 1994 Evidence for the presence of "metabolic sterols" in *Pneumocystis*: identification and initial characterization of *Pneumocystis carinii* sterols. J. Eukar. Microbiol. 41,78-85.

Kaneshiro, E.S., Amit Z., Swonger, M.M., Kreishman G.P., Brooks, E.E., Kreishman, M., Jayasimhulu, K., Parish, E.J., Sun, H., Kizito S.A., Beach, D.H., 1999. Pneumocysterol [(24Z)-ethylidenelanost-8-en-3β-ol], a rare sterol detected in the opportunistic pathogen Pneumocystis carinii hominis*:* Structural identity and chemical synthesis. Proceed. Nat. Acad. Sciences 96, 97-102.

Kasuga, T., White, T.J., Taylor, J.W., 2002. Estimation of nucleotide substitution rates in eurotiomycete fungi. Mol. Biol. Evol. 19, 2318-2324.

Keeling, P.J., Slamovits, C.H., 2005. Causes and effects of nuclear genome reduction. Curr. Opin. Genet. Dev. 15, 601-608.

Keeling, P.J., Corradi, N., Morrison, H.G., Haag, K.L., Ebert, D., Weiss, L.M., Akiyoshi, D.E., Tzipori, S., 2010. The reduced genome of the parasitic microsporidian *Enterocytozoon bieneusi* lacks genes for core carbon metabolism. Gen. Biol. Evol. 2, 304-309.

Keely, S. P., Fischer, J.M., Stringer, J.R., 2003a. Evolution and speciation of *Pneumocystis*. J. Eukaryot. Microbiol. 50 suppl, 624-626.

Keely, S.P., Cushion, M.T, Stringer S.J., Stringer, J.R., 2003b. Diversity at the locus associated with transcription of a variable surface antigen of *Pneumocystis carinii* as an index of population structure and dynamics in infected rats. Infect. Immun. 71, 47-60.

Keely, S.P., Fischer, J.M., Cushion, M.T., Stringer, J.R., 2004a. Phylogenetic identification of *Pneumocystis murina* sp. nov., a new species in laboratory mice. Microbiology 150, 1153-1165.

Keely, S.P., Renauld, H., Wakefield, A.E., Cushion, M.T., Smulian, A.G-, Fosker, N., Fraser, A., Harris, D., Murphy, L., Price, C., Quail, M.A., Seeger, K., Sharp, S., Tindal, C.J., Warren, T., Zuiderwijk, E., Barrell, B.G., Stringer, J.R., Hall, N., 2005. Gene arrays at *Pneumocystis carinii* telomeres. Genetics 170, 1589-600.

Keely, S.P., Linke M.J., Cushion M.T., Stringer, J.R., 2007. *Pneumocystis murina* MSG gene family and the structure of the locus associated with its transcription. Fung. Genet. Biol. 44, 905-919.

Keely, S.P., Stringer, J.R., 2009. Complexity of the MSG gene family of *Pneumocystis carinii*. BMC Genomics 10, 367.

Köhler, J.R., Hube, B., Puccia, R., Casadevall, A., Perfect, J.R., 2017. Fungi that Infect Humans. Microbiol Spectr. 5, doi: 10.1128-microbiolspec.FUNK-0014-2016.

Kottom, T.J., Limper, A.H., 2000. Cell wall assembly by *Pneumocystis carinii*. Evidence for a unique *gsc-1* subunit mediating beta -1,3-glucan deposition. J. Biol. Chem. 275, 40628-40634.

Kovacs, J.A., Powell, F., Edman, J.C., Lundgren, B., Martinez, A., Drew, B., Angus, W., 1993. Multiple genes encode the major surface glycoprotein of *Pneumocystis carinii*. J. Biol. Chem. 268, 6034-6040.

Kuo, C.H., Ochman, H., 2009. Deletional bias across the three domains of life. Gen. Biol. Evol. 1, 145-152.

Kutty, G., Ma, L., Kovacs, J.A., 2001. Characterization of the expression site of the major surface glycoprotein of human-derived *Pneumocystis carinii*. Mol. Microbiol. 42, 183–193.

Kutty, G., Maldarelli, F., Achaz, G., Kovacs, J.A., 2008. Variation in the major surface glycoprotein genes in *Pneumocystis jirovecii*. J. Infect. Dis. 198, 741-749.

Kutty, G., Davis, A. S., Ma, L., Taubenberger, J.K., Kovacs, J.A., 2015. *Pneumocystis* encodes a functional endo-beta-1,3-glucanase that is expressed exclusively in cysts. J. Infect. Dis. 211, 719-728.

Kutty, G., Davis, A.S., Ferreyra, G.A., Qiu, J., Huang, D.W., Sassi, M., Bishop, L., Handley, G., Sherman, B., Lempicki, R., Kovacs, J.A., 2016. Beta-glucans are masked but contribute to pulmonary inflammation during *Pneumocystis* pneumonia. J. Infect. Dis. 214, 782-791.

Laakkonen, J., 1998. *Pneumocystis carinii* in wildlife. Int. J. Parasitol. 28, 241-252.

Latgé, J.P., 2007. The cell wall: a carbohydrate armour for the fungal cell. Mol. Microbiol. 66, 279-290.

Latinne, A., Beze, F., Delhaes, L., Pottier, M., Gantois, N., Nguyen, J., Blasdell, K., Dei-Cas, E., Morand, S., Chabé, M., 2017. Genetic diversity and evolution of *Pneumocystis* fungi infecting wild Southeast Asian murid rodents. Parasitology 1-16. doi: 10.1017/S0031182017001883.

Latouche, S., Totet, A., Lacube, P., Bolognini, J., Nevez, G., Roux, P., 2001. Development of an RT-PCR on the heat shock protein 70 gene for viability detection of *Pneumocystis carinii* f. sp. *hominis* in patients with pneumocystosis and in air sample. J. Eukaryot. Microbiol. 48, S176-S177.

Le Gal, S., Damiani, C., Rouillé, A., Grall, A., Tréguer, L., Virmaux, M., Moalic, E., Quinio, D., Moal, M.C., Berthou, C., Saliou, P., Le Meur, Y., Totet, A., Nevez, G., 2012. A cluster of *Pneumocystis* infections among renal transplant recipients: molecular evidence of colonized patients as potential infectious sources of *Pneumocystis jirovecii*. Clin. Infect Dis. 54, e62-71.

Lee, C.H., Lu, J.J., Bartlett, M.S., Durkin, M.M., Liu, T.H., Wang, J., Jiang, B., Smith, J.W., 1993. Nucleotide sequence variation in *Pneumocystis carinii* strains that infect humans. J. Clin. Microbiol. 31, 754-757.

Liu, Y., Leibowitz, M.J., 1993. Variation and in vitro splicing of group I introns in rRNA genes of *Pneumocystis carinii*. Nucl. Ac. Res. 21, 2415-2421.

Liu, Y., Tidwell, R.R., Leibowitz, M.J., 1994. Inhibition of in vitro splicing of a group I intron of *Pneumocystis carinii*. J. Eukaryot. Microbiol. 41, 31-38.

Liu, Y., Leigh, J.W., Brinkmann, H., Cushion, M.T., Rodriguez-Ezpeleta, N., Philippe, H., Lang, B.F., 2009. Phylogenomic analyses support the monophyly of Taphrinomycotina, including *Schizosaccharomyces* fission yeasts. Mol. Biol. Evol. 26, 27-34.

Lu, J.-J., Bartlett, M.S., Smith, J.W., Lee, C.-H., 1995. Typing of *Pneumocystis carinii* strains with type-specific oligonucleotide probes derived from nucleotide sequences of internal transcribed spacers of rRNA genes. J. Clin. Microbiol. 33, 2973-2977.

Lugli, E.B., Bampton, E.T., Ferguson, D.J., Wakefield, A.E., 1999. Cell surface protease PRT1 identified in the fungal pathogen *Pneumocystis carinii*. Mol. Microbiol. 31, 1723-1733.

Lundgren, B., Cotton, R., Lundgren, J.D., Edman, J.C., Kovacs, J.A., 1990. Identification of *Pneumocystis carinii* chromosomes and mapping of five genes. Infect. Immun. 58, 1705-1710.

Ma, L., Kutty, G., Jia, Q., Imamichi, H., Huang, L., Atzori, C., Beckers, P., Groner, G., Beard, C.B., Kovacs, J.A., 2002. Analysis of variation in tandem repeats in the intron of the major surface glycoprotein expression site of the human form of *Pneumocystis carinii*. J. Infect. Dis. 186, 1647-1654.

Ma, L., Huang da, W., Cuomo, C.A., Sykes, S., Fantoni, G., Das, B., Sherman, B.T., Yang, J., Huber, C., Xia, Y., Davey, E., Kutty, G., Bishop, L., Sassi, M., Lempicki, R.A., Kovacs, J.A., 2013. "Sequencing and characterization of the complete mitochondrial genomes of three *Pneumocystis* species provide new insights into divergence between human and rodent *Pneumocystis*. Faseb J. 27, 1962-1972.

Ma, L., Chen, Z., Huang da, W., Kutty, G., Ishihara, M., Wang, H., Abouelleil, A., Bishop, L., Davey, E., Deng, R., Deng, X., Fan, L., Fantoni, G., Fitzgerald, M., Gogineni, E., Goldberg, J.M., Handley, G., Hu, X., Huber, C., Jiao, X., Jones, K., Levin, J.Z., Liu, Y., Macdonald, P., Melnikov, A., Raley, C., Sassi, M., Sherman, B.T., Song, X., Sykes, S., Tran, B., Walsh, L., Xia, Y., Yang, J., Young, S., Zeng, Q., Zheng, X., Stephens, R., Nusbaum, C., Birren, B.W., Azadi, P., Lempicki, R.A., Cuomo, C.A., Kovacs, J.A., 2016a. Genome analysis of three *Pneumocystis* species reveals adaptation mechanisms to life exclusively in mammalian hosts. Nat. Commun. 7, 10740.

Ma, L., Raley, C., Zheng, X., Kutty, G., Gogineni, E., Sherman, B.T., Sun, Q., Chen, X., Skelly, T., Jones, K., Stephens, R., Zhou, B., Lau, W., Johnson, C., Imamichi, T., Jiang, M., Dewar, R., Lempicki, R.A., Tran, B., Kovacs, J.A., Huang da, W., 2016b. Distinguishing highly similar gene isoforms with a clustering-based bioinformatics analysis of PacBio single-molecule long reads. BioData mining 9, 13.

Maher, N.H., Vermund, S., Welsh, D.A., Dillon, H.K., Awooda, A., Unnasch, T.R., 2001. Development and characterization of a molecular viability assay for *Pneumocystis carinii* f sp *hominis*. J. Infect. Dis. 183, 1825-1827.

Maitte, C., Leterrier, M., Le Pape, P., Miegeville, M., Morio, F., 2013. Multilocus sequence typing of *Pneumocystis jirovecii* from clinical samples: how many and which loci should be used? J. Clin. Microbiol. 51, 2843-2849.

Matos, O., Esteves, F., 2010. *Pneumocystis jirovecii* multilocus gene sequencing: findings and implications. Fut. Microbiol. 5, 1257-1267.

Martinez, A., Aliouat, E.M., Standaert-Vitse, A., Werkmeister, E., Pottier, M., Pincon, C., Dei-Cas, E., Aliouat-Deni,s C.M., 2011. Ploidy of cell-sorted trophic and cystic forms of *Pneumocystis carinii*. PLoS One 6, e20935.

Martinez, A., Halliez, M.C., Aliouat, E.M., Chabé, M., Standaert-Vitse, A., Frealle, E., Gantois, N., Pottier, M., Pinon, A., Dei-Cas, E., Aliouat-Denis, C.M., 2013. Growth and airborne transmission of cell-sorted life cycle stages of *Pneumocystis carinii*. PLoS One 8, e79958.

Matsumoto, Y., Yoshida, Y., 1984. Sporogony in *Pneumocystis carinii*: synaptonemal complexes and meiotic nuclear divisions observed in precysts. J. Protozool. 31, 420-428.

Mazars, E., Guyot, K., Durand, I., DeiCas, E., Boucher, S., BenAbderrazak, S., Banuls, A.L., Tibayrenc, M., Camus, D., 1997. Isoenzyme diversity in *Pneumocystis carinii* from rats, mice, and rabbits. J. Infect. Dis. 175, 655-660.

Medoff, G., Painter, A., Kobayashi, GS., 1987. Mycelial- to yeast-phase transitions of the dimorphic fungi *Blastomyces dermatitidis* and *Paracoccidioides brasiliensis*. J. Bacteriol. 169, 4055-4060.

Montes-Cano, M.A., Chabé, M., Fontillon-Alberdi, M., de-Lahorra, C., Respaldiza, N., Medrano, F.J., Varela, J.M., Dei-Cas, E., Calderon, E.J., 2009. Vertical transmission of *Pneumocystis jirovecii* in humans. Emerg. Infect. Dis. 15, 125-127.

Nahimana, A., Francioli, P., Blanc, D.S., Bille, J., Wakefield, A.E., Hauser, P.M., 2000a. Determination of the copy number of the nuclear rDNA and beta-tubulin genes of *Pneumocystis carinii* f. sp *hominis* using PCR multicompetitors. J. Eukar. Microbiol. 47, 368-372.

Nahimana, A., Blanc, D.S., Francioli, P., Bille, J., Hauser, P.M., 2000b. Typing of *Pneumocystis carinii* f. sp. *hominis* by PCR-SSCP to indicate a high frequency of co-infections. J. Med. Microbiol. 49, 753-758.

Nahimana, A., Cushion ,M.T., Blanc, D.S., Hauser, P.M., 2001. Rapid PCR-single-strand conformation polymorphism method to differentiate and estimate relative abundance of *Pneumocystis carinii* special forms infecting rats. J. Clin. Microbiol. 39, 4563-4565.

Nagy, L.G., Ohm, R.A., Kovács, G.M., Floudas, D., Riley, R., Gácser, A., Sipiczki, M., Davis, J.M., Doty, S.L., de Hoog, G.S., Lang, B.F., Spatafora, J.W., Martin, F.M., Grigoriev, I.V., Hibbett, D.S., 2014. Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. Nat. Commun. 18, 4471.

Nagy, G., 2017. Complex Multicellular Life with 5,500 Genes. Curr. Biol. 27, R609-R612.

Nguyen, T.A., Cissé, O.H., Yun Wong, J., Zheng, P., Hewitt, D., Nowrousian, M., Stajich, J.E., Jedd, G., 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota. Nat. Commun. 8, 14444.

Noble, L.M., Andrianopoulos, A., 2013. Fungal genes in context: genome architecture reflects regulatory complexity and function. Gen. Biol. Evol. 5, 1336-1352.

Nollstadt, K.H., Powles, M.A., Fujioka, H., Aikawa, M., Schmatz, D.M., 1994. Use of beta-1,3-glucan-specific antibody to study the cyst wall of *Pneumocystis carinii* and effects of pneumocandin B0 analog L-733,560. Ant. Ag. Chemoth. 38, 2258-2265.

O'Malley, M.A., Wideman, J.G., Ruiz-Trillo, I., 2016. Losing complexity: the role of simplification in macroevolution. Trends Ecol. Evol. 31, 608-621.

Otto, S.P., Lenormand, T., 2002. Resolving the paradox of sex and recombination. Nat. Rev. Genet. 3, 252-261.

Palmer, R.J., Settnes, O.P., Lodal, J., Wakefield, A.E., 2000. Population structure of rat-derived *Pneumocystis carinii* in Danish wild rats. Appl. Environ. Microbiol. 66, 4954-4961.

Papkou, A., Gokhale, C.S., Traulsen, A., Schulenburg, H., 2016. Host-parasite coevolution: why changing population size matters. Zoology 119, 330-338.

Parker, I.M., Gilbert, G.S., 2004. The evolutionary ecology of novel plant-pathogen interactions. An. Rev. Ecol. Evol. Syst. 35, 675-700.

Parobek, C.M., Jiang, L.Y., Patel, J.C., Alvarez-Martinez, M.J., Miro, J.M., Worodria, W., Andama, A., Fong, S., Huang, L., Meshnick, S.R., Taylor, S.M., Juliano, J.J., 2014. Multilocus microsatellite genotyping array for investigation of genetic epidemiology of *Pneumocystis jirovecii*. J. Clin. Microbiol. 52, 1391-1399.

Peters, S.E., English, K., Laakkonen, J., Gurnell, J., 1994. DNA analysis of *Pneumocystis carinii* infecting Finnish and English shrews. J. Eukaryot. Microbiol. 41, S108.

Peterson, J.C, Cushion, M.T., 2005. *Pneumocystis*: not just pneumonia. Curr. Opin. Microbiol. 8, 393-8

Pollak, E., 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117, 353-360.

Porollo, A., Sesterhenn, T.M., Collins, M.S., Welge, J.A., Cushion, M.T., 2014. Comparative genomics of *Pneumocystis* species suggests the absence of genes for myo-inositol synthesis and reliance on inositol transport and metabolism. mBio 5, e01834.

Poulin, R., Krasnov, B.R., Morand, S., 2006. Patterns of host specificity in parasites exploiting small mammals. Micromammals and Macroparasites: From Evolutionary Ecology to Management, 233-256.

Pringle, A, Baker, D.M., Platt, J.L., Wares, J.P., Latgé, J.P., Taylor, J.W., 2005. Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. Evolution 59, 1886-1899.

Pritchard, J. K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Restrepo, S., Tabima, J.F., M.F. Mideros, N.J. Grunwald, Matute, D.R., 2014. Speciation in fungal and oomycete plant pathogens. Annu. Rev. Phytopathol. 52. 289-316.

Richard, S., Almeida, J.M.G. C.F., Cissé, O.H., Luraschi, A., Nielsen, O., Pagni, M., Hauser, P.M., 2018. Functional and expression analyses of the *Pneumocystis MAT* genes suggest obligate sexuality through primary homothallism within host's lungs. mBio 9, e02201-17.

Rieseberg, L.H., 2001. Chromosomal rearrangements and speciation. Trend. Ecol. Evol. 16, 351-358.

Rhind N, Chen Z, Yassour M, Thompson DA, Haas BJ, Habib N, Wapinski I, Roy S, Lin MF, Heiman DI, Young SK, Furuya K, Guo Y, Pidoux A, Chen HM, Robbertse B, Goldberg JM, Aoki K, Bayne EH, Berlin AM, Desjardins CA, Dobbs E, Dukaj L, Fan L, FitzGerald MG, French C, Gujja S, Hansen K, Keifenheim D, Levin JZ, Mosher RA, Müller CA, Pfiffner J, Priest M, Russ C, Smialowska A, Swoboda P, Sykes SM, Vaughn M, Vengrova S, Yoder R, Zeng Q, Allshire R, Baulcombe D, Birren BW, Brown W, Ekwall K, Kellis M, Leatherwood J, Levin H, Margalit H, Martienssen R, Nieduszynski CA, Spatafora JW, Friedman N, Dalgaard JZ, Baumann P, Niki H, Regev A, Nusbaum C. 2011. Comparative functional genomics of the fission yeasts. Science 332, 930–936.

Rhodes, J., Beale, M.A,, Vanhove, M., Jarvis, J.N., Kannambath, S., Simpson, J.A., Ryan, A., Meintjes, G., Harrison, T.S., Fisher, M.C., Bicancic, T., 2017. A population genomics approach to assessing the genetic basis of within-host microevolution underlying recurrent cryptococcal meningitis infection. G3 (Bethes.) 3, 1165-1176.

Roy S.W., Gilbert W., 2005. Rates of intron loss and gain: Implications for early eukaryotic evolution. Proc Natl Acad Sci USA. 102, 5773–5778.

Sanchez, C.A., Chabé ,M., Aliouat, E.M., Durand-Joly, I., Gantois, N., Conseil, V., Lo´ pez, C., Duriez, T., Dei-Cas, E., Vargas, S.L., 2007. Exploring transplacental transmission of *Pneumocystis oryctolagi* in first-time pregnant and multiparous rabbit does. Med. Mycol. 45, 701–707.

Sasal, P., Trouve, S., Muller-Graf , C., Morand, S., 1999. Specificity and host predictability: a comparative analysis among monogenean parasites of fish. J. An. Ecol. 68, 437-444.

Schaffzin, J.K., Stringer, J.R., 2000. The major surface glycoprotein expression sites of two special forms of rat *Pneumocystis carinii* differ in structure. J. Infect. Dis. 181, 1729-1739.

Schmid-Siegert, E, Richard, S., Luraschi, A., Mühlethaler, K., Pagni, M., Hauser, P.M., 2017. Mechanisms of surface antigenic variation in the human pathogenic fungus *Pneumocystis jirovecii*. mBio 8, e01470-17.

Schoch, C.L. Sung, G.H., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., Hofstetter, V., Robbertse, B., Matheny, P.B., Kauff, F., Wang, Z., Gueidan, C., Andrie, R.M., Trippe, K., Ciufetti, L.M., Wynns, A., Fraker, E., Hodkinson, B.P., Bonito, G., Groenewald, J.Z., Arzanlou, M., de Hoog, G.S.,, Crous, P.W., Hewitt, D., Pfister, D.H., Peterson, K., Gryzenhout, M., Wingfield, M.J., Aptroot, A., Suh, S.O., Blackwell, M., Hillis, D.M., Griffith, G.W., Castlebury, L.A., Rossman, A.Y., Lumbsch, H.T., Lücking, R., Büdel, B., Rauhut, A., Diederich, P., Ertz, D., Geiser, D.M., Hosaka, K., Inderbitzin, P., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Mostert, L., O'Donnell, K., Sipman, H., Rogers, J.D., Shoemaker, R.A., Sugiyama, J., Summerbell, R.C., Untereiner, W., Johnston, P.R., Stenroos, S., Zuccaro, A., Dyer, P.S., Crittenden, P.D., Cole, M.S., Hansen, K., Trappe, J.M,, Yahr. R,, Lutzoni. F,, Spatafora. J,W., 2009. The Ascomycota

tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Systematic biology, 58, 224-239.

Schulze-Lefert, P., Panstruga, R., 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trend. Plant Sci. 16, 117-125.

Sesterhenn, T.M., Slaven, B.E., Keely, S.P., Smulian, A.G., Lang, B.F., Cushion, M.T., 2010. Sequence and structure of the linear mitochondrial genome of *Pneumocystis carinii*." Mol. Gen. Gen. 283, 63-72.

Shannon C. 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27:379–423.

Simpson, E.H., 1949. Measurement of diversity. Nature, 163, 688

Slaven, B.E., Meller, J., Porollo, A., Sesterhenn, T., Smulian, A.G.,. Cushion, M.T., 2006. Draft assembly and annotation of the *Pneumocystis carinii* genome. J. Eukaryot. Microbiol. 53 suppl, S89-91.

Smit, A.F.A,. Hubley, R., Green, P., 2013. *RepeatMasker Open-4.0*. <http: www.repeatmasker.org>.

Slamovits, C.H., Fast, N.M., Law, J.S., Keeling, P.J., 2004. Genome compaction and stability in microsporidian intracellular parasites. Curr. Biol, 14, 891-896.

Stajich, J.E., Dietrich, F.S., Roy, S.W., 2007. Comparative genomic analysis of fungal genomes reveals intron-rich ancestors. Gen. Biol. 8, R223.

Stringer, J.R., 1996. *Pneumocystis carinii*: what is it, exactly? Clinical Microbiol. Rev. 9, 489-498.

Stringer, J.R., 2007. Antigenic variation in *Pneumocystis*. J. Eukaryot. Microbiol. 54, 8 -13.

Stringer, J.R., Stringer, S.L., Zhang, J, Baughman, R., Smulian, A.G., Cushion, M.T., 1993. Molecular genetic distinction of *Pneumocystis carinii* from rats and humans. J. Eukaryot. Microbiol. 40, 733-741.

Stringer, J.R., Cushion, M.T., 1998. The genome of *Pneumocystis carinii*. FEMS Immunol. Med. Microbiol. 22, 15-26.

Sun, L., Huang, M., Wang, J., Xue, F., Hong, C., Guo, Z., Gu, J., 2015. Genotyping of *Pneumocystis jirovecii* isolates from human immunodeficiency virus-negative patients in China. Infect. Genet. Evol. 31, 209-215.

Sunkin, S.M., Stringer, S.L., Stringer, J.R., 1994. A tandem repeat of rat-derived *Pneumocystis carinii* genes encoding the major surface glycoprotein. J. Eukaryot. Microbiol. 41, 292-300.

Sunkin, S.M., Stringer, J.R, 1996. Translocation of surface antigen genes to a unique telomeric expression site in *Pneumocystis carinii*. Mol. Microbiol., 19, 283-295.

Sugiyama, J., Hosaka, K., Suh, S.O., 2006. Early diverging Ascomycota: phylogenetic divergence and related evolutionary enigmas. Mycologia 98, 996-1005.

Tanabe, K., Fuchimoto, M., Egawa, K., Nakamura, Y., 1988. Use of *Pneumocystis carinii* genomic DNA clones for DNA hybridization analysis of infected human lungs. J. Infect. Dis. 157, 593-596.

Tang, X., Bartlett, M.S., Smith, J.W., Lu, J.J., Lee, C.H., 1998. Determination of copy number of rRNA genes in *Pneumocystis carinii* f. sp. *hominis*. J .Clin. Microbiol. 36, 2491-2494.

Thomas, C.F., Jr., Leof, E.B., Limper, A.H., 1999. Analysis of *Pneumocystis carinii* introns. Infect. Immun. 67, 6157-6160.

Tibayrenc, M., Ayala, F.J., 2012. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc. Natl. Acad. Sci. USA 109, E3305-3313.

Tibayrenc, M., Ayala, F.J., 2014. *Cryptosporidium, Giardia, Cryptococcus, Pneumocystis* genetic gariability: cryptic biological species or clonal near-clades? PLoS Pathog. 10, e1003908.

Tsai, I.J., Tanaka, E., Masuya, H., Tanaka, R., Hirooka, Y., Endoh, R., Sahashi, N., Kikuchi, T., 2104. Comparative genomics of *Taphrina* fungi causing varying degrees of tumorous deformity in plants. *Gen. Biol. Evol.* 6, 861-872.

Valero, C., Buitrago, M.J., Gits-Muselli, M., Benazra, M., Sturny-Leclère, A., Hamane, S., Guigue, N., Bretagne, S., Alanio, A., 2016. Copy number variation of mitochondrial DNA genes in *Pneumocystis jirovecii* according to the fungal load in BAL specimens. Front. Microbiol. 7, 1413.

van der Does, H. C., Rep, M., 2007. Virulence genes and the evolution of host specificity in plant-pathogenic fungi. Mol. Plant-Microbe Interact. 20, 1175-1182.

Vasquez, J., Smulian, A.G., Linke, M.J., Cushion, M.T., 1996. Antigenic differences associated with genetically distinct *Pneumocystis carinii* from rats. Infect. Immun. 64, 290-297.

Venema, J., Tollervey, D., 1999. Ribosome synthesis in *Saccharomyces cerevisiae*. Annu. Rev. Genet. 33, 261-311.

Vestereng, V.H., Kovacs, J.A., 2004. Inability of *Pneumocystis* organisms to incorporate bromodeoxyuridine suggests the absence of a salvage pathway for thymidine. Microbiol. 150, 1179-1182.

Vohra, P.K., Park, J.G., Sanyal, B., Thomas, C.F., Jr., 2004. Expression analysis of PCSTE3, a putative pheromone receptor from the lung pathogenic fungus *Pneumocystis carinii*. Biochem. Biophys. Res. Commun. 319, 193-199.

Wada, M., Sunkin, S.M., Stringer, J.R., Nakamura, Y., 1995. Antigenic variation by positional control of major surface glycoprotein gene expression in *Pneumocystis carinii*. J. Infect. Dis. 171,1563-1568.

Wakefield, A.E., 1996. DNA sequences identical to *Pneumocystis carinii* f. sp. *carinii* and *Pneumocystis carinii* f. sp. *hominis* in samples of air spora. J. Clin. Microbiol. 34, 1754-1759.

Wakefield, A.E., 1998a. Genetic heterogeneity in *Pneumocystis carinii*: an introduction. FEMS Immunol. Med. Microbiol. 22, 5-13.

Wakefield, A.E. 1998b. Genetic heterogeneity in human-derived *Pneumocystis carinii*. FEMS Immunol. Med. Microbiol. 22, 59-65.

Wakefield, A.E., Fritscher, C.C., Malin, A.S., Gwanzura, L., Hughes, W.T., Miller, R.F. 1994. Genetic diversity in human-derived *Pneumocystis carinii* isolates from four geographical locations shown by analysis of mitochondrial rRNA gene sequences. J. Clin. Microbiol. 32, 2959-2961.

Wernegreen, J.J., 2015. Endosymbiont evolution: predictions from theory and surprises from genomes. Ann. NY Acad .Sci. 1360, 16-35.

Wolf, Y.I., Koonin, E.V., 2013. Genome reduction as the dominant mode of evolution. Bioessays 35, 829–837.

Wood, V,, Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat , N., Hayles. J,, Baker. S,, Basham, D., Bowman, S,, Brooks, K,. Brown, D., Brown, .S, Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Feltwell ,T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E.J,, Hunt. S,, Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neil, S., Pearson, D., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M,, Squares, R., Squares, S., Stevens, K,, Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R,, Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schäfer, M., Müller-Auer, S., Gabel, C., Fuchs, M., Düsterhöft, A., Fritzc, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dréano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sánchez, M,. del Rey, F., Benito, J., Domínguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T.,

McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G., Nurse, P., 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature 415, 871-880.

Wyder, M. A., Rasch, E. M. Kaneshiro, E.S., 1998. Quantitation of absolute *Pneumocystis carinii* nuclear DNA content. Trophic and cystic forms isolated from infected rat lungs are haploid organisms. J. Eukaryot. Microbiol. 45, 233-239.

Ye, D., Lee, C.H., Queener, S.F., 2001. Differential splicing of *Pneumocystis carinii* f. sp. *carinii* inosine 5'-monophosphate dehydrogenase pre-mRNA. Gene 263, 151-158.

Yoganathan, T., Lin, H., Buck, G.A., 1989. An electrophoretic karyotype and assignment of ribosomal genes to resolved chromosomes of *Pneumocystis carinii*. Mol. Microbiol. 3, 1473-1480.

Zhu, T., Niu, D.K., 2013. Mechanisms of intron loss and gain in the fission seast *Schizosaccharomyces*. PLOS ONE 8, e61683.

LEGENDS FIGURES

Figure 1 | Cell cycle

The whole cell cycle of *Pneumocystis* species would take place within the host's lungs, airborne asci ensuring transmission to new hosts. The cycle is thought to include two phases: sexual and asexual. The trophic forms tightly adhere to the host's alveolar epithelial pneumocytes type I, whereas asci are generally localized within the alveolar lumen. The ring shown in green might allow the formation of a rent upon contact with humidity and so the release of the ascospores. This ring may correspond to the parentheses-like structure visible on Figure 2. This Figure does not include new features relatively to models previously proposed.

Figure 2 | Cluster of *P. jirovecii* asci

Cluster of *P. jirovecii* asci stained with Grocott's Methenamine silver (Churukian and Schenk, 1977) within a patient's bronchoalveolar lavage. The structures darker than the rest of the wall on each ascus are the parentheses-like structure (picture from the Institute of Microbiology, Lausanne University Hospital).

Figure 3 | Genome composition of *Pneumocystis* and related fungi

Protein coding genes, intergenic spaces, and intron positions were obtained from NCBI (https://www.ncbi.nlm.nih.gov/, last accessed 2018-03-20). Curated *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* intron data were extracted respectively from Pombase database (Wood et al., 2002; https://www.pombase.org/downloads/intron-data, last accessed 2018-03-20) and *Saccharomyces* Genome database (Cherry et al., 1998; https://www.yeastgenome.org, last accessed 2018-03-20). Repeats include DNA transposons, retrotransposons, and simple low complexity repeats proportions as roughly estimated using

RepeatMasker (Smit et al., 2013) and RepBase database (Bao et al., 2015). The proportions of MSGs were calculated based on data from Ma et al. (2016a). Ribosomal DNA cassettes include each three genes (rDNA) and two internal transcribed spacers (ITS): 18S rDNA-ITS1-5.8 rDNA-ITS2-26S rDNA. *S. pombe* genome encodes roughly 140 copies of a cassette of a size of 5.8 kb. In *S. cerevisiae,* ca. 150 tandem copies of a 9.1 kb cassette are present (Venema and Tollervey, 1999). In contrast, *Pneumocystis* species harbor each a single rDNA cassette of 11 kb.

Figure 4 | Graphical overview of the hypothetical evolutionary history of *Pneumocystis* species *Pneumocystis* species divergence timing has been determined elsewhere (Keely et al., 2003a; 2004a; Beimforde et al., 2014). Losses of multiple metabolic pathways, as well as contraction and expansion of specific gene families are presented (Table 2). Note that the timing and order of losses is unknown. The gain and loss of specific functions for *Pneumocystis* is inferred here to have occurred in the last most recent ancestor common of *Pneumocystis* species (MRCA) because the underlying genes are absent in the genomes of all *Pneumocystis* sequenced to date. The MSG superfamily emerged in *Pneumocystis* ancestry and displays a substantial level of lineage specific divergence (represented by blue triangles). Intron loads are similar among *Pneumocystis* species, which might suggest a common origin. The fission yeast clade diverged ~250 MYA ago (Rhind et al. 2011) and has lost most of the introns acquired from an intron rich ancestor (Roy et al., 2005; Stajich et al. 2007; Rhind et al., 2011). Although there is no dating estimates for the intron loss in fission clade, the absence of recent intron gains and the low rates of intron loss (Zhu and Niu, 2013) suggest that the majority of introns were lost before the diversification of the fission yeast clade. The colors of the lines representing the evolving species signify different nutritional modes (dark green, saprophytism; light green, gradual shift from saprophytism to the parasitism; yellow, animal parasitism). We assume that the MRCA of

Taphrinomycota subphylum was a multicellular or dimorphic saprotroph based on ancestral traits reconstruction (Schoch et al., 2009; Nguyen et al., 2017). The phylogenetic relationship presented here is consistent with published phylogenies (Liu et al., 2009; Sugiyama et al., 2006). RRM correspond to RNA binding proteins harboring an RNA recognition motif.

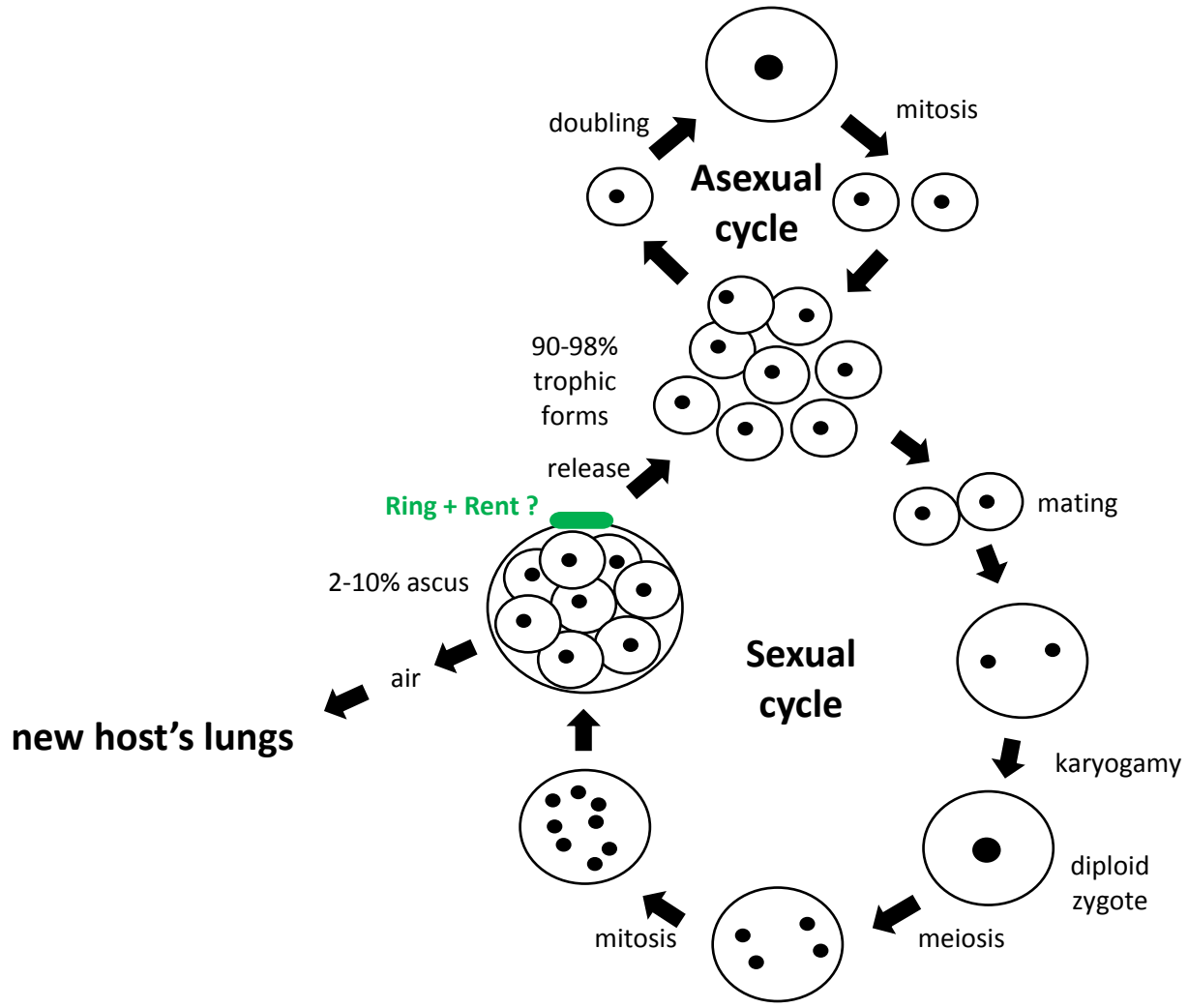
doubling
mitosis
Asexual cycle
90-98% trophic forms
release
Ring + Rent ?
mating
2-10% ascus
air
Sexual cycle
new host's lungs
karyogamy
diploid zygote
meiosis
mitosis

Figure 1.

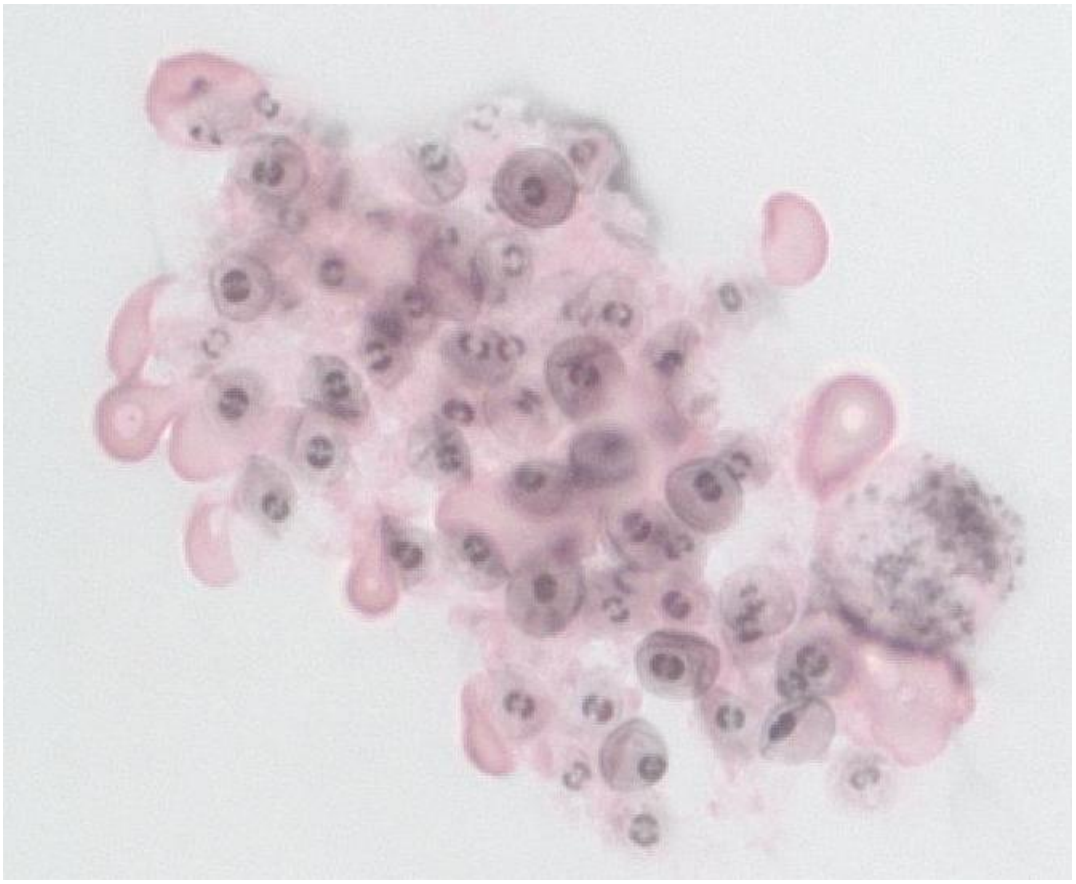

Figure 2.

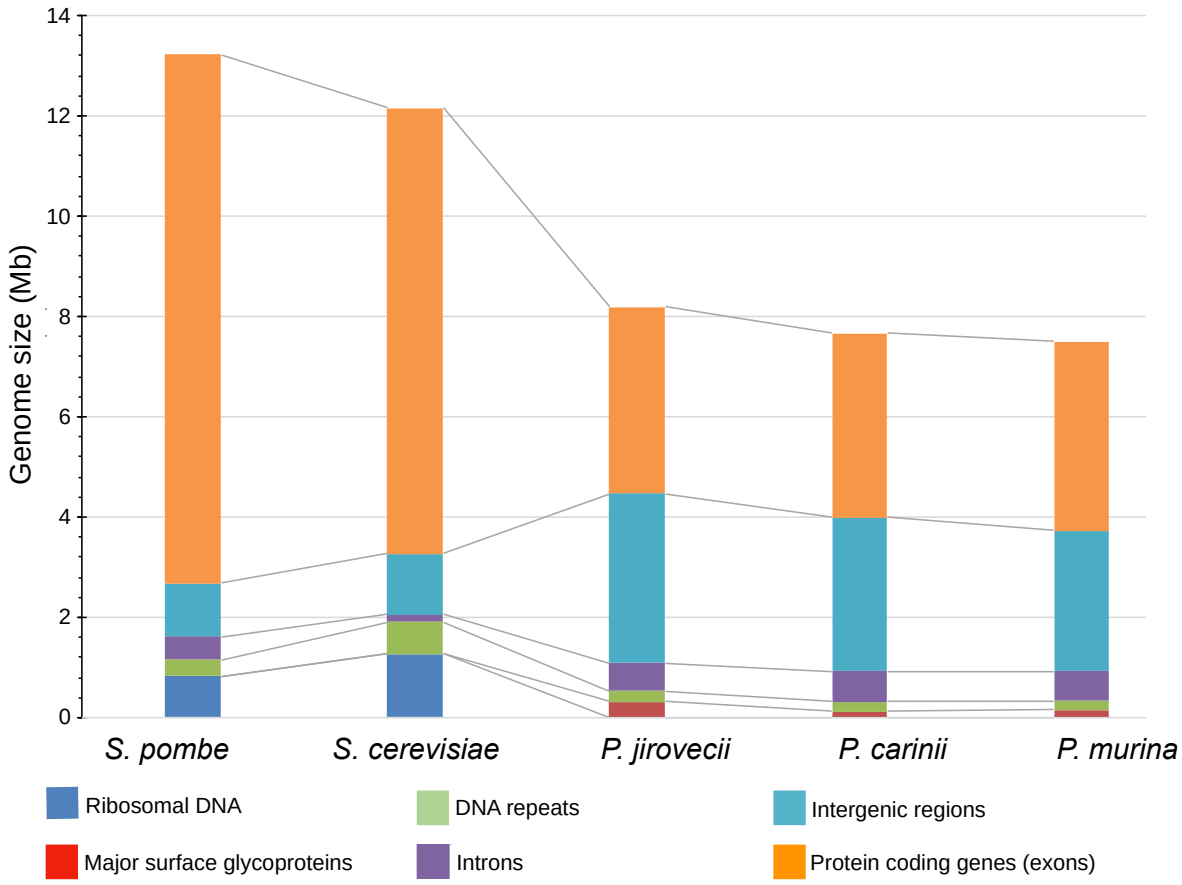
14
12
10
8
6
4
2
0
Genome size (Mb)
S. pombe
S. cerevisiae
P. jirovecii
P. carinii
P. murina
Ribosomal DNA
DNA repeats
Intergenic regions
Major surface glycoproteins
Introns
Protein coding genes (exons)

Figure 3.

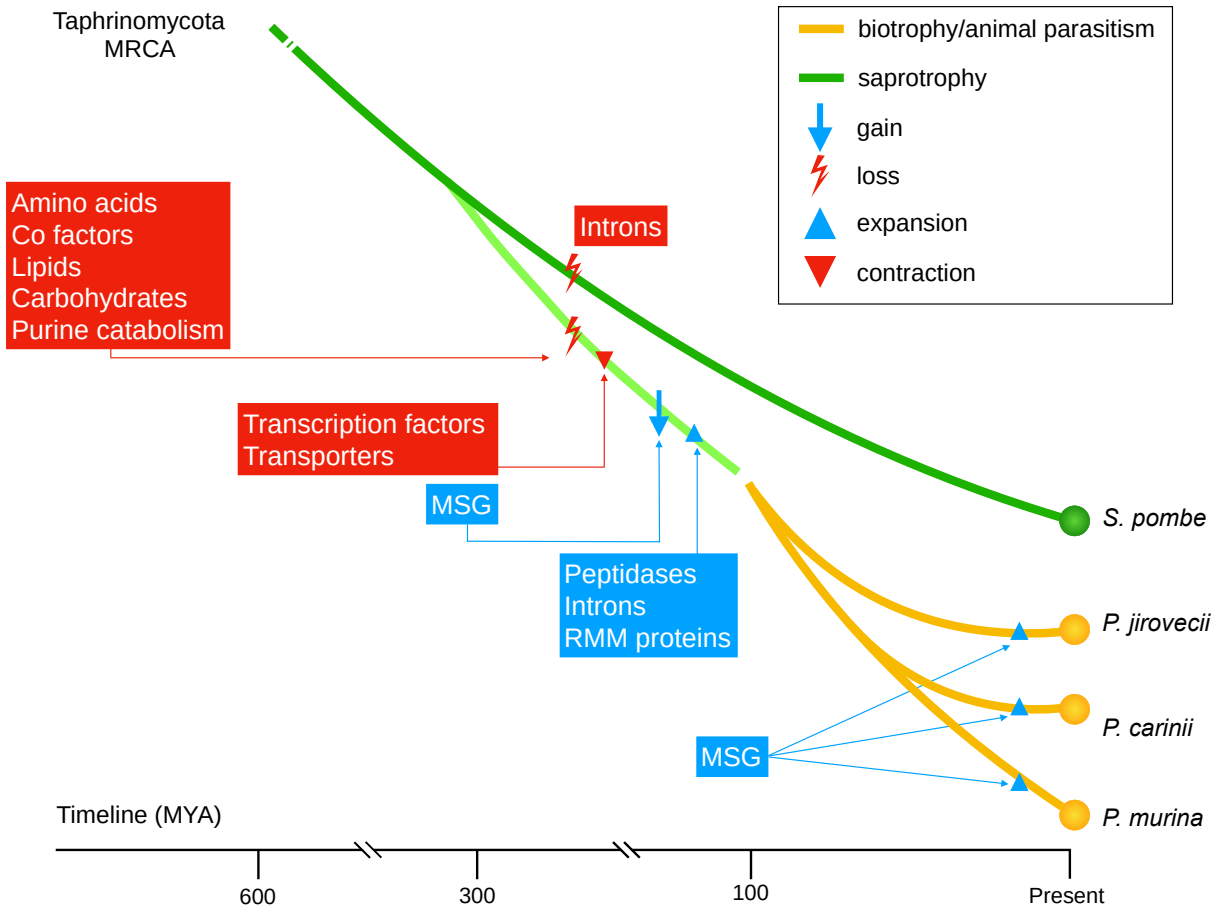
Taphrinomycota
MRCA
biotrophy/animal parasitism
saprotrophy
gain
loss
expansion
contraction
Amino acids
Co factors
Lipids
Carbohydrates
Purine catabolism
Introns
Transcription factors
Transporters
MSG
Peptidases
Introns
RMM proteins
MSG
S. pombe
P. jirovecii
P. carinii
P. murina
Timeline (MYA)
600
300
100
Present

Figure 4.

Table 1. Features of *Pneumocystis* genomes [a].

Species	Genome size Mb	%GC	Chromosomes no.	Protein coding genes no.
P. jirovecii	8.4	28.4	20	3,761
P. carinii	7.7	27.8	17	3,623
P. murina	7.5	27.0	17	3,646

[a] Genome data are from Ma et al. (2016a).

Table 2. Gene families expanded, contracted, and lost in the cellular and metabolic machineries of *Pneumocystis* species.

Event	Gene family / Pathway	Reference
Expansion		
	Major surface glycoproteins	Ma et al. 2016a; Schmid-Siegert et al. 2017
	S8 and M16 peptidases	Cissé et al., 2014 ; Ma et al. 2016a
	Proprotein convertase	Ma et al. 2016a
	Cystein rich CFEM_proteins	Ma et al. 2016a
	Kexin [a]	Ma et al. 2016a
Contraction		
	Transcription factors	Ma et al. 2016a
	Transporters	Cissé et al., 2012; Ma et al., 2016a
Loss		
	Co-factors coenzyme A, thiamine, biotin biosyntheses	Cissé et al., 2014; Ma et al., 2016a
	RNAi machinery [b]	Cissé et al., 2014
	Amino acids biosyntheses	Hauser et al., 2010; Cissé et al., 2012; Ma et al., 2016a

Steroids and *myo*-inositol biosyntheses	Porollo et al., 2014; Ma et al., 2016a
Inorganic sulfur and nitrogen assimilation	Cissé et al., 2014
Purines catabolism	Cissé et al., 2014
Nucleotide salvage pathways	Cushion et al., 2007; Cissé et al., 2012
Carbohydrate metabolism [c], lipids [d], and co-factors [e]	Kaneshiro et al., 1999; Vestereng and Kovacs, 2004; Cushion et al., 2007; Ma et al., 2016a

---

[a] Kexin protease family are only expanded in *P. carinii.* It might be involved in the processing of MSGs at the cell surface (Lugli et al., 1999).

[b] The RNA interference machinery includes the Dicer and Argonaute proteins.

[c] The lost carbohydrate pathways are those of gluconeogenesis, glyoxylate cycle, chitin, and hyper-mannose glycosylation (outer chain N-mannans).

[d] The lost lipids pathways are those of ergosterol, cholesterol, choline, ether lipids, sphingolipids, glycerol, and phosphatidylcholine.

[e] Co-factors includes co enzyme A, and vitamins H and B1.

Revised version: modified or new text is underlined.

Genomics and evolution of *Pneumocystis* species

Ousmane H. Cissé [1] and Philippe M. Hauser [2]

Correspondence: ousmane.cisse@nih.gov and Philippe.Hauser@chuv.ch

[1] Critical Care Medicine Department, NIH Clinical Center, National Institutes of Health, Bethesda, Maryland 20892, USA

[2] Institute of Microbiology, Lausanne University Hospital, Lausanne, Switzerland

TABLE OF CONTENTS

ABSTRACT

The genus *Pneumocystis* comprises highly diversified fungal species that cause severe pneumonia in individuals with a deficient immune system. These fungi infect exclusively mammals and present a strict host species specificity. These species have co-diverged with their hosts for long periods of time (> 100 MYA). Details of their biology and evolution are fragmentary mainly because of a lack of an established long-term culture system. Recent genomic advances have unlocked new areas of research and allow new hypotheses to be tested. We review here new findings of the genomic studies in relation with the evolutionary trajectory of these fungi and discuss the impact of genomic data analysis in the context of the population genetics. The combination of slow genome decay and limited expansion of specific gene families and introns reflect intimate interactions of these species with their hosts. The evolutionary adaptation of these organisms is profoundly influenced by their population structure, which in turn is determined by intrinsic features such as their self-fertilizing mating system, high host specificity, long generation times, and transmission mode. Essential key questions concerning their adaptation and speciation remain to be answered. The next cornerstone will consist in the establishment of a long-term culture system and genetic manipulation that should allow unravelling the driving forces of *Pneumocystis* species evolution.

Main text : 7482 words

BACKGROUND

History

Pneumocystis species form a group of opportunistic fungi that cause severe pulmonary infections in mammals with a deficient immune system. These organisms infect exclusively mammals. They were first described by Chagas (Chagas, 1909), and wrongly classified as special forms of trypanosomes. They were later identified as a *bona fide* separate species by the Delanoë couple at the Pasteur Institute in Paris (Delanoë and Delanoë, 1912). Their taxonomic classification remained then elusive because of a phenotypic resemblance with the protists. The issue was resolved using molecular phylogeny based on sequencing ribosomal DNA, which clearly indicated their fungal nature (Edman et al., 1988).

Phylogeny and taxonomy

Pneumocystis species belong to the subphylum of Taphrinomycotina within the Ascomycota (Eriksson, 1997; Sugiyama et al., 2006). The Taphrinomycotina subphylum is monophyletic and encompasses mostly plant-associated or soil-dwelling fungi (Liu et al., 2009). *Pneumocystis* closest relatives are *Schizosaccharomyces pombe* and *Taphrina deformans*, their common ancestor having diverged from the other Taphrinomycota members ca. 467 million years ago (MYA) (Beimforde et al., 2014).

Although all *Pneumocystis* species are ubiquitous, each mammal species can be infected with only one or two of them. Five species have been formally described so far based on the requirements of the International Code of Botanical Nomenclature (ICBN): *Pneumocystis*

jirovecii in *Homo sapiens* (Frenkel, 1999), *Pneumocystis carinii* in *Rattus norvegicus* (Frenkel, 1999), *Pneumocystis wakefieldiae* also in *Rattus norvegicus* (Cushion et al., 1993; Cushion et al., 2004), *Pneumocystis murina* in *Mus musculus* (Keely et al., 2004a), and *Pneumocystis oryctologi* in Old World rabbits (*Oryctolagus cuniculus*; Dei-cas et al., 2006). Antigenic and DNA based studies suggest the presence of distinct species also in macaques, ferrets, bats, shrews, horses, pigs, and dogs (Banerji et al., 1994; Peters et al., 1994; Christensen et al., 1996; English et al., 2001; Guillot et al., 2004).

P. jirovecii is the only species known to infect humans and has never been detected in any other animals. *P. carinii* is the best studied species because of the availability of protocols for experimental or natural infections in laboratory rats. *P. wakefieldiae* was reported either mixed with *P. carinii* (Cushion et al., 1993; Cushion, 1998; Cushion et al., 2004; Chabé et al., 2010), or alone (Palmer et al., 2000). The two latter species are different in terms of electrophoretic karyotypes, gene localization on the chromosomes, sequence identity (4-7% nucleotide divergence in seven orthologs; Cushion, 1998; Cushion et al., 2004), antigenic profiles (Vasquez et al., 1996), and major surface glycoproteins (MSG) expression (Schaffzin and Stringer, 2000). They might be competing against each other for resources when present together within the same rat (Icenhour et al., 2006a).

Species divergence

According to the evolutionary rates of several genomic loci, the radiation of the *Pneumocystis* genus occurred ca.100 MYA (Keely et al., 2003a; Keely et al., 2004a), which roughly overlaps with the radiation of the mammalian species (Holmes, 1991; dos Reis et al., 2015). *P. murina* would have diverged from *P. carinii* between 51 and 71 MYA (Keely et al., 2003a), while *P.*

carinii and *P. wakefieldiae* diverged between 15 and 22 MYA (Cushion et al., 2004; Fischer et al., 2006). The neat superposition of multiple *Pneumocystis* species phylogenetic trees with those of their respective hosts supports a co-evolution of these organisms (Guillot et al., 2001). Therefore, a plausible co-speciation scenario is that each species became physically separated from the other species, the hosts acting as barriers that led to the accumulation of genetic differences and the gradual reproductive isolation over time. The absence of gene flow or mating among the different species has been inferred based on linkage disequilibrium analysis consistent with an ancient reproductive isolation (Mazars et al., 1997; Keely et al., 2004a; Keely and Stringer, 2009). Furthermore, no evidence of hybridization was detected between *P. carinii* and *P. wakefieldiae*, even during co-infection of the same rat (Cushion, 1998; Cushion et al., 2004). However, caution is warranted because the absence of gene flow was inferred from a small set of conserved markers, which may have not allowed detecting all genetic events. Consequently, whole genome sequencing studies are necessary to validate these findings.

Life cycle

The life cycle of *Pneumocystis* organisms is still hypothetical and mostly derived from microscopic and molecular studies on *P. carinii* (Figure 1). As fungal organisms with an obligate parasitic behavior, the cycle would occur only inside host's lungs, and begin with the inhalation of infectious asci. Once inhaled, each ascus would release first eight ascospores which will evolve to what is known as trophic forms that bind to the type I pneumocytes of the alveolar epithelium. The cycle would then alternate between asexual multiplication of metabolically active trophic cells by binary fission, and sexual reproduction upon mating of two trophic cells that would culminate by the production of asci containing eight ascospores (Figure 2). Trophic

cells are amoeboid in shape and represent generally 90-98% of the populations in the infected lungs (Aliouat-Denis et al., 2009). These forms are mononuclear, 2-8 μm in diameter (Dei-Cas et al., 2004), and mostly haploid (Stringer and Cushion, 1998; Wyder et al., 1998; Martinez et al., 2011). Multiploid forms are rare and possibly caused by asymmetrical or post-mating divisions (Martinez et al., 2011). Trophic cell surface is composed of a single layer of electron dense material containing glycoproteins, but possibly no β-glucans. Indeed, the enzymes responsible for the synthesis of β-glucans and the associated endo-1,3-β glucanase are expressed almost exclusively in asci (Nollstadt et al., 1994; Kottom and Limper, 2000; Kutty et al., 2015). The presence of structural carbohydrate polymers of glucans in asci increases the physical strength of the cell wall, which might facilitate the survival outside the host. The doubling times are relatively long compared to free-living yeasts (~2 hours) and range from 1.5 to 10.5 days depending on the species (Aliouat et al., 1999; Keely et al., 2003b). The presence of a sexual cycle was initially supported by the ultrastructural observations of synaptonemal complexes (Matsumoto and Yoshida, 1984) and the expression of one pheromone receptor at the surface of *P. carinii* trophic cells (Vohra et al., 2004). Recent comparative genomic studies suggest that *Pneumocystis* species use primary homothallism (self-fertility) based on the genes number and arrangement on the chromosomes as a fusion of Plus and Minus mating type loci (Almeida et al., 2015). Thus, each strain would be able to produce asci on its own, without the need to find a compatible partner. Asci would be expelled by infected hosts and be the infectious stages because their specific inactivation or removal blocks the transmission chain (Cushion et al., 2010; Martinez et al., 2013). Consistently, recent analyses suggested that *Pneumocystis* sexuality is obligatory within host's lungs in order to complete the cell cycle and produce asci that are necessary for airborne transmission to new hosts (Richard et al., 2018). Furthermore, the

necessity of asci for transmission has been demonstrated by inhibition of the sexual cycle using echinocandins (Cushion et al., 2010), and by the fact that only purified asci could transmit the disease (Martinez et al., 2013). Recently, activation of sex-related genes upon treatment with echinocandins in RNA-seq analyses also suggested that sexuality is obligate (Cushion et al., 2018).

Transmission

Pneumocystis jirovecii pneumonia is a major public health problem with >400,000 cases per year worldwide and a mortality rate possibly as high as 80% if untreated (Brown et al., 2012). Epidemiological data for *Pneumocystis* species in animal populations are scarce, but investigations in shrews and rats suggest a pervasive low level of infections (Laakkonen, 1998; Chabé et al., 2010).

Pneumocystis organisms are transmitted via the air from infected individuals to new hosts (Hughes, 1982), including between individuals within hospitals (de Boer et al., 2011), but also possibly via the transplacental route (Ceré et al., 1997; Sanchez et al., 2007; Montes-Cano et al., 2009). The current hypothesis is that infections occur over short distance among infected and susceptible individuals (Chabé et al., 2011). The transfer of parasites from animals to humans is no longer considered as a valid hypothesis based on the strict host species specificity (Chabé et al., 2011). Consistently, no convincing evidence of an environmental source of *Pneumocystis* has been found so far, which strongly suggests that mammals constitute the only reservoir of these fungi. Furthermore, the erosion of metabolic capabilities evidenced by the genome sequencing studies suggests that these organisms are unable to live outside their hosts (see below, losses of metabolic machinery section). Finally, they apparently complete their whole cell cycle within host's lungs since sexuality occurs therein. Healthy infected hosts colonized by the organism are believed to contribute greatly to the transmission and circulation process (Chabé et al., 2004; Peterson and Cushion, 2005; Le Gal et al., 2012; Alanio and Bretagne, 2017).

Host specificity and biotrophy

The strict host species specificity of the *Pneumocystis* species means that the fungal cells can only infect or survive in the host in which they were isolated in the first place. This view is mainly supported by the systematic failure of cross-infection experiments involving severely combined immuno-deficient animals and nude rats (Aliouat et al., 1993; Furuta et al., 1993; Gigliotti et al., 1993; Aliouat et al., 1994; Atzori et al., 1999; Durand-Joly et al., 2002). The selective activation of trophic cells by their host seems to trigger the formation of cytoplasmic projections by *Pneumocystis* cells, the filopodia (Aliouat-Denis et al., 2008). Accordingly, *P. carinii*, the species naturally infecting rats, is unable to form filopodia and infect when inoculated in mice, whereas *P. murina*, the natural parasite of mice, produce filopodia and high parasite loads under the same conditions (Aliouat-Denis et al., 2008). The function of the filopodia remains elusive but these structures display ultrastructural differences that are species specific, and that might account for some aspects of the host specificity.

Another aspect of this host specificity is that *Pneumocystis* species are most probably obligate biotrophs (Cushion et al., 2007; Cushion and Stringer, 2010; Hauser, 2014; Ma et al., 2016a). The way fungal parasites scavenge nutrients from their host is an active research field and three modes are broadly recognized: (i) biotrophy, where the parasite acquires nutrients from a living cell, (ii) necrotrophy, where host cells are killed to release nutrients, and (iii) saprotrophy, where the organism feeds on dead or decaying organic material. Biotrophs do little damage to host cells and lack virulence factors (van der Does and Rep, 2007). *Pneumocystis* perfectly fits to the biotrophy definition because they cause no apparent cell death and lack any experimentally verified fungal virulence factors such as glyoxylate cycle, secondary metabolism, and secreted effectors (Cushion et al., 2007; Cissé et al., 2012; Cissé et al., 2014; Ma et al.,

2016a). This implies that they rely entirely on their host for their survival and thus have evolved close relationships that rendered them host species specific.

The physiological characteristics of the hosts are key determinants of parasite adaptation (Poulin et al., 2006). For example, micromammals are small bodied with short lifespans, high reproduction rates, and high population densities, whereas these distinctive features are reversed in large mammals such as humans. The co-evolution theory predicts that parasitic species infecting micromammals exhibit a weaker host specificity compared those adapted to long-lived hosts with more stable population densities (Poulin et al., 2006). This prediction has been validated in fish parasites, among which strong host specificity is favored in stable resources found in hosts with a large body size (Sasal et al., 1999; Desdevises et al., 2002). As far as *Pneumocystis* is concerned, humans are infected by only one species whereas rats can be co-infected by two (Cushion et al., 1993 and 2004; Icenhour et al., 2006a; Golab, 2009). The number of *Pneumocystis* species able to infect rodents might even be more important, as shown by the recent discovery of multiple lineages shared among species and genera of the Southeast Asian murid species (Latinne et al., 2017). These findings might indicate a relaxation of the strict host specificity in small mammals harboring *Pneumocystis*, although additional supporting data are needed to fully challenge the concept of widespread strict host specificity.

Co-phylogenetic studies of *Pneumocystis* species and their hosts suggest that the host specificity evolved as a continuous trait resulting from a long-lasting co-evolution (Demanche et al. 2001; Guillot et al., 2001; Hugot et al., 2003). Strict host specificity is rare in animal pathogens but widespread in plant fungal pathogens (Parker and Gilbert, 2004; Restrepo et al., 2014). In the latter, the ecological adaptation often results in a pronounced specialization to particular hosts (Clay and Kover, 1996). In these systems, host specificity acts as a reproductive

isolating mechanism because it favors higher rate of mating between individuals on the same host and reduced gene flow among populations from different hosts (Giraud, 2006). A rapid divergence of the virulence factors, the pathogen "effector repertoire", is often associated with the emergence of host specificity (Schulze-Lefert and Panstruga, 2011). The hypothesis of the latter authors states that changes in pathogen host range is driven by variation in the pathogen effector repertoire. This description fits the lineage specific expansion of the MSG superfamily in *Pneumocystis* species (Ma et al., 2016a; Ma et al., 2016b; Schmid-Siegert et al., 2017), which suggests that these proteins might account for some aspects of the host specificity.

GENOME ORGANIZATION

Genomic data acquisition

The quest for genome sequence data began with the successful cloning of *P. carinii* genomic fragments (Tanabe et al., 1988). Pulse field gradient gel electrophoreses have been then instrumental for karyotypic characterization of *Pneumocystis* genomes and evidenced 12 to 20 chromosomes according to the species totaling ca. 8 Mb (Hong et al., 1990; Yoganathan et al., 1989; Stringer and Cushion, 1998). Differences in karyotype profiles determined that the species infecting humans and rats are genetically distinct (Stringer et al., 1993). Significant genome size variations among species have been reported, *e.g.* that of ferret *Pneumocystis* would be ca. 1.7 times bigger than that of *P. carinii* (Stringer and Cushion, 1998). A draft of *P. carinii* genome covering ca. 70% of genome was generated in 2006 using a clone-based Sanger sequencing approach from infected laboratory rats (Slaven et al., 2006). In 2012, the first draft of *P. jirovecii* genome was obtained from a single bronchoalveolar lavage of a patient with pneumonia (Cissé et al., 2012). This assembly encompasses 358 contigs capturing 90 to 95% of the genome, but the repetitive subtelomeric and centromeric regions could not be resolved. The centromeres have not been discovered yet in *Pneumocystis*, whereas the subtelomeric regions were resolved using Sanger sequencing of cosmids (Keely et al. 2005), and more recently assembled using sequencing generating long reads (Ma et al., 2016a; Ma et al., 2016b; Schmid-Siegert et al., 2017) (see below, chromosomal ends section). Chromosomal level assemblies of *P. jirovecii*, *P. carinii*, and *P. murina* were recently published revealing genome sizes ranging from 7.4 to 8.3 Mb (Table 1; Ma et al., 2016a). It became evident that the genomes of *Pneumocystis* species had undergone an important reduction relative to *S. pombe* (7.5 to 8.3 Mb versus 12.5 Mb).

Nuclear Genome content

The analysis of the *Pneumocystis* genome assemblies validated the presence of single copy ribosomal DNA reported previously for *P. jirovecii* (Giuntoli et al. 1994; Stringer, 1996; Tang et al., 1998; Nahimana et al., 2000a). This is similar to *Taphrina deformans* (Cissé et al., 2013), but contrasts with most fungi which harbors commonly tens or hundreds of copies of the locus. Figure 3 shows the genome compositions of *Pneumocystis* species compared to related fungi. These data highlight the contraction of the protein coding regions as compared to free-living yeasts, which reflect massive gene losses. Figure 3 also evidences the expansions of the MSG superfamily, of introns, as well as of the cumulative length of the intergenic regions (IGR). We previously reported that IGR in *P. jirovecii* occupy a larger genome fraction as compared to free living yeasts *Saccharomyces cerevisiae* and *S. pombe* despite a significantly smaller genome (Cissé et al., 2014). This observation holds when we re-evaluate here IGRs in the newly published full-length genomes of *P. jirovecii*, *P. carinii,* and *P. murina* (Ma et al. 2016a). This strongly suggests that genome streamlining in *Pneumocystis* species is driven by gene deletions rather than reduction of IGRs. This observation seems counterintuitive because genome reduction is almost always associated to a reduction of introns and IGRs in parasites (Keeling and Slamovits, 2005). Alternatively, large IGRs might favor chromosomal re-arrangements by increasing the number of possible breakpoints, as hypothesized in fungal microsporidian parasites (Slamovits et al., 2004; Keeling and Slamovits, 2005). The other characteristics of these genomes are discussed in the following sections.

Chromosomal ends

Subtelomeres in microbial parasites are often enriched with multi-copy surface glycoprotein gene families (Deitsch et al. 2009). These genomic regions are prone to (i) gene silencing that can be used for mutually exclusive expression, (ii) enhanced mutagenesis, and (iii) ectopic recombinations facilitated by the formation of clusters of telomeres at the nuclear periphery (Barry et al., 2003). These regions correspond to an important proportion of the *Pneumocystis* genomes (ca. 5%), and harbor a superfamily including five to six families of highly polymorphic multi-copy proteins called major surface glycoproteins (MSG) that are believed to be crucial for the fungus' lifestyle (Ma et al., 2016a; Schmid-Siegert et al., 2017). These *msg* genes exist only in *Pneumocystis* species and all species of *Pneumocystis* have their own repertoire, which suggest they have been acquired in a common ancestor, although their origin is not known. The absence of homology of these MSGs outside *Pneumocystis* lineages might indicate a transfer from an unknown species or a gene co-option. *msg* families have been first described and studied in *P. carinii* (Kovacs et al., 1993; Sunkin et al., 1994; Sunkin et al., 1996; Keely et al., 2005; Keely and Stringer, 2009), and subsequently analyzed in *P. jirovecii* and *P. murina* (Haidaris et al., 1998; Kutty et al., 2008; Ma et al., 2016a; Schmid-Siegert et al., 2017). Important differences exist among *Pneumocystis* species in terms of *msg* gene copy numbers, 60 to 140 copies per cell, and protein divergence (Ma et al., 2016a).Moreover, one MSG family is present only in *P. jirovecii* (*msg*-IV or -B), whereas another one is present only in *P. carinii* and *P. murina* (MSR family, *i.e.* MSG-related). MSGs are believed to be involved in antigenic variation (Stringer, 2007). MSGs would also mask glucans at the asci surface from the immune recognition (Kutty et al., 2016). The antigenic diversity seems to be created via intra-family recombination of *msg* genes encoding different isoforms, creating mosaic genes, as well as through increased

mutagenesis (Kutty et al., 2008; Keely and Stringer, 2009; Schmid-Siegert et al., 2017). The expression of the most abundant MSG family (*msg*-I or –A1) that is present in all species is subject to mutually exclusive expression of a single isoform in each cell by using a single copy transcription promoter (the upstream conserved sequence, UCS) (Edman et al., 1996; Kutty et al., 2001; Sunkin et al., 1996; Wada et al., 1995). The UCS ends by the conserved recombination joint element (CRJE) which is also present at the beginning of each *msg*-I gene and may serve as recombination breakpoint (Stringer, 2007). The CRJE would be larger in *P. wakefieldiae* (ca. 330 bps) than in *P. murina* (132), which in turn is larger than in *P. carinii* and *P. jirovecii* (28 and 33, respectively) (Keely et al., 2007). On the other hand, at least in *P. jirovecii,* members of the other five families possess each their own promoter (Schmid-Siegert et al., 2017), but their expression patterns remain to be characterized. Recently, one family has been shown in *P. murina* to be expressed only in ascospores within asci and young trophic forms (Bishop et al., 2018).

Introns

Introns are extremely abundant in *Pneumocystis* genes and are as many as several tens per gene with a mean of five, and more than 40% of genes are interrupted by at least four introns (Stringer and Cushion, 1998; Ma et al., 2016a). Their presence can be equally explained by massive gains in *Pneumocystis* most recent common ancestry, or retention of ancestral elements that would have been lost in some Taphrinomycotina lineages such as *Schizosaccharomyces*. The introns are short (average length of 48 nucleotides), have a strong adenine and thymine bias, and present typical donor, acceptor and branch site patterns (Slaven et al., 2006). *Pneumocystis* introns cannot be processed by *S. pombe* and *S. cerevisiae* spliceosomes because of the divergence in intron-exon boundaries and branching sites within the introns (Thomas et al., 1999). RNA-seq

data indicate that intron retention affects ca. 45% of all introns (Ma et al., 2016a). *Pneumocystis* species contain self-splicing group I introns that are absent in higher eukaryotes such as humans (Liu et al., 1994), which renders them a prime target for the development of new drugs. These latter introns catalyze their own excision from RNA transcripts, a reaction that is inhibited by the drug pentamidine that is used against *Pneumocystis* (Liu and Leibowitz, 1993).

Given the important genome reduction at the *Pneumocystis* genus level, the presence of a high intron density per gene suggests a selective constraint to conserve them. Intron loss is dominant in fungi (Stajich et al., 2007), and this tendency is even more pronounced in some parasites such as microsporidia (Keeling et al., 2010). The intron history is highly flexible within the Taphrinomycota, with the plant-associated *Neolecta* having a high intron density similar to *Pneumocystis* (Nguyen et al., 2017), and the intron-poor free-living yeast *S. pombe* (Wood et al., 2002). The non-sense-mediated mRNA decay machinery is conserved in *Pneumocystis* species (Ma et al., 2016a). Under neutral scenario (no advantage) and widespread intron retention, most of the introns would produce non-functional transcripts tagged for destruction. This would be an incredible waste of resources in absence of another function. The latter could consist in alternative splicing increasing transcript diversity and regulating gene transcription or mRNA stability. Consistently, the *P. carinii* inosine 5'-monophosphate dehydrogenase pre-mRNA is differentially spliced, which was suggested to reflect changes in environmental stresses (Ye et al., 2001). These considerations suggest that introns might be neutral elements involved in many cellular processes via a greater proteome diversity, possibly including acting as a favorable substrate to facilitate shifts in lifestyle (*i.e.* parasite transition from one host species to another, or from plant to animal).

Mitogenomes

The mitochondrial genomes of *P. carinii* (Sesterhenn et al., 2010; Ma et al., 2013), *P. jirovecii* (Cissé et al., 2012; Ma et al., 2013), and *P. murina* (Ma et al., 2013) have been sequenced. The mitogenome sizes range from 24 to 35-kb with a substantial size variability among isolates in all species. *P. carinii* and *P. murina* mitogenomes end with single-stranded loop sequences that would allow forming linear concatemers and protecting the ends of the molecule. The presence of these repeats might account for the variable size of *P. carinii* mitogenomes. *P. jirovecii* mitochondrial genome is circular since it lacks inverted terminal repeat allowing circulation. The significance of circularity versus linearity is unknown. Related Taphrinomycota of the genera *Schizosaccharomyces, Taphrina,* and *Neolecta* have circular genomes (Bullerwell et al., 2003; Cissé et al., 2013; Tsai et al., 2014; Nguyen et al., 2017), which might indicate that the circular form is ancestral. Interestingly, *P. carinii* and *P. murina* mitogenomes are highly co-linear whereas *P. jirovecii* mitogenome presents some re-arrangements, similarly to the nuclear genomes (see below Chromosomal re-arrangement section). The gene content is highly conserved among the three *Pneumocystis* species, although there is a substantial nucleotide divergence among species (27 to 31%) (Ma et al., 2013). These mitogenomes encode ca. 17 genes commonly found in mitochondrial fungal genomes such as ATP synthases, cytochrome *c* oxidases, NADH dehydrogenases, and the full repertoire of at least 20 transfer RNAs. Reports investigating the dynamics of the mitochondrial genes during infection have revealed that mitogenomes would be very plastic in terms of copy number variations (Valero et al., 2016), and of genetic diversity including heteroplasmy (Alanio et al., 2016). The subsequent sections of this review focus on nuclear genomes.

EVOLUTION

Comparison of the gene families and pathways present in *Pneumocystis* genomes to those in selected fungi has revealed numerous losses / contractions and relatively few expansions (Table 2). The hypothetical evolutionary history of *Pneumocystis* species derived from these observations is represented in Figure 4 and discussed in the following sections.

Losses in the metabolic and cellular machineries

Massive gene losses suggest that *Pneumocystis* species are auxotroph for essential nutrients, which might explain the recurrent failures of *in vitro* culturing attempts. The lost pathways include basic components of metabolic machinery such as the synthesis of amino acids or carbohydrates (Table 2). The loss of purines catabolism seems unique to *Pneumocystis* (Chitty and Fraser, 2017). *Pneumocystis* species are able to synthesize fecosterol and episterol but lack enzymes to convert them into ergosterol. Consequently, their membranes contain cholesterol instead of ergosterol, which probably explains their resilience to azole treatment. *Pneumocystis* organisms are also able to synthetize a unique class of sterols, the "pneumocysterols" (Kaneshiro et al. 1994; Kaneshiro et al. 1999; Florin-Christensen et al. 1994; Giner et al., 2002). It interesting to note the early steps of the sterol biosynthetic pathway leading to the formation of pneumocysterol and episterol are conserved in *Pneumocystis* species, and only the final steps toward ergosterol/cholesterol production are missing (Joffrion et al., 2010). This is exemplified by the fact that key enzymes for the formation of ergosterol (*i.e. erg3*, *erg4* and *erg5*) are not identifiable within the genomes. Analysis of the sterol biosynthesis machinery suggest that these species may be able to synthetize ergosterol/cholesterol precursors such as zymosterol, fecosterol

and episterol. Thus, the sterol pathway may have been re-routed and branch to form unique sterols, the pneumocystisterols.

Overall, these observations are consistent with the idea that losses of metabolic genes correlate with an increased dependency of the parasite on its host. Therefore, nutrients need to be scavenged from the host, which often mechanistically involves large batteries expanded transporters (*e.g.* as observed in microsporidia [Cuomo et al., 2012]). This is not the case in *Pneumocystis* since transporters families are also greatly reduced (Cissé et al., 2014; Ma et al., 2016a). For instance, the amino acid permeases and transporters that can respectively carry amino acids and oligopeptides are greatly reduced relatively to other Taphrinomycota (one copy of general amino acid permease versus 21 copies in *S. pombe*). Transmembrane proteins such as those of the major facilitator superfamily, sugar transporters, or more specific transporters (*e.g.* efflux pumps) are significantly reduced in *Pneumocystis*. The reduction of the transporters battery might be compensated by the use of highly selective transporters for critical compounds. The recent discovery of the import of *myo*-inositol in *Pneumocystis* cells via a low affinity but highly selective system supports this idea (Cushion et al., 2016). Unfortunately, high affinity transporters cannot be identified solely by computational means. Alternatively, simple diffusion across the membrane may occur, as evidenced in *P. carinii* for amino acids uptake using *in vitro* experiments (Basselin et al, 2001a; Basselin et al, 2001b). Basic cellular machinery is also affected by the loss of several fungal specific transcription factor families and the RNA interference machinery (Table 2).

Evolutionary basis of gene loss

Gene loss is a common trend in parasitic and symbiotic species, which often harbor a small sized genome (Keeling and Slamovits, 2005; Wolf and Koonin, 2013). The driving factors are often unknown or specific to the lifestyle of the species under study. A central question in evolutionary biology is whether gene loss is neutral or adaptive. In *Pneumocystis* species, there are footprints of both processes and we discuss here a few examples.

The neutral theory is usually sufficient to explain gene loss in parasites (O'Malley et al., 2016). Organisms with narrow host niche such as *Pneumocystis* are predicted to have small sized populations with increased **genetic drift** (bold: see glossary) (Papkou et al., 2016). The main mechanisms for gene loss are pseudogenization and sudden DNA deletions. Pseudogenization consists in the accumulation of deleterious mutations in non-essential genes ultimately leading to the loss (Kuo and Ochman, 2009; Wernegreen, 2015). The proportion of pseudogenes in *P. jirovecii* is low and equivalent to that present in free-living yeasts (0.02 pseudogene per protein-coding gene [Cissé et al., 2014]). This observation might indicate that pseudogenization is not the main driver of gene loss in this species. The following considerations do not undermine this observation but suggest that caution must be exercised: (i) this rate of pseudogenization is valid only for *P. jirovecii* and for the single isolate which genome was sequenced (Cissé et al., 2012), and (ii) only genes including stop codons were considered, that is, other types of gene inactivation were not considered (*e.g.* untranslated RNA genes or unfixed mutations). Gene loss can also be result of deletions independent of selection such as the movement of transposable or integrated viruses (reviewed by Albalat and Canestro, 2016).

The adaptive theory of gene loss implies a selective advantage and has been demonstrated to have occurred in many pathogenic lineages, for example for the Allergen 1 in

Cryptococcus neoformans (Jain et al., 2009), and for the *de novo* biosynthesis of nicotinic acid genes in *Candida glabrata* (Domergue et al., 2005). In *Pneumocystis*, the loss of chitin might have been lost to allow avoiding recognition from the host immune system (Ma et al., 2016a). The gene families and pathways cited in Table 2 are missing in the three *Pneumocystis* genomes available (Ma et al 2016a), which suggests that these losses occurred before the radiation of the genus. An unexpected consequence is that the observed gene losses might not reflect the current selective forces, and therefore might not be relevant for the host specificity.

Chromosomal re-arrangement

The chromosome level assemblies revealed that an important chromosomal re-arrangement occurred among *Pneumocystis* species (Ma et al., 2016a). The re-arrangement, however, followed the species tree, that is, the macrosynteny is broken between rodents infecting *Pneumocystis* (*P. carinii* and *P. murina*) and the humans infecting species (*P. jirovecii*), whereas *P. carinii* and *P. murina* genomes are highly collinear. Nevertheless, the gene order is conserved in syntenic regions among the three species (>92% of the genes), and ca. 83% of gene families are orthologous, with 4 to 30% of divergence at the nucleotide sequence level. The high gene conservation among the three species suggest that re-arrangements occurred mostly in the intergenic regions (IGR). In fungi, IGRs are often enriched in regulatory functions such as signal transduction or binding sites of transcription factors (Noble and Andrianopoulos, 2013). Chromosomal translocations impact gene expression as well as long-distance gene-to-gene contact via chromatin interactions, and thus might be involved in speciation (Rieseberg, 2001; Bakloushinskaya, 2016). Protein evolution is also faster in re-arranged chromosomes than collinear chromosomes because re-arrangements reduce homologous recombination and

facilitate positive selection (Rieseberg, 2001). A key question here is whether chromosomal re-arrangements are involved in the adaptation of each *Pneumocystis* species to its host. Future studies are required to probe an eventual role of these re-arrangements in *Pneumocystis* evolution.

Loss of complex multicellularity

The recent sequencing of the *Neolecta irregularis* genome revealed that the Taphrinomycotina last common ancestor was probably multicellular (Nguyen et al., 2017). These findings suggest that *Pneumocystis* organisms evolved from a plant-associated or soil-adapted multicellular organism. The shift in cell morphology to single celled organisms is associated with the deletion of an ancestral morphogenic kit that included many cell differentiation and cell-to-cell signaling genes. These losses are not specific to *Pneumocystis* and were observed in a wide range of unrelated yeasts (Nguyen et al., 2017; Nagy et al., 2014; Nagy, 2017), which suggests a convergent evolution. The transition from a hyphal to yeast form takes place in many fungal lineages and is often triggered by a thermal stimulus (Köhler et al., 2017), CO_2 levels (Hall et al., 2010), or pH (Davis, 2009), and is directly linked to the ability to invade hosts. Notable examples include the dimorphic human pathogenic fungi *Histoplasma*, *Blastomyces*, *Coccidioides*, and *Paracoccidioides* (Beaman et al., 1981; Medoff et al., 1987; Inglis et al., 2013).

The ancestral morphogenic kit for complex multicellularity (fruiting bodies) is lost in *Pneumocystis*. However, *Pneumocystis* species are able to produce biofilms (Cushion et al., 2009), which is an undifferentiated form of aggregative multicellularity often seen in bacteria (Claessen et al., 2014). Inversely, the yeast *Saitoella complicata* grows primarily by budding

(Goto et al., 1987), despite having the cellular machinery for the production of fruiting bodies (Nguyen et al., 2017). Comparative genomics and epigenomics would be extremely valuable to explore the molecular process underlying the loss of the multicellular phenotype. These considerations highlight the fact that phenotypes cannot be explained solely by gene loss and gain balance, and that other subtle mechanisms need to be considered.

POPULATION GENETICS

Strain typing

Given the high homogeneity of genomic sequences at the nucleotide sequence level among *P. carinii* isolates, strain typing for this species relied on chromosomes' size analyses which allowed identifying numerous different karyotypic forms (Lundgren et al., 1990; Cushion, 1998; Wakefield, 1998a; Nahimana et al., 2001). On the other hand, the low but significant heterogeneity in many genomic loci among *P. jirovecii* isolates allowed using multilocus sequence typing (Wakefield, 1998b). The latter method represents nowadays the most used technique for *P. jirovecii* strains identification. The discrimination power of eight distinct loci has been validated and extensively used for epidemiological studies of *P. jirovecii* pneumonia (Maitte et al., 2013). Genotypes identification is performed by PCR of multiple loci followed by direct DNA sequencing (Sanger), restriction fragment length polymorphism, single-strand conformation polymorphism, type-specific oligonucleotide hybridization, tandem repeats number analysis, or high-throughput amplicon sequencing (Hauser et al., 1997; Hauser et al., 1998; Lee et al, 1993; Lu et al., 1995; Ma et al., 2002; Alanio et al., 2016; Esteves et al., 2016).

Genetic diversity

The conclusions drawn from the studies concerning *Pneumocystis* genetic diversity were often contradictory. Low levels of genetic diversity as defined by Shannon diversity and Simpson indexes (Shannon, 1948; Simpson 1949) have been reported at the *P. jirovecii* and *P. carinii* internal transcribed spacers of the nuclear rDNA operon using PCR-based Sanger sequencing (Palmer et al., 2000; Beser et al., 2011). On the other hand, moderate to important levels of diversity measured in term of DNA polymorphisms in *P. jirovecii* using multilocus sequence

typing have been reported (Matos and Esteves, 2010; Jarboui et al., 2013; Sun et al., 2015; Alanio et al., 2017). The lack of whole genome sequence data, differences in sampling strategies, differences in interpretation, as well as the likely frequent *in vitro* formation of PCR chimeras (Beser et al., 2007), make difficult the reconciliation of these conclusions.

Moreover, sexual recombination could explain partly these conflicting conclusions. Indeed, sexual reproduction is one of the main mechanisms to generate genetic diversity in fungi. It is believed to favor adaptation in fluctuating conditions while purging deleterious alleles (Heitman, 2010). *Pneumocystis* are probably homothallic species (Almeida et al., 2015; see life cycle section), and self-fertilization favors mating by avoiding the search of a compatible partner, a strategy thought to be favorable to and adopted by several human pathogens such as *Cryptococcus* and *Candida* species (Heitman, 2010). Sexual reproduction is based on classical Mendelian segregation, which supports both cross- and self-fertilization (Buscaglia et al., 2015). *Pneumocystis* would be able to perform both clonal and sexual propagation with various degrees of inbreeding or outcrossing. These variations in the multiplication process could explain the conflicting patterns of genetic diversity reported.

Polymorphism rates change substantially across loci and chromosomes in various species, including fungi, plants, and animals (Ellegren and Galtier, 2016). Genetic diversity is influenced by three main forces: mutation, demography (migration and **bottlenecks**), and selection (**selective sweeps** or **clonal interference**). Demography and selection create differences in the effective population size, whilst variations in mutation rate may create differences in the level of genetic diversity according to the geographical location. Neutral mutation rates in eurotiomycetes are typically between 1×10^{-8} and 1×10^{-9} substitutions per site per year (Kasuga et al., 2002), and a rate of 1.2×10^{-10} for the 18S rDNA has been used to

estimate *Pneumocystis* species divergence (Keely et al., 2003a). However, the genome-wide mutation rates for these species are unknown and expected to fluctuate greatly among genomic regions. For example, subtelomeric regions harboring MSGs have high substitution rates (Keely and Stringer, 2009; Schmid-Siegert et al., 2017), whereas ribosomal regions display a normal rate (Fischer et al., 2006). Moreover, given their likely variations according to the host, the mutation rates for each species must be determined independently. Care must be taken inferring these rates because recombination can be mutagenic and its impact as well as other confounder effects need to be addressed.

The size of the populations of *Pneumocystis* species are not known, but they are expected to be small because of their narrow host ranges. *P. jirovecii* would have a small population size relative to the species infecting micro-mammals, thus reflecting the small size of human populations relative to those of rodents. Variations in population size over time affect the genetic diversity, *e.g.* a strong population **bottleneck** creates a loss of allele diversity due to increased **genetic drift**. Using non-recombining neutral loci, realistic mutation rates, and appropriate molecular clock models, past population history can be traced back using coalescent theory based applications such as skyline plot methods (Drummond et al., 2005; Heled and Drummond, 2008). These demographic reconstructions would provide key metrics such as ancestral population sizes and evolutionary rates.

Interestingly, the strongest prediction of genetic diversity in many species is the life history, not the population history (Ellegren and Galtier, 2016). This means that there is a strong correlation between phenotypic traits (*e.g.* mating system, generation times) and the genetic diversity. For example, homothallism is expected to have long term evolutionary cost fitness because selfing populations experience reduced recombination rates and size, which ultimately

reduce the strength of purifying selection and increase genetic drift (Charlesworth and Wright, 2001; Hill and Robertson, 1966; Otto and Lenormand, 2002; Pollak, 1987). The homothallism used by *Pneumocystis* species is also often associated to higher probability to experience population **bottlenecks** via founder effects and linked selection (Jarne, 1995; Charlesworth and Wright, 2001). There is a complex interplay between demographic, selective factors, and genetic diversity. Alternative scenarios, such as purifying selection purging deleterious alleles, which is known as "background selection" (Charlesworth, 1994), need also to be considered. In conclusion, many factors may have influenced genetic diversity of *Pneumocystis* species, which remains unclear.

Population structure

The population structure of *Pneumocystis* species is also controversial. Indeed, data support an absence of strong subdivision in *P. jirovecii* (Parobek et al., 2014) and *P. carinii* (Palmer et al., 2000), whereas other data support possible geographical clusters in *P. jirovecii* (Esteves et al., 2016; Alanio et al., 2017). Importantly, Matos and Esteves (2010) noted that the infections are not necessarily clonal and recombination between multi-locus genotypes is possible. All these inferences are based on a relatively small number of markers (*e.g.* ITS, mitochondrial large subunit rDNA), and need to be validated at the genome scale using appropriate Bayesian methods based on unlinked multi-allelic genotypes, such as STRUCTURE (Pritchard et al., 2000). In the meantime, interesting clues can be extracted from the biological cycle. The question is whether the fluctuation of the population structure is caused by variations in spore dispersal or in sexual recombination. The asci are 4-6 μm in size, which is small enough to be airborne dispersed efficiently over long distances. The asci cell wall is enriched with

glycoproteins, melanin, β-glucans, and mannans without outer chain (Kottom and Limper, 2000; Icenhour et al., 2003; Icenhour et al., 2006b; Ma et al., 2016a), which might allow them to resist desiccation and UV irradiation usually fatal to many fungal spores (Golan and Pringle, 2017; Latgé, 2007). The viability of spores for extended periods of time is supported by the detection of *P. jirovecii* mRNA in hospital air samples (Latouche et al., 2001; Maher et al., 2001). Their resistance to physical assaults is suggested by their detection in air spora trapped in rural locations (Wakefield, 1996).

Dispersal of fungi can occur in two modes: (i) multiple sequential short-distance dispersal, and (ii) a single successful long-distance move of spores ultimately coinciding with optimal conditions for the growth of the fungus (Golan and Pringle, 2017). The former option produces a strong population subdivision, while the latter ends up with no or weak population structure because the same genotype(s) will be spread over large geographical distance. If long distance dispersal occurs on a global scale, it will result in a global population structure (Pringle et al., 2005). Rare long-distance dispersal would involve stochastic founding events, which can be revealed by population structures with an excess of rare alleles. Future studies combining genetic and geography are needed to fully access the population structures of *Pneumocystis* species.

Clonal evolution or predominant sex/recombination?

P. jirovecii infections are most often caused by multiple populations co-infecting the same individual (Hauser et al., 1997; Nahimana et al., 2000b; Palmer et al., 2000; Ma et al., 2002; Alanio et al., 2016). Multilocus genotypes (MLGs), which refer to a unique combination of alleles, can persist over long periods of time (4 to 9 years), and be observed across different

countries (Wakefield et al., 1994; Esteves et al., 2010). Recombination was also detected among MLGs (Esteves et al., 2010), which would explain the reported lack of strong population subdivision, at least in *P. jirovecii* (Parobek et al., 2014). Under **panmictic** population conditions, MLGs should not persist in the population because they will be disturbed consistently by recombination.

Buscaglia et al. (2015) proposed that "a highly structured (*i.e.* clonal) population indicates that the main mode of reproduction for such a species lacks genetic exchange (*i.e.* is primarily asexual) or sex occurs only rarely". In *Pneumocystis*, MLGs do recombine which would indicate a limited global population structure (Esteves et al., 2010; Parobek et al., 2014). Thus, the definition proposed by Buscaglia et al. would suggest a widespread sexual reproduction in *Pneumocystis* species. However, some MLGs persist over time, which suggests that these species might be mostly clonal and only rarely engage to sexual events. This latter scenario would be consistent with the theory of predominant clonal evolution (Tibayrenc and Ayala, 2012; Tibayrenc and Ayala, 2014), which proposes that restrained recombination is not strong enough to disturb the pattern of clonal structure. The frequency of recombination events at the genome level is unknown in *Pneumocystis*, which currently prevents reaching definitive conclusions.

Intra individual short-term evolution

Infections are usually caused by multiple *P. jirovecii* strains acquired from different origins (infections *de novo* but also possibly re-activation of organisms). The balance between different strains will likely change over the course of the disease because of either drug treatment, pressures from the host immune system, and/or varying metabolism and fitness among the strains

present. Other pathogens such as *Candida* and *Cryptococcus* species evolve rapidly within their hosts by acquiring new mutations or changes in genomic heterozygosity associated with drug resistance (Ford et al., 2015; Chen et al., 2017). It is unclear if theses mutations result from positive selection or DNA repair errors in *Cryptococcus* (Rhodes et al., 2017), although it is also possible that DNA repair errors are selected by positive selection. Competition among multiclonal parasite populations within the same host can, in theory, promotes parasite diversity (Bashey, 2015). The full extent of *Pneumocystis* short-term evolution within their host is unknown. Interestingly, Alanio and colleagues used a set of markers to evidence changes in population composition during *P. jirovecii* infections (Alanio et al. 2016). Multiple strains infections are frequently found in pathogens and may have clinically relevant consequences (Balmer and Tanner, 2011). Different strains might have different susceptibility to treatment or evolve differently so that they may escape detection by the immune system or diagnostics tools. We anticipate that the characterization of multiclonal infections will have serious implications for the treatment and the management of *P. jirovecii* pneumonia. Experimental setups will become realistic when long-term *in vitro* culture method will become widely reproducible.

PERSPECTIVES AND CONCLUSIONS

The research on *Pneumocystis* is still in its infancy mainly because of the lack of culture *in vitro*, but the availability of genomic data will help exploring the mysteries of their evolution. The next cornerstone will be the establishment of a long-term culture system and genetic manipulation. The upcoming expectation goes far beyond the *Pneumocystis* research community and will allow exploring key questions in evolutionary cell biology such as the evolution of parasitism and multicellularity. The study of *Pneumocystis* organisms has the unique interest that they are the only strictly mammalian-adapted fungal pathogens. Thus, determining the molecular basis of their adaptation and speciation are of uttermost importance. The key questions are: what are the determinants of the genome reduction? What are the molecular determinants of the host specificity and speciation? Why introns are so abundant and what are their function(s)? What are the impact of multiclonal infections and short-term evolution within host in the context of drug resistance and development of vaccines? How do natural populations of *Pneumocystis* evolve in different hosts?

ACKNOWLEDGMENTS

OHC is currently supported by an US National Institutes of Health Fogarty fellowship. PHM's research is presently funded by the Swiss National Science Foundation (grant 310030_165825). This research was supported by the Intramural Research Program of the NIH, Clinical Center. We thank Joe Kovacs and Liang Ma for critical reading of the manuscript and fruitful comments.

GLOSSARY

Background selection

Reduction of genetic diversity at linked loci owing to selection against deleterious mutations.

Bottleneck

A sharp and rapid reduction in the size of a population.

Clonal interference

Phenomenon in population genetics of organisms with significant linkage disequilibrium (*i.e.* absence of recombination), especially in asexual organisms. It occurs when two (or more) different beneficial mutations arise independently in different individuals.

Effective population size

The size that a theoretical population evolving under a Wright-Fisher model would need to be in order to match aspects of the observed genetic data.

Genetic drift

Fluctuation of allele frequency among generations in a population owing to the randomness of survival and reproduction of individuals, irrespective of selective pressures.

Haploid selfing

Refers to true homothallic species. A species able to accomplish their entire sexual reproduction without the need of a partner.

Heterozygosity

Measure of the genetic diversity, which represents the presence of different alleles at one or more loci on homologous chromosomes. Often presented as a probability that two randomly sampled gene copies in a population carry distinct alleles.

Linkage disequilibrium

Nonrandom association of alleles at two loci often but not always due physical linkage. Such association is broken over time by recombination.

Panmictic population

Random mating among individuals in an idealized population.

Selective sweep

Elimination or reduction of genetic diversity in the neighborhood of a beneficial allele that increases in frequency in the population, typically after an environmental change.

Selective sweeps

Elimination or reduction of genetic diversity in the neighborhood of a beneficial allele that increases in frequency in the population, typically after an environmental change.

REFERENCES

Alanio, A., Gits-Muselli, M., Mercier-Delarue, S., Dromer, F., Bretagne, S., 2016. Diversity of *Pneumocystis jirovecii* during infection revealed by ultra-deep pyrosequencing. Front. Microbiol. 7, 733.

Alanio, A., Bretagne, S., 2017. *Pneumocystis jirovecii* detection in asymptomatic patients: what does its natural history tell us? F1000 Faculty Rev. 6, 739.

Alanio, A., Gits-Muselli, M., Guigue, N., Desnos-Ollivier, M., Calderon, E. J., Di Cave, D., Dupont, D., Hamprecht, A., Hauser, P.M., Helweg-Larsen, J., Kicia, M., Lagrou, K., Lengerova, M., Matos, O., Melchers, W.J.G., Morio, F., Nevez, G., Totet, A., White, L. P., Bretagne, S., 2017. Diversity of *Pneumocystis jirovecii* across Europe: a multicentre observational study. EBioMedicine 22, 155-163.

Albalat, R., Canestro, C., 2016. Evolution by gene loss. Nat. Rev. Genet. 17, 379-391.

Aliouat, E.M., Mazars, E., Dei-Ca, E., Cesbron, J., Camus, D., 1993. Intranasal inoculation of mouse, rat or rabbit-derived *Pneumocystis* in SCID mice. J. Protozool. Res. 3, 94-98.

Aliouat, E.M., Mazars, E., Dei-Cas, E., Delcourt, P., Billaut, P., Camus, D., 1994. *Pneumocystis* cross infection experiments using SCIDmice and Nude rats as recipient host, showed strong hostspecies specificity. J. Eukaryot. Microbiol. 41, S71.

Aliouat, E.M., Dujardin, L., Martinez, A., Duriez, T., Ricard, I., Dei-Cas, E., 1999. *Pneumocystis carinii* growth kinetics in culture systems and in hosts: involvement of each life cycle parasite stage. J. Eukaryot. Microbiol. 46, S116-S117.

Aliouat-Denis, C.M., Chabé, M., Demanche, C., Aliouat, E.M., Viscogliosi, E., Guillot, J., Delhaes, L., Dei-Cas, E., 2008. *Pneumocystis* species, co-evolution and pathogenic power. Infect. Genet. Evol. 8, 708-726.

Aliouat-Denis, C.M., Martinez, A., E. Aliouat, M. Pottier, N. Gantois, Dei-Cas, E., 2009. The *Pneumocystis* life cycle. Memor. Instit. Oswaldo Cruz 104, 419-426.

Almeida, J.M.G.C.F., Cissé, O.H., Fonseca, A., Pagni, M., Hauser, P.M., 2015. Comparative genomics suggests primary homothallism of *Pneumocystis* species. mbio 6, e02250-14.

Atzori C, A.F., Angeli, E., Mainini, A., Micheli, V., Cargnel, A., 1999. *P. carinii* host specificity: attempt of cross infections with human derived strains in rats. J. Eukaryot. Microbiol. 46, S112.

Bakloushinskaya, I.Y., 2016. Chromosomal rearrangements, genome reorganization, and speciation. Biol. Bull. 43, 759-775.

Balmer, O., Tanner, M., 2011. Prevalence and implications of multiple-strain infections. Lanc. Infect. Dis. 11, 868-878.

Banerji, S., Lugli, E.B., Wakefield, A.E., 1994. Identification of two genetically distinct strains of *Pneumocystis carinii* in infected ferret lungs. J. Eukaryot. Microbiol. 41, S73.

Bao, W., Kojima, K.K., Kohany, O., 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA 6, 11.

Barry, J.D., Ginger, M.L., Burton, P., McCulloch, R., 2003. Why are parasite contingency genes often associated with telomeres? Int. J. Parasitol. 33, 29-45.

Bashey, F., 2015. Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. Phil. Trans. R. Soc. B 370, 20140301.

Basselin, M., Lipscomb, K.J., Qiu, Y.H., Kaneshiro, E.S., 2001a. Transport of aspartic acid, arginine, and tyrosine by the opportunistic protist *Pneumocystis carinii*. Biochim. Biophys. Acta 1511, 349-359.

Basselin, M., Qiu, Y.H., Lipscomb, K.J., Kaneshiro, E.S., 2001b. Uptake of the neutral amino acids glutamine, leucine, and serine by *Pneumocystis carinii*. Arch. Biochem. Biophys. 391, 90-98.

Beaman, L., Benjamini, E., Pappagianis, D., 1981. Role of lymphocytes in macrophage-induced killing of *Coccidioides immitis* in vitro. *Infect. Immun.* 34, 347-353.

Beimforde, C., Feldberg, K., Nylinder, S., Rikkinen, J., Tuovila, H., Dorfelt, H., Gube, M., Jackson, D.J., Reitner, J., Seyfullah, L.J., Schmidt, A.R., 2014. Estimating the Phanerozoic history of the Ascomycota lineages: combining fossil and molecular data. Mol. Phylogenet. Evol.78, 386-398.

Beser, J., Hagblom, P., Fernandez, V., 2007. Frequent in vitro recombination in internal transcribed spacers 1 and 2 during genotyping of *Pneumocystis jirovecii*. J. Clin. Microbiol. 45, 881-886.

Beser, J., Botero-Kleiven, S., Lebbad, M., Hagblom, P., Fernandez, V., 2011. A limited number of ITS haplotypes defines the diversity of *Pneumocystis jirovecii* strains in Sweden. Infect. Genet. Evol. 11, 948-954.

Bishop, L.R., Davis, A.S., Bradshaw, K., Gamez, M., Cissé, O.H., Wang, H., Ma, L., Kovacs, J.A., 2018. Characterization of p57, a stage-specific antigen of *Pneumocystis murina*. J. Infect. Dis. doi: 10.1093/infdis/jiy099. [Epub ahead of print].

Brown, G.D., Denning, D.W., Gow, N.A., Levitz, S.M., Netea, M.G., White, T.C., 2012. Hidden killers: human fungal infections. Sci. Transl. Med. 4, 165rv113.

Bullerwell, C.E., Leigh, J., Forget, L., Lang, B.F., 2003. A comparison of three fission yeast mitochondrial genomes. Nucl. Ac. Res., 31, 759-768.

Buscaglia, C.A., Kissinger, J.C., Aguero, F., 2015. Neglected tropical diseases in the post-genomic era. Trends Genet. 31, 539-555.

Ceré , N., Drouet-Viard, F., Dei-Cas, E., Chanteloup, N., Coudert, P., 1997. *In utero* transmission of *Pneumocystis carinii* sp. f. *oryctolagi*. Parasite 4, 325–330.

Chabé, M,. Dei-Cas, E., Creusy, C., Fleurisse, L., Respaldiza, N., Camus, D., Durand-Joly, I., 2004. Immunocompetent hosts as a reservoir of *Pneumocystis* organisms: histological and RT-PCR data demonstrate active replication. Eur. J. Clin.. Microbiol. 23, 89-97.

Chabé, M., Herbreteau, V., Hugot, J.P., Bouzard, N., Deruyter, L., Morand, S., Dei-Cas, E., 2010. *Pneumocystis carinii* and *Pneumocystis wakefieldiae* in wild *Rattus norvegicus* trapped in Thailand. J. Eukaryot. Microbiol. 57, 213-217.

Chabé, M., Aliouat-Denis, C. M., Delhaes, L., Aliouat, E.M., Viscogliosi, E., Dei-Cas, E., 2011. *Pneumocystis*: from a doubtful unique entity to a group of highly diversified fungal species. FEMS Yeast Res. 11, 2-17.

Chagas, C., 1909. Nova tripanozomiaze humana: estudos sobre a morfolojia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., ajente etiolojico de nova entidade morbida do homem. Mem. Inst. Oswaldo Cruz 1, 159-218.

Charlesworth, B., 1994. The effect of background selection against deleterious mutations on weakly selected linked variants. Genet. Res. 63, 213-227.

Charlesworth, D., Wright, S.I., 2001. Breeding systems and genome evolution. . Curr. Opin. Genet. Dev. 11, 685-690.

Chen Y, F. R., Giamberardino C, Sakthikumar S, Jones A, and e. a. Yang T (2017). Microevolution of serial clinical isolates of *Cryptococcus neoformans* var. *grubii* and *C. gattii*. MBio **8**: e00166–00117.

Cherry, J.M., Adler, C., Ball, C., Chervitz, S.A., Dwight, S.S., Hester, E.T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., Weng, S., Botstein, D., 1998. SGD: *Saccharomyces* Genome Database. *Nucl. Ac. Res.* 26, 73-79.

Chitty, J.L., Fraser, J.A., 2017. Purine acquisition and synthesis by human fungal pathogens. Microorg. 5, E33. doi: 10.3390/microorganisms5020033.

Christensen, C.B.V., Settnes, O.P., Bille-Hansen, V., Jorsal, S.E., Henriksen, S.A., Lundgren, B., 1996. *Pneumocystis carinii* from pigs and humans are antigenically distinct. J. Med. Vet. Mycol. 34, 431-433.

Churukian, C.J., Schenk, E.A., 1977. Rapid Grocott's methenamine silver nitrate method for fungi and *Pneumocystis carinii*. Am. J. Clin. Pathol. 68, 427-428

Cissé, O.H., Pagni, M., Hauser, P.M., 2012. *De novo* assembly of the *Pneumocystis jirovecii* genome from a single bronchoalveolar lavage fluid specimen from a patient. mBio 4, e00428-00412.

Cissé, O.H., Almeida, J.M., Fonseca, A., Kumar, A.A., Salojarvi, J., Overmyer, K., Hauser, P.M., Pagni, M., 2013. Genome sequencing of the plant pathogen *Taphrina deformans*, the causal agent of peach leaf curl. mBio 4, e00055-00013.

Cissé, O.H., Pagni, M., Hauser, P.M., 2014. Comparative genomics suggests that the human pathogenic fungus *Pneumocystis jirovecii* acquired obligate biotrophy through gene loss. Gen. Biol. Evol. 6, 1938-1948.

Claessen, D., Rozen, D.E., Kuipers, O.P., Sogaard-Andersen, L., van Wezel, G.P., 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies" Nat. Rev. Microbiol. 12, 115-124.

Clay, K., Kover, P. X., 1996. The Red Queen Hypothesis and plant-pathogen interactions. Ann. Rev. Phytopathol. 34, 29-50.

Cuomo, C.A., Desjardins, C.A., Bakowski, M.A., Goldberg, J., Ma, A.T., Becnel, J.J., Didier, E.S., Fan, L., Heiman, D.I., Levin, J.Z., Young, S., Zeng, Q., Troemel, E.R. 2012. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Gen. Res. 22, 2478-2488.

Cushion, M.T., Zhang, J., Kaselis, M., Giuntoli, D., Stringer, S.L., Stringer, J.R., 1993. Evidence for two genetic variants of *Pneumocystis carinii* coinfecting laboratory rats. J. Mycol. Med. 8, 1-6.

Cushion, M.T., 1998. Genetic heterogeneity of rat-derived *Pneumocystis*. FEMS Immunol. Med. Microbiol. 22, 51-58.

Cushion, M.T., Keely, S.P., Stringer, J.R., 2004. Molecular and phenotypic description of *Pneumocystis wakefieldiae* sp. nov., a new species in rats. Mycologia 96, 429-438.

Cushion, M.T., Smulian, A.G., Slaven, B.E., Sesterhenn, T., Arnold, J., Staben, C., Porollo, A., Adamczak, R., Meller, J., 2007. Transcriptome of *Pneumocystis carinii* during fulminate infection: carbohydrate metabolism and the concept of a compatible parasite. Plos One 2, e423.

Cushion, M.T., Collins, M.S., Linke, M.J., 2009. Biofilm formation by *Pneumocystis* spp. Eukaryot. Cell 8, 197-206.

Cushion, M.T., Linke, M.J., Ashbaugh, A., Sesterhenn, T., Collins, M.S., Lynch, K., Brubaker, R., Walzer, P D., 2010. Echinocandin treatment of *Pneumocystis* pneumonia in rodent models depletes cysts leaving trophic burdens that cannot transmit the infection. PLoS One 5, e8524.

Cushion, M.T., Stringer, J.R., 2010. Stealth and opportunism: alternative lifestyles of species in the fungal genus *Pneumocystis*. Annu. Rev. Microbiol. 64, 431-452.

Cushion, M.T., Collins, M.S., Sesterhenn, T., Porollo, A., Vadukoot, A.K., Merino, E.J., 2016. Functional Characterization of *Pneumocystis carinii* Inositol Transporter 1. mBio 7, e01851-16.

Cushion, M.T., Ashbaugh, A., Hendrix, K., Linke, M.J., Tisdale, N., Sayson, S.G., Porollo, A., 2018. Gene expression of *Pneumocystis murina* after treatment with anidulafungin results in strong signals for sexual reproduction, cell wall integrity, and cell cycle arrest, indicating a requirement for ascus formation for proliferation. Antimicrob. Agents Chemother. 62, e02513-17
Davis, D.A., 2009. How human pathogenic fungi sense and adapt to pH: the link to virulence. Curr. Opin. Microbiol. 12, 365–70

de Boer, M.G., de Fijter, J.W., Kroon, F.P., 2011. Outbreaks and clustering of *Pneumocystis* pneumonia in kidney transplant recipients: a systematic review.
Med. Mycol. 49, 673-80.

Dei-Cas, E., Aliouat, E.M., Cailliez, J.C., 2004. Cellular structure. In: Walzer, P.D., Cushion, M.T. (Eds.). *Pneumocystis* Pneumonia, Thrid ed. Marcel Dekker Inc., New York, pp. 61-94.

Dei-Cas, E., Chabé, M., Moukhlis, R., Durand-Joly, I., Aliouat, E.M., Stringer, J.R., Cushion, M.T., Noël, C., de Hoog, G.S., Guillot, J., Viscogliosi, E., 2006. *Pneumocystis oryctolagi* sp. nov., an uncultured fungus causing pneumonia in rabbits at weaning: review of current knowledge, and description of a new taxon on genotypic, phylogenetic and phenotypic bases. FEMS Microbiol. Rev. 80, 853-871.

Deitsch, K.W., Lukehart, S.A., Stringer, J.R., 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat. Rev. Microbiol. 7, 493-503.

Delanoë, P., Delanoë, E., 1912. Sur les rapports des kystes de carinii du poumon des rats avec le *Trypanosoma lewisi*. CR Acad. Sci. 155, 658-660.

Demanche, C., Berthelemy, M., Petit, T., Polack, B., Wakefield, A. E., Dei-Cas, E., Guillot, J., 2001. Phylogeny of *Pneumocystis carinii* from 18 primate species confirms host specificity and suggests coevolution. J. Clin .Microbiol. 39, 2126-2133.

Desdevises, Y., Morand, S., Legendre, P., 2002. Evolution and determinants of host specificity in the genus Lamellodiscus (Monogenea). Biol. J. . Linnean Soc. 77, 431-443.

Domergue, R., Castano, I., De Las Penas, A., Zupancic, M., Lockatell, V., Hebel, J.R., Johnson, D., Cormack, B.P., 2005. Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTI. Science 308, 866-870.

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M.J., Donoghue, P.C.J., Yang, Z.H., 2015. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. Curr. Biol. 25, 2939-2950.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 2, 1185-1192.

Durand-Joly, I., Aliouat, E.M., Recourt, C., Guyot, K., Francois, N., Wauquier, M., Camus, D., Dei-Cas, E., 2002. *Pneumocystis carinii* f. sp. *hominis* is not infectious for SCID mice. J. Clin. Microbiol. 40, 1862-1865.

Edman, J.C., Kovacs, J.A., Masur, H., Santi, D.V., Elwood, H.J., Sogin, M.L., 1988. Ribosomal RNA sequence shows *Pneumocystis carinii* to be a member of the fungi. Nature 334, 519-522.

Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus,C.W., Kovacs, J.A., 1996. A single expression site with a conserved leader sequence regulates variation of expression of the *Pneumocystis carinii* family of major surface glycoprotein genes. DNA Cell Biol 15, 989 –999

Ellegren, H., Galtier, N., 2016. Determinants of genetic diversity. Nat. Rev. Genet. 17, 422.

Eriksson, O.E., Winka, K., 1997. Supraordinal taxa of Ascomycota. Myconet 1, 1-16.

Esteves, F., Gaspar, J., Tavares, A., Moser, I., Antunes, F., Mansinho, K., Matos, O., 2010. Population structure of *Pneumocystis jirovecii* isolated from immunodeficiency virus-positive patients. Infect. Genet. Evol. 10, 192-199.

Esteves, F., de Sousa, B., Calderon, E.J., Huang, L., Badura, R., Maltez, F., Bassat, Q., de Armas, Y., Antunes, F., Matos, O., 2016. Multicentre study highlighting clinical relevance of new high-throughput methodologies in molecular epidemiology of *Pneumocystis jirovecii* pneumonia. Clin. Microbiol. Infect. 22, 566.e9-566.e19.

English, K., Peters, S.E., Maskell, D.J., Collins, M.E., 2001. DNA analysis of *Pneumocystis* infecting a Cavalier King Charles spaniel. J. Eukaryot. Microbiol. 48, S106.

Fischer, J.M., Keely, S.P., Stringer, J.R., 2006. Evolutionary rate of ribosomal DNA in *Pneumocystis* species is normal despite the extraordinarily low copy-number of rDNA genes. J. Eukaryot. Microbiol. 53, S156-S158.

Florin-Christensen, M., Florin-Christensen, J., Wu, Y.P., Zhou, L., Gupta, A., Rudney, H.,Kaneshiro, E.S., 1994. Occurrence of specific sterols in *Pneumocystis carinii*. Biochem. Bioph. Res. Co. 198, 236–242.

Ford, C.B., Funt, J., Abbey, D., Issi, L., Guiducci, C., Martinez, D.A., Delorey, T., White, T.C., Cuomo, C., Rao, R.P., Berman, J., Thompson, D.A., Regey, A., 2015. The evolution of drug resistance in clinical isolates of *Candida albicans*. elife 4, e00662.

Frenkel, J., 1999. *Pneumocystis* pneumonia, an immunodeficiency-dependent disease (IDD): a critical historical overview. J. Eukaryot. Microbiol. 46, S89–S92.

Furuta, T., Fujita, M., Mukai, R., Sakakibara, I., Sata, T., Miki, K., Hayami, M., Kojima, S., Yoshikawa, Y., 1993. Severe pulmonary pneumocystosis in simian acquired immunodeficiency syndrome induced by simian immunodeficiency virus: its characterization by the polymerase-chain-reaction method and failure of experimental transmission to immunodeficient animals. Parasitol. Res. 69, 624-628.

Gigliotti, F., Harmsen, A.G., Haidaris, C.G., Haidaris, P.J., 1993. *Pneumocystis carinii* is not universally transmissible between mammalian species. Infect. Immun. 61, 2886-2890.

Giner, J-L., H. Zhao, D. H. Beach, E. J. Parish, K. Jayasimhulu, Kaneshiro, E. S., 2002. Comprehensive and definitive structural identities of *Pneumocystis carinii* sterols. J. Lipid Res. 43, 1114–1124.

Giraud, T., 2006. Speciation: selection against migrant pathogens: the immigrant inviability barrier in pathogens. Heredity 97, 316-318.

Giuntoli, D., Stringer, S.L., Stringer, J.R., 1994. Extraordinarily low number of ribosomal RNA genes in P. carinii. J. Eukaryot. Microbiol. 41, 88S.

Gołab, E., 2009. Presence of *Pneumocystis carinii* and *Pneumocystis wakefieldiae* DNA in the extrapulmonary tissues of immunocompromised laboratory rats. Wiad. Parazytol. 55,167-71.

Golan, J.J., Pringle, A., 2017. Long-Distance Dispersal of Fungi. Microbiol. Spectr. 5, doi: 10.1128/microbiolspec.FUNK-0047-2016.

Goto, S., Sugiyama, J., Hamamoto, M., Komagata, K., 1987. *Saitoella*, a new anamorph genus in the Cryptococcaceae to accommodate two himalayan yeast isolates formerly identified as *Rhodotorula glutinis*. J. Gen. Appl. Microbiol. 33, 75-85.

Guillot, J., Demanche, C., Hugot, J.P., Berthelemy, M., Wakefield, A.E., Dei-Cas, E., Chermette, R., 2001. Parallel phylogenies of *Pneumocystis* species and their mammalian hosts. J. Eukaryot. Microbiol. suppl, 113s-115s.

Guillot, J., Demanche, C., Norris, K., Wildschutte, H., Wanert, F., Berthelemy, M., Tataine, S., Dei-Cas, E., Chermette, R., 2004. Phylogenetic relationships among *Pneumocystis* from Asian macaques inferred from mitochondrial rRNA sequences. Mol. Phylogenet. Evol. 31, 988-996.

Haidaris, C.G., Medzihradsky, O., Gigliotti, F., Simpson-Haidaris, P.J., 1998. Molecular characterization of mouse *Pneumocystis carinii* surface glycoprotein A. DNA Res. 5, 77-85.

Hall, R.A., De Sordi, L., MacCallum, D.M., Topal, H., Eaton, R., et al., 2010. CO2 Acts as a Signalling Molecule in Populations of the Fungal Pathogen *Candida albicans*. PLOS Pathogens 6, e1001193.

Hauser, P.M., Francioli, P., Bille, J., Telenti, A., Blanc, D.S., 1997. Typing of *Pneumocystis carinii* f. sp. *hominis* by single-strand conformation polymorphism of four genomic regions. J. Clin. Microbiol. 35, 3086-3091.

Hauser, P.M. , Blanc, D.S., Bille, J., Francioli, P., 1998. Typing methods to approach *Pneumocystis carinii* genetic heterogeneity. FEMS Immunol. Med. Microbiol. 22, 27-35

Hauser, P.M., Burdet, F.X., Cissé, O.H., Keller, L., Taffe, P., Sanglard, D., Pagni, M., 2010. Comparative genomics suggests that the fungal pathogen *Pneumocystis* is an obligate parasite scavenging amino acids from its host's lungs. Plos One 5, e15152.

Hauser, P.M., 2014. Genomic insights into the fungal pathogens of the genus *Pneumocystis*: obligate biotrophs of humans and other mammals. PLoS Pathog. 10, e1004425.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 289.

Heitman, J., 2010. Evolution of eukaryotic microbial pathogens via covert sexual reproduction. Cell Host Microbe 8, 86-99.

Hill, W.G., Robertson, A., 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8, 269-294.

Holmes, E.C., 1991. Different Rates of Substitution May Produce Different Phylogenies of the Eutherian Mammals. J. Mol. Evol. 33, 209-215.

Hong, S.T., Steele, P.E., Cushion, M.T., Walzer, P.D., Stringer, S.L., Stringer, J.R., 1990. *Pneumocystis carinii* karyotypes. J. Clin. Microbiol. 28, 1785-1795.

Hughes, W.T., 1982. Natural mode of acquisition for *de novo* infection with *Pneumocystis carinii*. J. Infect. Dis. 145, 842-848.

Hugot, J. P., Demanche, C., Barriel, V., Dei-Cas, E., Guillot, J., 2003. Phylogenetic systematics and evolution of primate-derived *Pneumocystis* based on mitochondrial or nuclear DNA sequence comparison. Syst. Biol. 52, 735-744.

Icenhour, C.R., Kottom, T.J., Limper, A.H., 2003. Evidence for a melanin cell wall component in *Pneumocystis carinii*. Infect. Immun. 71, 5360-5363.

Icenhour, C.R., Arnold, J., Medvedovic, M., Cushion, M.T., 2006a. Competitive coexistence of two *Pneumocystis* species. Infect. Genet. Evol. 6, 177-186.

Icenhour, C.R., Kottom, T.J., Limper, A.H., 2006b. *Pneumocystis* melanins confer enhanced organism viability. Eukaryot. Cell 5, 916-923.

Inglis, D.O., Voorhies, M., Hocking Murray, D.R., Sil, A., 2013. Comparative transcriptomics of infectious spores from the fungal pathogen *Histoplasma capsulatum* reveals a core set of transcripts that specify infectious and pathogenic states. *Eukaryot. Cel*l 12, 828-852.

Jain, N., Li, L., Hsueh, Y.P., Guerrero, A., Heitman, J., Goldman, D.L., Fries, B.C., 2009. Loss of allergen 1 confers a hypervirulent phenotype that resembles mucoid switch variants of *Cryptococcus neoformans*. Infect. Immun. 77, 128-140.

Jarboui, M.A., Mseddi, F., Sellami, H., Sellami, A., Makni, F., Ayadi, A., 2013. Genetic diversity of *Pneumocystis jirovecii* strains based on sequence variation of different DNA region. Med. Mycol. 51, 561-567.

Jarne, P., 1995. Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. Genet. Res. 65, 193-207.

Kaneshiro, E. S., Ellis, J. E., Jayasimhulu, K. and Beach, D. H., 1994 Evidence for the presence of "metabolic sterols" in *Pneumocystis*: identification and initial characterization of *Pneumocystis carinii* sterols. J. Eukar. Microbiol. 41,78-85.

Kaneshiro, E.S., Amit Z., Swonger, M.M., Kreishman G.P., Brooks, E.E., Kreishman, M., Jayasimhulu, K., Parish, E.J., Sun, H., Kizito S.A., Beach, D.H., 1999. Pneumocysterol [(24Z)-ethylidenelanost-8-en-3β-ol], a rare sterol detected in the opportunistic pathogen Pneumocystis carinii hominis*:* Structural identity and chemical synthesis. Proceed. Nat. Acad. Sciences 96, 97-102.

Kasuga, T., White, T.J., Taylor, J.W., 2002. Estimation of nucleotide substitution rates in eurotiomycete fungi. Mol. Biol. Evol. 19, 2318-2324.

Keeling, P.J., Slamovits, C.H., 2005. Causes and effects of nuclear genome reduction. Curr. Opin. Genet. Dev. 15, 601-608.

Keeling, P.J., Corradi, N., Morrison, H.G., Haag, K.L., Ebert, D., Weiss, L.M., Akiyoshi, D.E., Tzipori, S., 2010. The reduced genome of the parasitic microsporidian *Enterocytozoon bieneusi* lacks genes for core carbon metabolism. Gen. Biol. Evol. 2, 304-309.

Keely, S. P., Fischer, J.M., Stringer, J.R., 2003a. Evolution and speciation of *Pneumocystis*. J. Eukaryot. Microbiol. 50 suppl, 624-626.

Keely, S.P., Cushion, M.T, Stringer S.J., Stringer, J.R., 2003b. Diversity at the locus associated with transcription of a variable surface antigen of *Pneumocystis carinii* as an index of population structure and dynamics in infected rats. Infect. Immun. 71, 47-60.

Keely, S.P., Fischer, J.M., Cushion, M.T., Stringer, J.R., 2004a. Phylogenetic identification of *Pneumocystis murina* sp. nov., a new species in laboratory mice. Microbiology 150, 1153-1165.

Keely, S.P., Renauld, H., Wakefield, A.E., Cushion, M.T., Smulian, A.G-, Fosker, N., Fraser, A., Harris, D., Murphy, L., Price, C., Quail, M.A., Seeger, K., Sharp, S., Tindal, C.J., Warren, T., Zuiderwijk, E., Barrell, B.G., Stringer, J.R., Hall, N., 2005. Gene arrays at *Pneumocystis carinii* telomeres. Genetics 170, 1589-600.

Keely, S.P., Linke M.J., Cushion M.T., Stringer, J.R., 2007. *Pneumocystis murina* MSG gene family and the structure of the locus associated with its transcription. Fung. Genet. Biol. 44, 905-919.

Keely, S.P., Stringer, J.R., 2009. Complexity of the MSG gene family of *Pneumocystis carinii*. BMC Genomics 10, 367.

Köhler, J.R., Hube, B., Puccia, R., Casadevall, A., Perfect, J.R., 2017. Fungi that Infect Humans. Microbiol Spectr. 5, doi: 10.1128-microbiolspec.FUNK-0014-2016.

Kottom, T.J., Limper, A.H., 2000. Cell wall assembly by *Pneumocystis carinii*. Evidence for a unique *gsc-1* subunit mediating beta -1,3-glucan deposition. J. Biol. Chem. 275, 40628-40634.

Kovacs, J.A., Powell, F., Edman, J.C., Lundgren, B., Martinez, A., Drew, B., Angus, W., 1993. Multiple genes encode the major surface glycoprotein of *Pneumocystis carinii*. J. Biol. Chem. 268, 6034-6040.

Kuo, C.H., Ochman, H., 2009. Deletional bias across the three domains of life. Gen. Biol. Evol. 1, 145-152.

Kutty, G., Ma, L., Kovacs, J.A., 2001. Characterization of the expression site of the major surface glycoprotein of human-derived *Pneumocystis carinii*. Mol. Microbiol. 42, 183–193.

Kutty, G., Maldarelli, F., Achaz, G., Kovacs, J.A., 2008. Variation in the major surface glycoprotein genes in *Pneumocystis jirovecii*. J. Infect. Dis. 198, 741-749.

Kutty, G., Davis, A. S., Ma, L., Taubenberger, J.K., Kovacs, J.A., 2015. *Pneumocystis* encodes a functional endo-beta-1,3-glucanase that is expressed exclusively in cysts. J. Infect. Dis. 211, 719-728.

Kutty, G., Davis, A.S., Ferreyra, G.A., Qiu, J., Huang, D.W., Sassi, M., Bishop, L., Handley, G., Sherman, B., Lempicki, R., Kovacs, J.A., 2016. Beta-glucans are masked but contribute to pulmonary inflammation during *Pneumocystis* pneumonia. J. Infect. Dis. 214, 782-791.

Laakkonen, J., 1998. *Pneumocystis carinii* in wildlife. Int. J. Parasitol. 28, 241-252.

Latgé, J.P., 2007. The cell wall: a carbohydrate armour for the fungal cell. Mol. Microbiol. 66, 279-290.

Latinne, A., Beze, F., Delhaes, L., Pottier, M., Gantois, N., Nguyen, J., Blasdell, K., Dei-Cas, E., Morand, S., Chabé, M., 2017. Genetic diversity and evolution of *Pneumocystis* fungi infecting wild Southeast Asian murid rodents. Parasitology 1-16. doi: 10.1017/S0031182017001883.

Latouche, S., Totet, A., Lacube, P., Bolognini, J., Nevez, G., Roux, P., 2001. Development of an RT-PCR on the heat shock protein 70 gene for viability detection of *Pneumocystis carinii* f. sp. *hominis* in patients with pneumocystosis and in air sample. J. Eukaryot. Microbiol. 48, S176-S177.

Le Gal, S., Damiani, C., Rouillé, A., Grall, A., Tréguer, L., Virmaux, M., Moalic, E., Quinio, D., Moal, M.C., Berthou, C., Saliou, P., Le Meur, Y., Totet, A., Nevez, G., 2012. A cluster of *Pneumocystis* infections among renal transplant recipients: molecular evidence of colonized patients as potential infectious sources of *Pneumocystis jirovecii*. Clin. Infect Dis. 54, e62-71.

Lee, C.H., Lu, J.J., Bartlett, M.S., Durkin, M.M., Liu, T.H., Wang, J., Jiang, B., Smith, J.W., 1993. Nucleotide sequence variation in *Pneumocystis carinii* strains that infect humans. J. Clin. Microbiol. 31, 754-757.

Liu, Y., Leibowitz, M.J., 1993. Variation and in vitro splicing of group I introns in rRNA genes of *Pneumocystis carinii*. Nucl. Ac. Res. 21, 2415-2421.

Liu, Y., Tidwell, R.R., Leibowitz, M.J., 1994. Inhibition of in vitro splicing of a group I intron of *Pneumocystis carinii*. J. Eukaryot. Microbiol. 41, 31-38.

Liu, Y., Leigh, J.W., Brinkmann, H., Cushion, M.T., Rodriguez-Ezpeleta, N., Philippe, H., Lang, B.F., 2009. Phylogenomic analyses support the monophyly of Taphrinomycotina, including *Schizosaccharomyces* fission yeasts. Mol. Biol. Evol. 26, 27-34.

Lu, J.-J., Bartlett, M.S., Smith, J.W., Lee, C.-H., 1995. Typing of *Pneumocystis carinii* strains with type-specific oligonucleotide probes derived from nucleotide sequences of internal transcribed spacers of rRNA genes. J. Clin. Microbiol. 33, 2973-2977.

Lugli, E.B., Bampton, E.T., Ferguson, D.J., Wakefield, A.E., 1999. Cell surface protease PRT1 identified in the fungal pathogen *Pneumocystis carinii*. Mol. Microbiol. 31, 1723-1733.

Lundgren, B., Cotton, R., Lundgren, J.D., Edman, J.C., Kovacs, J.A., 1990. Identification of *Pneumocystis carinii* chromosomes and mapping of five genes. Infect. Immun. 58, 1705-1710.

Ma, L., Kutty, G., Jia, Q., Imamichi, H., Huang, L., Atzori, C., Beckers, P., Groner, G., Beard, C.B., Kovacs, J.A., 2002. Analysis of variation in tandem repeats in the intron of the major surface glycoprotein expression site of the human form of *Pneumocystis carinii*. J. Infect. Dis. 186, 1647-1654.

Ma, L., Huang da, W., Cuomo, C.A., Sykes, S., Fantoni, G., Das, B., Sherman, B.T., Yang, J., Huber, C., Xia, Y., Davey, E., Kutty, G., Bishop, L., Sassi, M., Lempicki, R.A., Kovacs, J.A., 2013. "Sequencing and characterization of the complete mitochondrial genomes of three *Pneumocystis* species provide new insights into divergence between human and rodent *Pneumocystis*. Faseb J. 27, 1962-1972.

Ma, L., Chen, Z., Huang da, W., Kutty, G., Ishihara, M., Wang, H., Abouelleil, A., Bishop, L., Davey, E., Deng, R., Deng, X., Fan, L., Fantoni, G., Fitzgerald, M., Gogineni, E., Goldberg, J.M., Handley, G., Hu, X., Huber, C., Jiao, X., Jones, K., Levin, J.Z., Liu, Y., Macdonald, P., Melnikov, A., Raley, C., Sassi, M., Sherman, B.T., Song, X., Sykes, S., Tran, B., Walsh, L., Xia, Y., Yang, J., Young, S., Zeng, Q., Zheng, X., Stephens, R., Nusbaum, C., Birren, B.W., Azadi, P., Lempicki, R.A., Cuomo, C.A., Kovacs, J.A., 2016a. Genome analysis of three *Pneumocystis* species reveals adaptation mechanisms to life exclusively in mammalian hosts. Nat. Commun. 7, 10740.

Ma, L., Raley, C., Zheng, X., Kutty, G., Gogineni, E., Sherman, B.T., Sun, Q., Chen, X., Skelly, T., Jones, K., Stephens, R., Zhou, B., Lau, W., Johnson, C., Imamichi, T., Jiang, M., Dewar, R., Lempicki, R.A., Tran, B., Kovacs, J.A., Huang da, W., 2016b. Distinguishing highly similar gene isoforms with a clustering-based bioinformatics analysis of PacBio single-molecule long reads. BioData mining 9, 13.

Maher, N.H., Vermund, S., Welsh, D.A., Dillon, H.K., Awooda, A., Unnasch, T.R., 2001. Development and characterization of a molecular viability assay for *Pneumocystis carinii* f sp *hominis*. J. Infect. Dis. 183, 1825-1827.

Maitte, C., Leterrier, M., Le Pape, P., Miegeville, M., Morio, F., 2013. Multilocus sequence typing of *Pneumocystis jirovecii* from clinical samples: how many and which loci should be used? J. Clin. Microbiol. 51, 2843-2849.

Matos, O., Esteves, F., 2010. *Pneumocystis jirovecii* multilocus gene sequencing: findings and implications. Fut. Microbiol. 5, 1257-1267.

Martinez, A., Aliouat, E.M., Standaert-Vitse, A., Werkmeister, E., Pottier, M., Pincon, C., Dei-Cas, E., Aliouat-Deni,s C.M., 2011. Ploidy of cell-sorted trophic and cystic forms of *Pneumocystis carinii*. PLoS One 6, e20935.

Martinez, A., Halliez, M.C., Aliouat, E.M., Chabé, M., Standaert-Vitse, A., Frealle, E., Gantois, N., Pottier, M., Pinon, A., Dei-Cas, E., Aliouat-Denis, C.M., 2013. Growth and airborne transmission of cell-sorted life cycle stages of *Pneumocystis carinii*. PLoS One 8, e79958.

Matsumoto, Y., Yoshida, Y., 1984. Sporogony in *Pneumocystis carinii*: synaptonemal complexes and meiotic nuclear divisions observed in precysts. J. Protozool. 31, 420-428.

Mazars, E., Guyot, K., Durand, I., DeiCas, E., Boucher, S., BenAbderrazak, S., Banuls, A.L., Tibayrenc, M., Camus, D., 1997. Isoenzyme diversity in *Pneumocystis carinii* from rats, mice, and rabbits. J. Infect. Dis. 175, 655-660.

Medoff, G., Painter, A., Kobayashi, GS., 1987. Mycelial- to yeast-phase transitions of the dimorphic fungi *Blastomyces dermatitidis* and *Paracoccidioides brasiliensis*. J. Bacteriol. 169, 4055-4060.

Montes-Cano, M.A., Chabé, M., Fontillon-Alberdi, M., de-Lahorra, C., Respaldiza, N., Medrano, F.J., Varela, J.M., Dei-Cas, E., Calderon, E.J., 2009. Vertical transmission of *Pneumocystis jirovecii* in humans. Emerg. Infect. Dis. 15, 125-127.

Nahimana, A., Francioli, P., Blanc, D.S., Bille, J., Wakefield, A.E., Hauser, P.M., 2000a. Determination of the copy number of the nuclear rDNA and beta-tubulin genes of *Pneumocystis carinii* f. sp *hominis* using PCR multicompetitors. J. Eukar. Microbiol. 47, 368-372.

Nahimana, A., Blanc, D.S., Francioli, P., Bille, J., Hauser, P.M., 2000b. Typing of *Pneumocystis carinii* f. sp. *hominis* by PCR-SSCP to indicate a high frequency of co-infections. J. Med. Microbiol. 49, 753-758.

Nahimana, A., Cushion ,M.T., Blanc, D.S., Hauser, P.M., 2001. Rapid PCR-single-strand conformation polymorphism method to differentiate and estimate relative abundance of *Pneumocystis carinii* special forms infecting rats. J. Clin. Microbiol. 39, 4563-4565.

Nagy, L.G., Ohm, R.A., Kovács, G.M., Floudas, D., Riley, R., Gácser, A., Sipiczki, M., Davis, J.M., Doty, S.L., de Hoog, G.S., Lang, B.F., Spatafora, J.W., Martin, F.M., Grigoriev, I.V., Hibbett, D.S., 2014. Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. Nat. Commun. 18, 4471.

Nagy, G., 2017. Complex Multicellular Life with 5,500 Genes. Curr. Biol. 27, R609-R612.

Nguyen, T.A., Cissé, O.H., Yun Wong, J., Zheng, P., Hewitt, D., Nowrousian, M., Stajich, J.E., Jedd, G., 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota. Nat. Commun. 8, 14444.

Noble, L.M., Andrianopoulos, A., 2013. Fungal genes in context: genome architecture reflects regulatory complexity and function. Gen. Biol. Evol. 5, 1336-1352.

Nollstadt, K.H., Powles, M.A., Fujioka, H., Aikawa, M., Schmatz, D.M., 1994. Use of beta-1,3-glucan-specific antibody to study the cyst wall of *Pneumocystis carinii* and effects of pneumocandin B0 analog L-733,560. Ant. Ag. Chemoth. 38, 2258-2265.

O'Malley, M.A., Wideman, J.G., Ruiz-Trillo, I., 2016. Losing complexity: the role of simplification in macroevolution. Trends Ecol. Evol. 31, 608-621.

Otto, S.P., Lenormand, T., 2002. Resolving the paradox of sex and recombination. Nat. Rev. Genet. 3, 252-261.

Palmer, R.J., Settnes, O.P., Lodal, J., Wakefield, A.E., 2000. Population structure of rat-derived *Pneumocystis carinii* in Danish wild rats. Appl. Environ. Microbiol. 66, 4954-4961.

Papkou, A., Gokhale, C.S., Traulsen, A., Schulenburg, H., 2016. Host-parasite coevolution: why changing population size matters. Zoology 119, 330-338.

Parker, I.M., Gilbert, G.S., 2004. The evolutionary ecology of novel plant-pathogen interactions. An. Rev. Ecol. Evol. Syst. 35, 675-700.

Parobek, C.M., Jiang, L.Y., Patel, J.C., Alvarez-Martinez, M.J., Miro, J.M., Worodria, W., Andama, A., Fong, S., Huang, L., Meshnick, S.R., Taylor, S.M., Juliano, J.J., 2014. Multilocus microsatellite genotyping array for investigation of genetic epidemiology of *Pneumocystis jirovecii*. J. Clin. Microbiol. 52, 1391-1399.

Peters, S.E., English, K., Laakkonen, J., Gurnell, J., 1994. DNA analysis of *Pneumocystis carinii* infecting Finnish and English shrews. J. Eukaryot. Microbiol. 41, S108.

Peterson, J.C, Cushion, M.T., 2005. *Pneumocystis*: not just pneumonia. Curr. Opin. Microbiol. 8, 393-8

Pollak, E., 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117, 353-360.

Porollo, A., Sesterhenn, T.M., Collins, M.S., Welge, J.A., Cushion, M.T., 2014. Comparative genomics of *Pneumocystis* species suggests the absence of genes for myo-inositol synthesis and reliance on inositol transport and metabolism. mBio 5, e01834.

Poulin, R., Krasnov, B.R., Morand, S., 2006. Patterns of host specificity in parasites exploiting small mammals. Micromammals and Macroparasites: From Evolutionary Ecology to Management, 233-256.

Pringle, A, Baker, D.M., Platt, J.L., Wares, J.P., Latgé, J.P., Taylor, J.W., 2005. Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. Evolution 59, 1886-1899.

Pritchard, J. K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Restrepo, S., Tabima, J.F., M.F. Mideros, N.J. Grunwald, Matute, D.R., 2014. Speciation in fungal and oomycete plant pathogens. Annu. Rev. Phytopathol. 52. 289-316.

Richard, S., Almeida, J.M.G. C.F., Cissé, O.H., Luraschi, A., Nielsen, O., Pagni, M., Hauser, P.M., 2018. Functional and expression analyses of the *Pneumocystis MAT* genes suggest obligate sexuality through primary homothallism within host's lungs. mBio 9, e02201-17.

Rieseberg, L.H., 2001. Chromosomal rearrangements and speciation. Trend. Ecol. Evol. 16, 351-358.

Rhind N, Chen Z, Yassour M, Thompson DA, Haas BJ, Habib N, Wapinski I, Roy S, Lin MF, Heiman DI, Young SK, Furuya K, Guo Y, Pidoux A, Chen HM, Robbertse B, Goldberg JM, Aoki K, Bayne EH, Berlin AM, Desjardins CA, Dobbs E, Dukaj L, Fan L, FitzGerald MG, French C, Gujja S, Hansen K, Keifenheim D, Levin JZ, Mosher RA, Müller CA, Pfiffner J, Priest M, Russ C, Smialowska A, Swoboda P, Sykes SM, Vaughn M, Vengrova S, Yoder R, Zeng Q, Allshire R, Baulcombe D, Birren BW, Brown W, Ekwall K, Kellis M, Leatherwood J, Levin H, Margalit H, Martienssen R, Nieduszynski CA, Spatafora JW, Friedman N, Dalgaard JZ, Baumann P, Niki H, Regev A, Nusbaum C. 2011. Comparative functional genomics of the fission yeasts. Science 332, 930–936.

Rhodes, J., Beale, M.A,, Vanhove, M., Jarvis, J.N., Kannambath, S., Simpson, J.A., Ryan, A., Meintjes, G., Harrison, T.S., Fisher, M.C., Bicancic, T., 2017. A population genomics approach to assessing the genetic basis of within-host microevolution underlying recurrent cryptococcal meningitis infection. G3 (Bethes.) 3, 1165-1176.

Roy S.W., Gilbert W., 2005. Rates of intron loss and gain: Implications for early eukaryotic evolution. Proc Natl Acad Sci USA. 102, 5773–5778.

Sanchez, C.A., Chabé ,M., Aliouat, E.M., Durand-Joly, I., Gantois, N., Conseil, V., Lo´ pez, C., Duriez, T., Dei-Cas, E., Vargas, S.L., 2007. Exploring transplacental transmission of *Pneumocystis oryctolagi* in first-time pregnant and multiparous rabbit does. Med. Mycol. 45, 701–707.

Sasal, P., Trouve, S., Muller-Graf , C., Morand, S., 1999. Specificity and host predictability: a comparative analysis among monogenean parasites of fish. J. An. Ecol. 68, 437-444.

Schaffzin, J.K., Stringer, J.R., 2000. The major surface glycoprotein expression sites of two special forms of rat *Pneumocystis carinii* differ in structure. J. Infect. Dis. 181, 1729-1739.

Schmid-Siegert, E, Richard, S., Luraschi, A., Mühlethaler, K., Pagni, M., Hauser, P.M., 2017. Mechanisms of surface antigenic variation in the human pathogenic fungus *Pneumocystis jirovecii*. mBio 8, e01470-17.

Schoch, C.L. Sung, G.H., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., Hofstetter, V., Robbertse, B., Matheny, P.B., Kauff, F., Wang, Z., Gueidan, C., Andrie, R.M., Trippe, K., Ciufetti, L.M., Wynns, A., Fraker, E., Hodkinson, B.P., Bonito, G., Groenewald, J.Z., Arzanlou, M., de Hoog, G.S.,, Crous, P.W., Hewitt, D., Pfister, D.H., Peterson, K., Gryzenhout, M., Wingfield, M.J., Aptroot, A., Suh, S.O., Blackwell, M., Hillis, D.M., Griffith, G.W., Castlebury, L.A., Rossman, A.Y., Lumbsch, H.T., Lücking, R., Büdel, B., Rauhut, A., Diederich, P., Ertz, D., Geiser, D.M., Hosaka, K., Inderbitzin, P., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Mostert, L., O'Donnell, K., Sipman, H., Rogers, J.D., Shoemaker, R.A., Sugiyama, J., Summerbell, R.C., Untereiner, W., Johnston, P.R., Stenroos, S., Zuccaro, A., Dyer, P.S., Crittenden, P.D., Cole, M.S., Hansen, K., Trappe, J.M,, Yahr. R,, Lutzoni. F,, Spatafora. J,W., 2009. The Ascomycota

tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Systematic biology, 58, 224-239.

Schulze-Lefert, P., Panstruga, R., 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trend. Plant Sci. 16, 117-125.

Sesterhenn, T.M., Slaven, B.E., Keely, S.P., Smulian, A.G., Lang, B.F., Cushion, M.T., 2010. Sequence and structure of the linear mitochondrial genome of *Pneumocystis carinii*." Mol. Gen. Gen. 283, 63-72.

Shannon C. 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27:379–423.

Simpson, E.H., 1949. Measurement of diversity. Nature, 163, 688

Slaven, B.E., Meller, J., Porollo, A., Sesterhenn, T., Smulian, A.G.,. Cushion, M.T., 2006. Draft assembly and annotation of the *Pneumocystis carinii* genome. J. Eukaryot. Microbiol. 53 suppl, S89-91.

Smit, A.F.A,. Hubley, R., Green, P., 2013. *RepeatMasker Open-4.0*. <http: www.repeatmasker.org>.

Slamovits, C.H., Fast, N.M., Law, J.S., Keeling, P.J., 2004. Genome compaction and stability in microsporidian intracellular parasites. Curr. Biol, 14, 891-896.

Stajich, J.E., Dietrich, F.S., Roy, S.W., 2007. Comparative genomic analysis of fungal genomes reveals intron-rich ancestors. Gen. Biol. 8, R223.

Stringer, J.R., 1996. *Pneumocystis carinii*: what is it, exactly? Clinical Microbiol. Rev. 9, 489-498.

Stringer, J.R., 2007. Antigenic variation in *Pneumocystis*. J. Eukaryot. Microbiol. 54, 8 -13.

Stringer, J.R., Stringer, S.L., Zhang, J, Baughman, R., Smulian, A.G., Cushion, M.T., 1993. Molecular genetic distinction of *Pneumocystis carinii* from rats and humans. J. Eukaryot. Microbiol. 40, 733-741.

Stringer, J.R., Cushion, M.T., 1998. The genome of *Pneumocystis carinii*. FEMS Immunol. Med. Microbiol. 22, 15-26.

Sun, L., Huang, M., Wang, J., Xue, F., Hong, C., Guo, Z., Gu, J., 2015. Genotyping of *Pneumocystis jirovecii* isolates from human immunodeficiency virus-negative patients in China. Infect. Genet. Evol. 31, 209-215.

Sunkin, S.M., Stringer, S.L., Stringer, J.R., 1994. A tandem repeat of rat-derived *Pneumocystis carinii* genes encoding the major surface glycoprotein. J. Eukaryot. Microbiol. 41, 292-300.

Sunkin, S.M., Stringer, J.R, 1996. Translocation of surface antigen genes to a unique telomeric expression site in *Pneumocystis carinii*. Mol. Microbiol., 19, 283-295.

Sugiyama, J., Hosaka, K., Suh, S.O., 2006. Early diverging Ascomycota: phylogenetic divergence and related evolutionary enigmas. Mycologia 98, 996-1005.

Tanabe, K., Fuchimoto, M., Egawa, K., Nakamura, Y., 1988. Use of *Pneumocystis carinii* genomic DNA clones for DNA hybridization analysis of infected human lungs. J. Infect. Dis. 157, 593-596.

Tang, X., Bartlett, M.S., Smith, J.W., Lu, J.J., Lee, C.H., 1998. Determination of copy number of rRNA genes in *Pneumocystis carinii* f. sp. *hominis*. J .Clin. Microbiol. 36, 2491-2494.

Thomas, C.F., Jr., Leof, E.B., Limper, A.H., 1999. Analysis of *Pneumocystis carinii* introns. Infect. Immun. 67, 6157-6160.

Tibayrenc, M., Ayala, F.J., 2012. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc. Natl. Acad. Sci. USA 109, E3305-3313.

Tibayrenc, M., Ayala, F.J., 2014. *Cryptosporidium, Giardia, Cryptococcus, Pneumocystis* genetic gariability: cryptic biological species or clonal near-clades? PLoS Pathog. 10, e1003908.

Tsai, I.J., Tanaka, E., Masuya, H., Tanaka, R., Hirooka, Y., Endoh, R., Sahashi, N., Kikuchi, T., 2104. Comparative genomics of *Taphrina* fungi causing varying degrees of tumorous deformity in plants. *Gen. Biol. Evol.* 6, 861-872.

Valero, C., Buitrago, M.J., Gits-Muselli, M., Benazra, M., Sturny-Leclère, A., Hamane, S., Guigue, N., Bretagne, S., Alanio, A., 2016. Copy number variation of mitochondrial DNA genes in *Pneumocystis jirovecii* according to the fungal load in BAL specimens. Front. Microbiol. 7, 1413.

van der Does, H. C., Rep, M., 2007. Virulence genes and the evolution of host specificity in plant-pathogenic fungi. Mol. Plant-Microbe Interact. 20, 1175-1182.

Vasquez, J., Smulian, A.G., Linke, M.J., Cushion, M.T., 1996. Antigenic differences associated with genetically distinct *Pneumocystis carinii* from rats. Infect. Immun. 64, 290-297.

Venema, J., Tollervey, D., 1999. Ribosome synthesis in *Saccharomyces cerevisiae*. Annu. Rev. Genet. 33, 261-311.

Vestereng, V.H., Kovacs, J.A., 2004. Inability of *Pneumocystis* organisms to incorporate bromodeoxyuridine suggests the absence of a salvage pathway for thymidine. Microbiol. 150, 1179-1182.

Vohra, P.K., Park, J.G., Sanyal, B., Thomas, C.F., Jr., 2004. Expression analysis of PCSTE3, a putative pheromone receptor from the lung pathogenic fungus *Pneumocystis carinii*. Biochem. Biophys. Res. Commun. 319, 193-199.

Wada, M., Sunkin, S.M., Stringer, J.R., Nakamura, Y., 1995. Antigenic variation by positional control of major surface glycoprotein gene expression in *Pneumocystis carinii*. J. Infect. Dis. 171,1563-1568.

Wakefield, A.E., 1996. DNA sequences identical to *Pneumocystis carinii* f. sp. *carinii* and *Pneumocystis carinii* f. sp. *hominis* in samples of air spora. J. Clin. Microbiol. 34, 1754-1759.

Wakefield, A.E., 1998a. Genetic heterogeneity in *Pneumocystis carinii*: an introduction. FEMS Immunol. Med. Microbiol. 22, 5-13.

Wakefield, A.E. 1998b. Genetic heterogeneity in human-derived *Pneumocystis carinii*. FEMS Immunol. Med. Microbiol. 22, 59-65.

Wakefield, A.E., Fritscher, C.C., Malin, A.S., Gwanzura, L., Hughes, W.T., Miller, R.F. 1994. Genetic diversity in human-derived *Pneumocystis carinii* isolates from four geographical locations shown by analysis of mitochondrial rRNA gene sequences. J. Clin. Microbiol. 32, 2959-2961.

Wernegreen, J.J., 2015. Endosymbiont evolution: predictions from theory and surprises from genomes. Ann. NY Acad .Sci. 1360, 16-35.

Wolf, Y.I., Koonin, E.V., 2013. Genome reduction as the dominant mode of evolution. Bioessays 35, 829–837.

Wood, V,, Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat , N., Hayles. J,, Baker. S,, Basham, D., Bowman, S,, Brooks, K,. Brown, D., Brown, .S, Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Feltwell ,T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E.J,, Hunt. S,, Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neil, S., Pearson, D., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M,, Squares, R., Squares, S., Stevens, K,, Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R,, Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schäfer, M., Müller-Auer, S., Gabel, C., Fuchs, M., Düsterhöft, A., Fritzc, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dréano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sánchez, M,. del Rey, F., Benito, J., Domínguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T.,

McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G., Nurse, P., 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature 415, 871-880.

Wyder, M. A., Rasch, E. M. Kaneshiro, E.S., 1998. Quantitation of absolute *Pneumocystis carinii* nuclear DNA content. Trophic and cystic forms isolated from infected rat lungs are haploid organisms. J. Eukaryot. Microbiol. 45, 233-239.

Ye, D., Lee, C.H., Queener, S.F., 2001. Differential splicing of *Pneumocystis carinii* f. sp. *carinii* inosine 5'-monophosphate dehydrogenase pre-mRNA. Gene 263, 151-158.

Yoganathan, T., Lin, H., Buck, G.A., 1989. An electrophoretic karyotype and assignment of ribosomal genes to resolved chromosomes of *Pneumocystis carinii*. Mol. Microbiol. 3, 1473-1480.

Zhu, T., Niu, D.K., 2013. Mechanisms of intron loss and gain in the fission seast *Schizosaccharomyces*. PLOS ONE 8, e61683.

LEGENDS FIGURES

Figure 1 | Cell cycle

The whole cell cycle of *Pneumocystis* species would take place within the host's lungs, airborne asci ensuring transmission to new hosts. The cycle is thought to include two phases: sexual and asexual. The trophic forms tightly adhere to the host's alveolar epithelial pneumocytes type I, whereas asci are generally localized within the alveolar lumen. The ring shown in green might allow the formation of a rent upon contact with humidity and so the release of the ascospores. This ring may correspond to the parentheses-like structure visible on Figure 2. This Figure does not include new features relatively to models previously proposed.

Figure 2 | Cluster of *P. jirovecii* asci

Cluster of *P. jirovecii* asci stained with Grocott's Methenamine silver (Churukian and Schenk, 1977) within a patient's bronchoalveolar lavage. The structures darker than the rest of the wall on each ascus are the parentheses-like structure (picture from the Institute of Microbiology, Lausanne University Hospital).

Figure 3 | Genome composition of *Pneumocystis* and related fungi

Protein coding genes, intergenic spaces, and intron positions were obtained from NCBI (https://www.ncbi.nlm.nih.gov/, last accessed 2018-03-20). Curated *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* intron data were extracted respectively from Pombase database (Wood et al., 2002; https://www.pombase.org/downloads/intron-data, last accessed 2018-03-20) and *Saccharomyces* Genome database (Cherry et al., 1998; https://www.yeastgenome.org, last accessed 2018-03-20). Repeats include DNA transposons, retrotransposons, and simple low complexity repeats proportions as roughly estimated using

RepeatMasker (Smit et al., 2013) and RepBase database (Bao et al., 2015). The proportions of MSGs were calculated based on data from Ma et al. (2016a). Ribosomal DNA cassettes include each three genes (rDNA) and two internal transcribed spacers (ITS): 18S rDNA-ITS1-5.8 rDNA-ITS2-26S rDNA. *S. pombe* genome encodes roughly 140 copies of a cassette of a size of 5.8 kb. In *S. cerevisiae,* ca. 150 tandem copies of a 9.1 kb cassette are present (Venema and Tollervey, 1999). In contrast, *Pneumocystis* species harbor each a single rDNA cassette of 11 kb.

Figure 4 | Graphical overview of the hypothetical evolutionary history of *Pneumocystis* species *Pneumocystis* species divergence timing has been determined elsewhere (Keely et al., 2003a; 2004a; Beimforde et al., 2014). Losses of multiple metabolic pathways, as well as contraction and expansion of specific gene families are presented (Table 2). Note that the timing and order of losses is unknown. The gain and loss of specific functions for *Pneumocystis* is inferred here to have occurred in the last most recent ancestor common of *Pneumocystis* species (MRCA) because the underlying genes are absent in the genomes of all *Pneumocystis* sequenced to date. The MSG superfamily emerged in *Pneumocystis* ancestry and displays a substantial level of lineage specific divergence (represented by blue triangles). Intron loads are similar among *Pneumocystis* species, which might suggest a common origin. The fission yeast clade diverged ~250 MYA ago (Rhind et al. 2011) and has lost most of the introns acquired from an intron rich ancestor (Roy et al., 2005; Stajich et al. 2007; Rhind et al., 2011). Although there is no dating estimates for the intron loss in fission clade, the absence of recent intron gains and the low rates of intron loss (Zhu and Niu, 2013) suggest that the majority of introns were lost before the diversification of the fission yeast clade. The colors of the lines representing the evolving species signify different nutritional modes (dark green, saprophytism; light green, gradual shift from saprophytism to the parasitism; yellow, animal parasitism). We assume that the MRCA of

Taphrinomycota subphylum was a multicellular or dimorphic saprotroph based on ancestral traits reconstruction (Schoch et al., 2009; Nguyen et al., 2017). The phylogenetic relationship presented here is consistent with published phylogenies (Liu et al., 2009; Sugiyama et al., 2006). RRM correspond to RNA binding proteins harboring an RNA recognition motif.

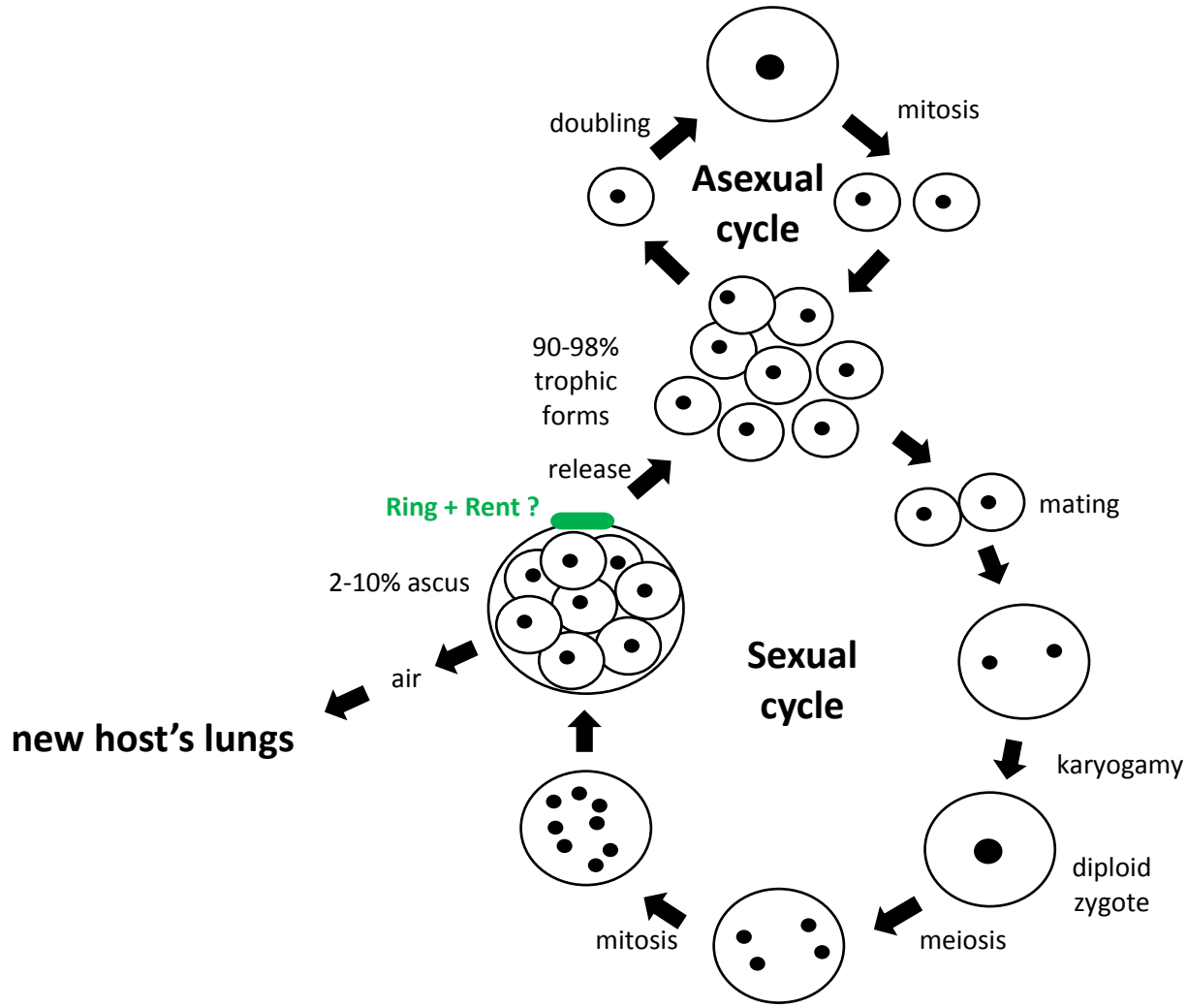
doubling
mitosis
Asexual cycle
90-98% trophic forms
release
Ring + Rent ?
mating
2-10% ascus
air
new host's lungs
Sexual cycle
karyogamy
diploid zygote
meiosis
mitosis

Figure 1.

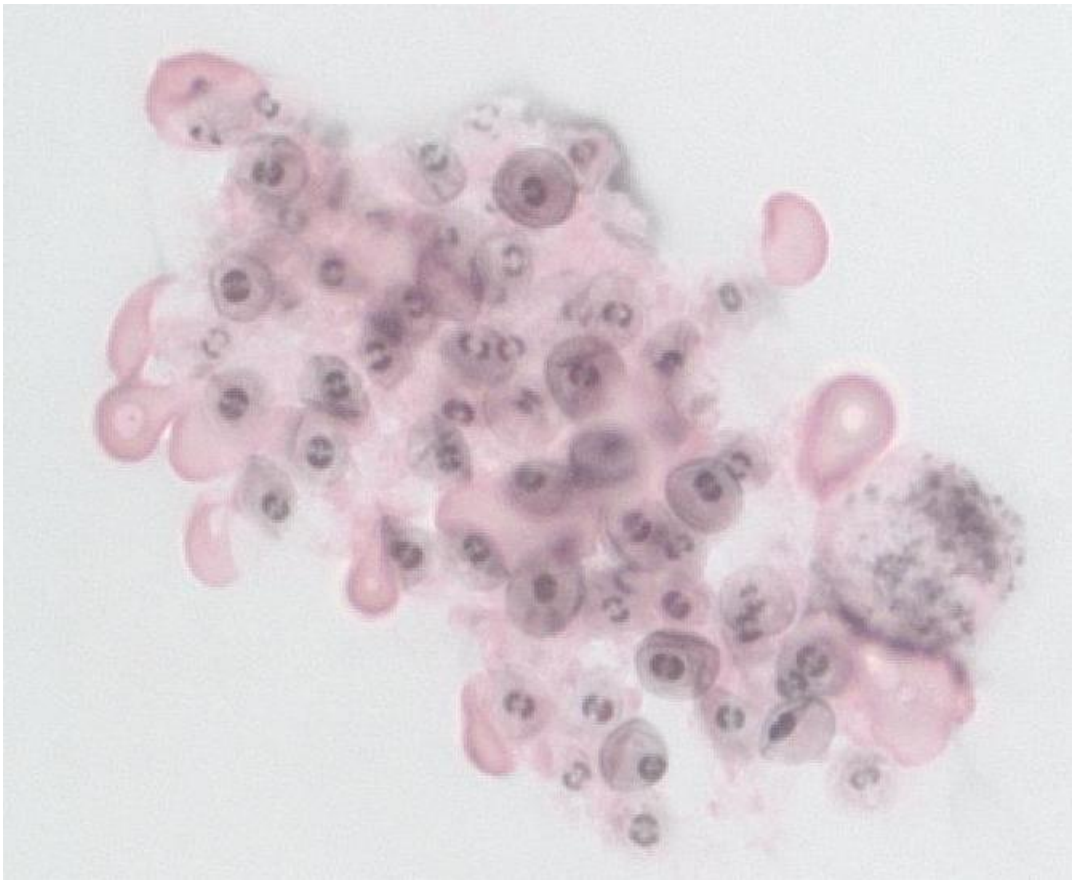

Figure 2.

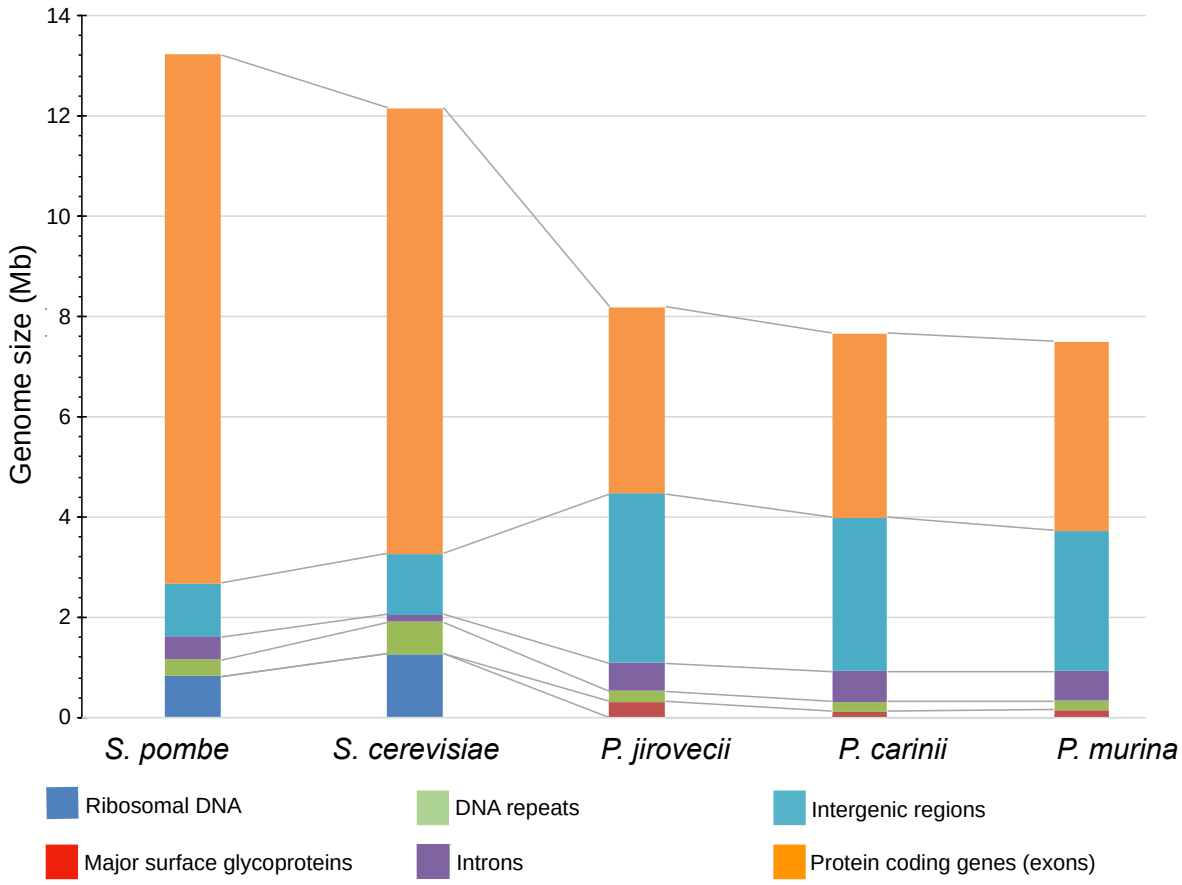
14
12
10
8
6
4
2
0
Genome size (Mb)
S. pombe
S. cerevisiae
P. jirovecii
P. carinii
P. murina
Ribosomal DNA
DNA repeats
Intergenic regions
Major surface glycoproteins
Introns
Protein coding genes (exons)

Figure 3.

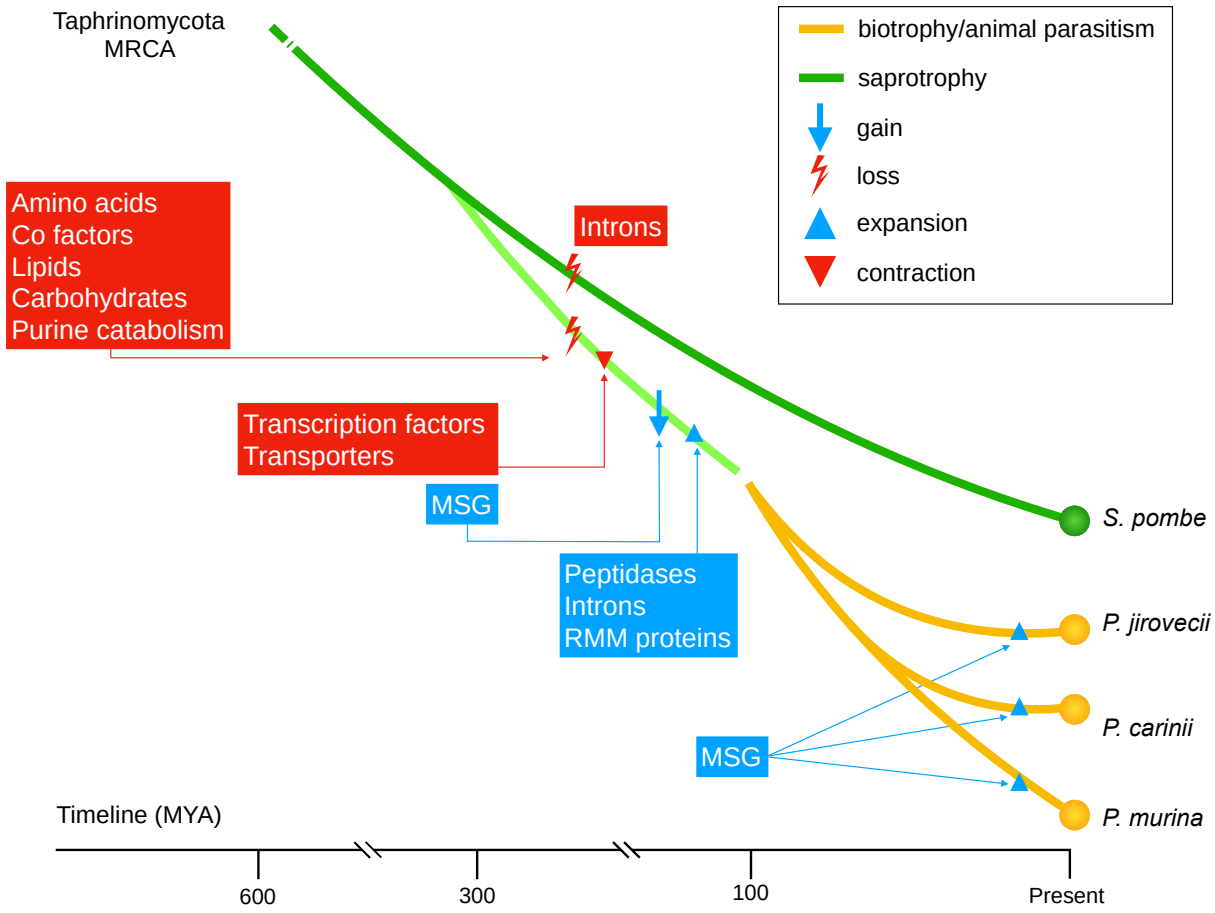
Taphrinomycota
MRCA
biotrophy/animal parasitism
saprotrophy
gain
loss
expansion
contraction
Amino acids
Co factors
Lipids
Carbohydrates
Purine catabolism
Introns
Transcription factors
Transporters
MSG
Peptidases
Introns
RMM proteins
MSG
S. pombe
P. jirovecii
P. carinii
P. murina
Timeline (MYA)
600
300
100
Present

Figure 4.