UNIVERSITA' DEGLI STUDI DI SASSARI

SCUOLA DI DOTTORATO IN SCIENZE BIOMOLECOLARI E BIOTECNOLOGICHE

(Intenational PhD School in Biomolecular and Biotechnological Sciences)

Indirizzo: Microbiologia molecolare e clinica

Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*

Direttore della scuola:
Prof. Masala Bruno

Relatore:
Prof. Pier Luigi Fiori

Correlatore:
Dott. Eric Viscogliosi

Tesi di Dottorato:
Dionigia Meloni

XXIV CICLO

UNIVERSITA' DEGLI STUDI DI SASSARI

SCUOLA DI DOTTORATO IN SCIENZE BIOMOLECOLARI E BIOTECNOLOGICHE

(Intenational PhD School in Biomolecular and Biotechnological Sciences)

Indirizzo: Microbiologia molecolare e clinica

Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*

Direttore della scuola:
Prof. Masala Bruno

Relatore:
Prof. Pier Luigi Fiori

Correlatore:
Dott. Eric Viscogliosi

Tesi di Dottorato:
Dionigia Meloni

XXIV CICLO

Abstract

My thesis was conducted on the study of two groups of protozoa: the Parabasalia and *Blastocystis*. The first part of my work was focused on the identification, pathogenicity, and phylogeny of parabasalids. We showed that *Pentatrichomonas hominis* is a possible zoonotic species with a significant potential of transmission by the waterborne route and could be the aetiological agent of gastrointestinal troubles in children. We also confirmed the frequent presence of *Trichomonas tenax* and *Trichomonas vaginalis* outside their natural habitats in humans such as in the lungs and intestinal tract. In parallel, in complement to the SSU rRNA gene, we inferred the phylogenetic relationships among parabasalids on the basis of protein sequences (GAPDH, actin, EF1a). Our data indicated that increasing the number of taxa as well as the addition of new molecular markers greatly improved the robustness of the parabasalid tree and suggested that the ancestral parabasalid cell exhibited a morphologically simple structure. In the second part of my work, the first molecular epidemiological surveys of human *Blastocystis* isolates were conducted in Italy and Lebanon. The frequency of each *Blastocystis* subtype was determined in both populations. In addition, we first demonstrated mixed intra- and inter-subtype infections with *Blastocystis* in a same individual exhibiting a high potential risk of mixed infections with this parasite according to her lifestyle in rural area and long history of travelling.

Riassunto

La mia tesi si basa sullo studio di due gruppi di protozoi: Parabasalia e *Blastocystis*. La prima parte del mio lavoro si é focalizzata sull'identificazione, la patogenicità e filogenesi dei Parabsalia. Abbiamo dimostrato che *Pentatrichomonas hominis* è una possibile specie zoonotica con un significativo potenziale di trasmissione attraverso l'acqua e che potrebbe essere l'agente eziologico dei problemi gastrointestinali nei bambini. Abbiamo anche confermato che, nell'uomo, *Trichomonas tenax* e *Trichomonas vaginalis* sono presenti con elevata frequenza al di fuori dei loro habitat naturali, come nei polmoni e nel tratto intestinale. Inoltre, a complemento del gene SSU rRNA, abbiamo dedotto le relazioni filogenetiche tra i parabasalidi sulla base di altre sequenze proteiche (GAPDH, actina, EF1a). I nostri dati mostrano che l'aumento del numero dei taxa così come l'aggiunta di nuovi marcatori molecolari migliorano notevolmente la solidità dell' albero dei parabasalidi e ci suggeriscono che la cellula ancestrale dei Parabasalia mostrava una struttura morfologicamente semplice. Nella seconda parte del mio lavoro, sono stati condotti i primi studi epidemiologico-molecolari per individuare la presenza di *Blastocystis* in Italia e Libano. In entrambe le popolazioni é stata determanata la frequenza di ciascun sottotipo di *Blastocystis*. Inoltre, per la prima volta, abbiamo dimostrato la presenza di infezioni miste intra-e inter-sottotipo di *Blastocystis* in un medesimo individuo che presenta un elevato rischio di infezioni da parte di questo parassita, a causa del suo stile di vita, come il vivere in una zona rurale e il viaggiare frequantemente.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus _Blastocystis_
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Introduction

I Introduction

Protozoa (in Greek *proto* = first and *zoa* = animals) are heterotrophic, single-celled or colonial eukaryotes. There are over 30,000 different species of protozoa documented by scientists. Beyond this broad description, it is difficult to define protozoa because they are so diverse and only distantly related to each other. Despite the fact that protozoa is not a proper taxonomic name (this assemblage is not monophyletic in molecular phylogenies), it is a useful and functional term. Protozoa are ubiquitous and can be observed in oceans, lakes, rivers, glaciers, ponds, hot springs, and terrestrial environments. Their cysts can be found in even the most inhospitable parts of the biosphere (extreme heat or cold). Protozoa are important components of aquatic and soil ecosystems, and have a major ecological importance through their role in food chain and carbon cycling. For instance they help control the population of other organisms in maintaining ecological balance and diversity and are involved in the decomposition of the remains of dead plants and animals. Photosynthetic protozoa also contribute in the mitigation of global warming by recycling carbon dioxide. Most are free-living and eat bacteria, algae, or other protozoa but many species are in mutual relationship with other animals. Moreover, there are several protozoa of medical and economic importance including the flagellate *Trypanosoma,* which causes African sleeping sickness and *Plasmodium* species, which cause malaria.

Given their ecological interest as well as their yet unknown extending biodiversity, protozoa have long fascinated researchers. Moreover, regarding their economic and medical interest, the analysis of emerging protozoa and better identification of each individual potential pathogen has become one of the greatest challenges of the next decades in life sciences. Therefore projects in many research groups around the world are focused on these microorganisms. This is the case of the two laboratories in France and Italy jointly involved in my thesis project. Indeed, since several years, these two laboratories are working on the biology and diversity of two groups of protozoa: the Parabasalia and *Blastocystis* sp. Parabasalia or commonly parabasalids comprises a monophyletic but complex assemblage of diverse species of flagellated protozoa (more than 80 genera and 400 species have been described so far). Most parabasalids inhabit the digestive tract of animal hosts as commensals, parasites, or symbionts. In particular, symbiotic parabasalids found in the gut of termites and wood-eating cockroaches play a central role in the digestion of cellulose. This symbiotic relationship is considered a key element in the evolution of social behavior in the hosts. Several parabasalids are also of considerable medical and veterinary importance as pathogens i.e. *Trichomonas vaginalis* in humans and *Tritrichomonas foetus* in bovids.

Regarding *Blastocystis* sp., this is an enteric protozoa commonly occurring in a wide range of animals. To date, *Blastocystis* sp. is the most common intestinal parasite found in human feces and considered an emerging parasite with a worldwide distribution. It remained widely debated in the literature whether *Blastocystis* sp. is a truly pathogenic organism. However, recent *in vitro* and *in vivo* data together with those of the analysis of its genome, allowed proposing a model for pathogenesis of this parasite. Accumulating reports suggest an association between *Blastocystis* sp. and a variety of gastrointestinal disorders and skin lesions. Molecular phylogenies revealed a considerable genetic diversity among *Blastocystis* sp. isolates from humans and animals with the identification of numerous subtypes. However the correlation between subtypes and pathogenicity has not yet been clearly established.

Despite recent significant advances in the knowledge of these two groups of protozoa, several aspects of their biology, epidemiology, molecular diversity and pathogenicity remained to be clarified. For Parabasalia, questions arise concerning the pathogenicity of the gastrointestinal species *Pentatrichomonas hominis*, the frequency and nature of trichomonad infections in unusual locations (outside the gastrointestinal and urogenital tracts) or even the phylogeny and systematic of this complex group. Similarly for *Blastocystis* sp., its prevalence in the human population and its zoonotic potential raise questions about the impact of this parasite in public health and the significance of its genetic diversity on pathogenicity of variants. Consequently epidemiological studies in human populations are needed to report the frequency of different subtypes, permit tracking potential contamination sources and test a possible correlation between the subtype and the presence of intestinal and/or cutaneous symptoms. To provide answers to these questions my thesis project focused on the identification and pathogenicity of trichomonads in humans and the systematic and evolution of Parabasalia based on various molecular indicators. Concerning *Blastocystis* sp., I performed the first epidemiological studies in Italy and Lebanon and reported a first case of multiple co-infections with *Blastocystis* sp. isolates in a French patient.

Literature Review

II Parabasalia: a complex assemblage of diverse species of protozoa

II-1 Diversity and size

On the basis of marked morphological differences, mostly linked to the structure and development of the cytoskeleton, Parabasalia (or commonly parabasalids) have historically been divided into two classes: Trichomonada (or trichomonads) and Hypermastigia (or hypermastigids) (Brugerolle and Lee 2001). This group of protozoa includes more than 80 genera and 400 species identified so far (Grassé 1952, Honigberg 1963, Brugerolle 1976; Yamin 1979). Trichomonad cells are usually smaller and simpler than those of hypermastigids **(Figure 1)**. They have an average size between 10 and 40 µm, with up to six flagella. However there are some exceptions: in the genus *Dientamoeba*, basal bodies and flagella are absent and in the genera *Tricercomitus* and *Trichomitopsis*, cell size is of around 2 µm and 150 µm, respectively. Hypermastigids are easily distinguished from trichomonads by their size (until 500 µm), the hyper-development of their cytoskeleton and the multiplication of their flagella (until several thousands).

II-2 Cytoskeleton of Parabasalia

Almost all parabasalids are characterized by a complex and diversified cytoskeleton (for reviews see Honigberg 1963; Brugerolle 1976; Honigberg and Brugerolle 1990; Brugerolle and Lee 2001) **(Figure 2)** composed of:

a) Microtubular structures, such as the axostyle-pelta complex and the anterior and recurrent flagella. No free cytoplasmic microtubules have been found in these protozoa.

 b) Microfibrillar structures, such as the costa, the preaxostylar and parabasal fibers, and the undulating membrane. They all are attached to the kinetosomes and are composed of 2-4 nm filaments which are widely distributed in protozoa (Grain 1986).

Parabasal fibers are common to all species of Parabasalia, while the costa or the undulating membrane, are specific to a genus or family and have been commonly used as taxonomic characters in systematics.

II-2-1 Microtubular structures

Trichomonads usually have four flagella (ex: genus *Tritrichomonas*): three are anteriorly directed (anterior flagella) while the fourth, the recurrent flagellum, can be free or adherent to the cell body and conventionally determines the dorsal side of the cell. The kinetosomes (or basal bodies) of the anteriorly directed flagella are orthogonal to the
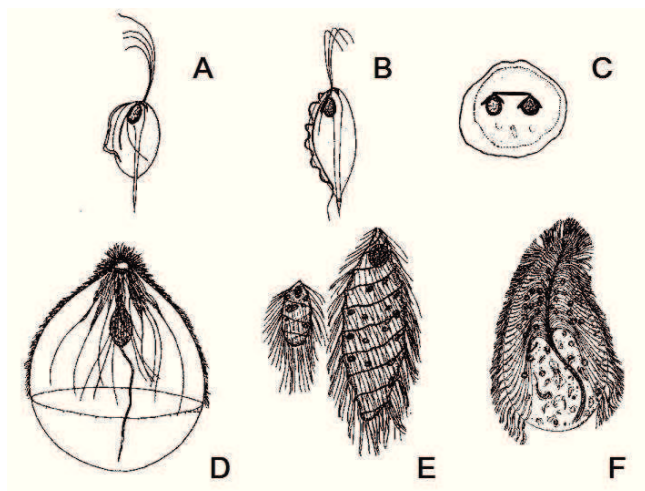
Figure 1: Morphological variability of Parabasalia. **Trichomonads: A:** *Trichomonas* sp.; **B:** *Tritrichomonas* sp.; **C:** *Dientamoeba fragilis.* **Hypermastigids: D:** *Deltotrichonympha* sp.; **E:** *Holomastigotes* sp.; **F:***Snyderella tabogae*.
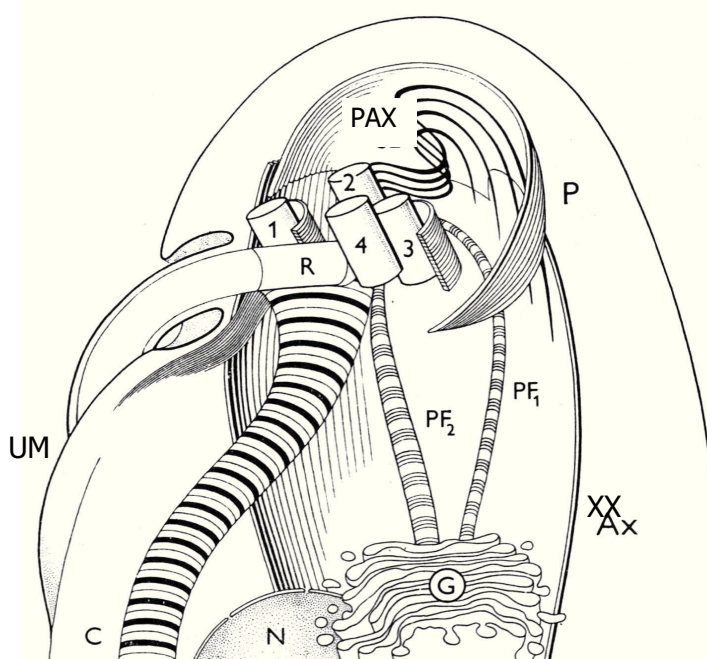
Figure 2: Schematic diagram of trichomonad mastigont (from Kulda et al. 1986).

<u>Abbreviations:</u> AX: axostyle; P: pelta (together AX and P form the axostyle-pelta complex); C: costa; G: dictyosome; N: nucleus; PAX: preaxostylar fibers; PF1, PF2: parabasal fibers; R: kinetosome of recurrent flagellum; UM: undulating membrane; 1, 2, 3, 4: kinetosomes of anterior flagella.
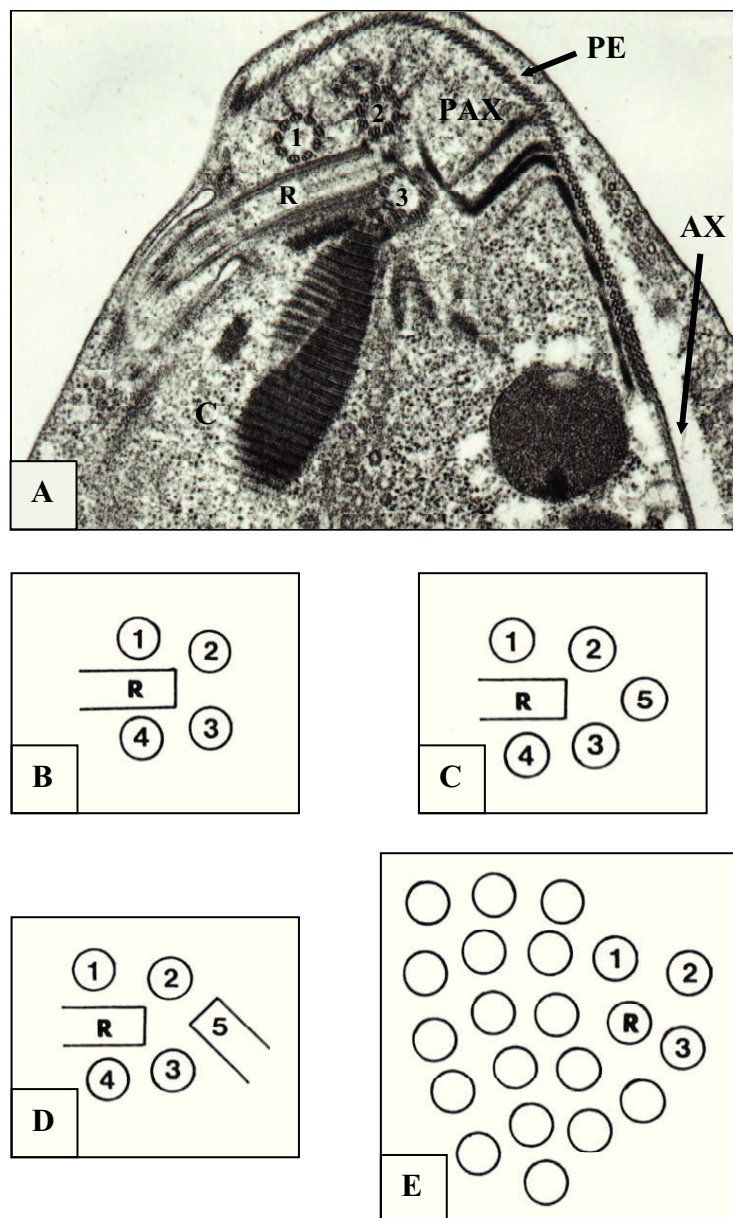
Figure 3: Basal body arrangement in Parabasalia. **A:** *Tritrichomonas foetus* (from Viscogliosi and Brugerolle, 1993b); **B:** *Trichomonas* and *Tetratrichomonas;* **C:** *Hexamastix* and *Pentatrichomonoides*; **D:** *Pentatrichomonas*; **E:** Flagellar area in the hypermastigid *Lophomonas.* **Abbreviations:** AX: axostyle; C: costa; PAX: preaxostylar fiber; PE: pelta; R, 1, 2, 3, 4, 5: kinetosome of flagella

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

kinetosome of the recurrent flagellum (four basic or privileged basal bodies) **(Figure 3)**. The kinetosomes are identifiable and numbered according to microfibrillar structures associated with them (1, 2, 3, and R). The genera *Trichomonas* and *Tetratrichomonas* possess an additional kinetosome (4) giving rise to a fourth anterior flagellum. A fifth anterior flagellum (5) is present in the genera *Hexamastix* and *Pentatrichomonoides* (Brugerolle et al. 1994). *Pentatrichomonas* also presents 6 flagella but the sixth kinetosome is orthogonal to kinetosomes of anterior flagella. Therefore, the sixth flagellum of this genus emerges on the ventral surface of the cell (Honigberg et al. 1968). Two particular and interesting genera: *Dientamoeba* which lacks kinetosomes and flagella (Camp et al. 1974), and *Histomonas* which has a single flagellum and 3 kinetosomes withouth flagella (Rybicki et al. 1972). In hypermastigids, there is an increase in the number of flagella (up to several thousands) forming the flagellar area (Hollande et Carruette-Valentin 1969, 1971, 1972; Kubai 1973).

The axostyle-pelta complex is a true axial skeleton that maintains the shape of the cell **(Figure 4)**. This complex is formed by the assembly of a large number of microtubules connected to each other and arranged in a parallel array running from the anterior to the posterior region of the cell. The axostyle is rolled into a cylindrical trunk which opens in a nuclear area to form a spoon-like capitulum. The axostylar trunk may protrude from the posterior cell surface. In *Devescovina*, the axostylar trunk can roll up on itself and describe several rounds of turns (Brugerolle 1976). It becomes even more powerful in hypermastigids (Hollande and Valentin 1969). On the contrary, this structure is much less developed or is lacked in amoeboid trichomonads such as *Histomonas* (Rybicka et al. 1972) or *Dientamoeba* (Camp et al. 1974).

II-2-2 <u>Microfibrillar structures</u>

Typical trichomonad cells exhibit different types of microfibrillar structures **(Figure 5)**. Kinetosome 2 is connected to the axostyle-pelta complex by several preaxostylar fibers (or sigmoid fibers), which are in varying numbers (4-18). These fibers course from the ventral surface of kinetosome 2 to the axostyle-pelta junction. Two parabasal fibers (striated fibers) are also attached to kinetosome 2 and support the dictyosomes without obvious connection with these organelles **(Figure 6)**.

The parabasal apparatus is composed of a dictyosome attached to a parabasal fiber. The periodicity of this striated fiber (about 42 nm) as well as its ultrastructure is similar to that of the costa (see below A-type costa). In hypermastigids, the periodicity of parabasal fibers is
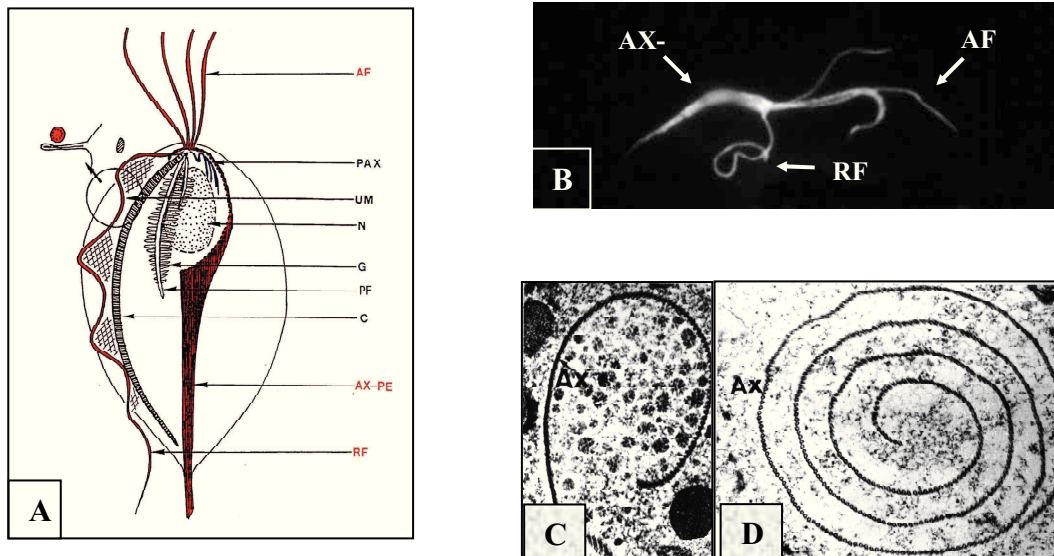
Figure 4: Microtubular structures in trichomonads. **A:** Schematic representation of cytoskeleton of *Trichomonas vaginalis*. Microtubular structures are indicated in red. **B:** Microtubular structures during interphase observed by immunofluorescence using an anti-tubulin antibody from *Trichomonas vaginalis* (from Noël et al. 2003). **C, D:** Transversal section of the axostyle from *Tritrichomonas* (**C**) and *Devescovina* (**D**) (from Brugerolle 1976). **Abbreviations:** AF: anterior flagella; AX: axostyle; AX-PE: axostyle-pelta complex; C: costa; G: dictyosome; N: nucleus; PAX: preaxostylar fibers, PF: parabasal fibers; RF: recurrent flagella; UM: undulating membrane.
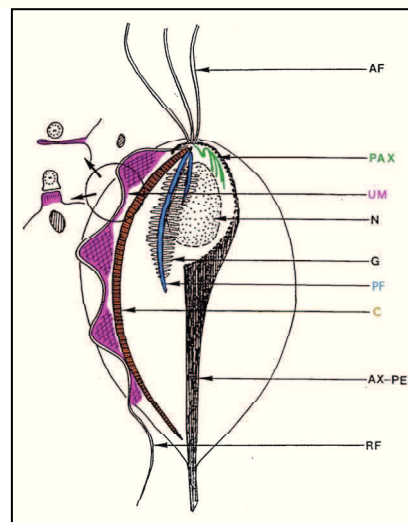
Figure 5: Schematic representation of cytoskeleton of trichomonads and location of microfibrillar structures. **Abbreviations:** AF: anterior flagella; AX: axostyle; AX-PE: axostyle-pelta complex; C: costa; G: dictyosome; N: nucleus; PAX: preaxostylar fibers, PF: parabasal fibers; RF: recurrent flagella; UM: undulating membrane.
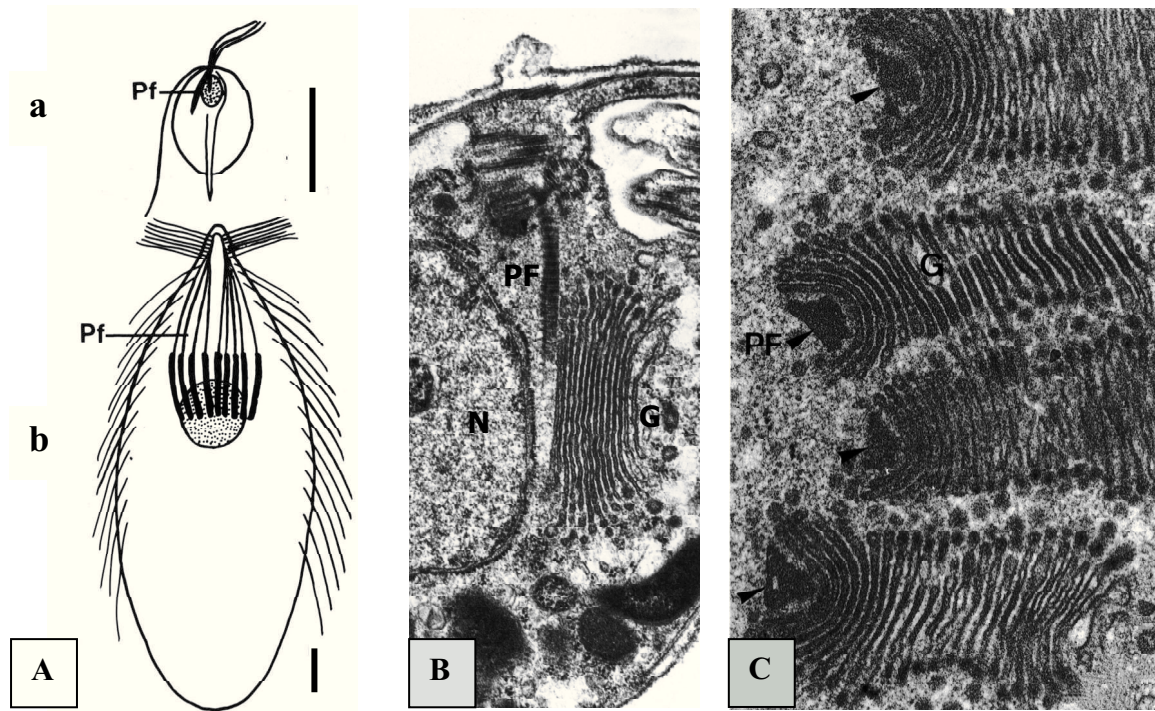
Figure 6: Parabasal fibers in Parabasalia (from Brugerolle and Viscogliosi 1994). **A:** Number and location of parabasal fibers in trichomonads (*Monocercomonas*) (a) and hypermastigids (*Trichonympha*) (b) (scale bar = 10 µm); **B:** In trichomonads (ex: *Monocercomonas*), the parabasal apparatus is composed of a parabasal fiber and a dictyosome; **C:** Multiplication of parabasal fibers in the hypermastigid *Joenia*. Each of the parabasal fibers is associated to a dictyosome.

<u>Abbreviations:</u> G: dictyosome; N: nucleus; PF: parabasal fibers.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

variable. Moreover, in these organisms, there is a multiplication of these fibers, each of them supporting a dictyosome (Hollande et Valentin 1969; Hollande et Carruette-Valentin 1971).

The undulating membrane is formed by the recurrent flagellum adhering to the cell body and a cytoplasmic fold including microfibrillar structures between the costa and the cytoplasmic membrane **(Figure 7)**. In typical trichomonads, the undulating membrane represents the principal mean of locomotion (Kulda et al. 1986). Three different types of undulating membranes have been described: 1) *Trichomonas*-type (lamellar-like form), 2) *Tritrichomonas*-type (rail-like form), and 3) *Devescovina*-type involving a microfibrillar structure, the cresta (Kirby 1941; Joyon et al. 1969; Brugerolle, 1976). The cresta is also present in some hypermastigids. This morphological criterion is the basis for the creation of the order Cristamonadida by Brugerolle and Patterson (2001).

Finally, the costa, a broad striated fiber connected to the kinetosomes, has been described as consisting of bundles of 2-4 nm filaments **(Figure 8)**. The costa occurs only in trichomonads which possess an undulating membrane. Indeed, the costa is assumed to provide mechanical support to the undulating membrane (Kulda et al. 1986). Two major costa types can be distinguished based on their insertion site on kinetosomes and band pattern although each type exhibits the same periodicity (about 42 nm). The first type or A-type (*Tritrichomonas*) is attached to kinetosome 2 while B-type (*Trichomonas*) is connected to kinetosome R. In the A-type costae, minor striations are visible in addition to the two major bands also seen in B-type costae. Moreover, A-type costae show no longitudinal filaments composing the lattice structure as in the B-type costae of *Pentatrichomonas* (Honigberg et al. 1968) and *Pentatrichomonoides* (Brugerolle et al. 1994). In fact, this dichotomy is not so simple. For instance, the costa of the genus *Trichomitopsis* is of A-type according to its striation but of B-type regarding its insertion site on kinetosomes (Hollande et Valentin 1968; Mattern and Honigberg 1971). Moreover, the costa can be motile in the genera *Trichomitopsis* and *Pseudotrypanosoma* (Amos et al. 1979).

II-3 <u>Division and hydrogenosomes</u>

All parabasalid genera studied to date exhibit a special type of closed mitosis called cryptopleuromitosis by Grassé (1952), characterized by the persistence of the nuclear envelope nuclear and the presence of an extra-nuclear spindle or paradesmosis (Brugerolle 1975) (**Figure 9**). This division process has been well described from light (Grasse 1952;

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
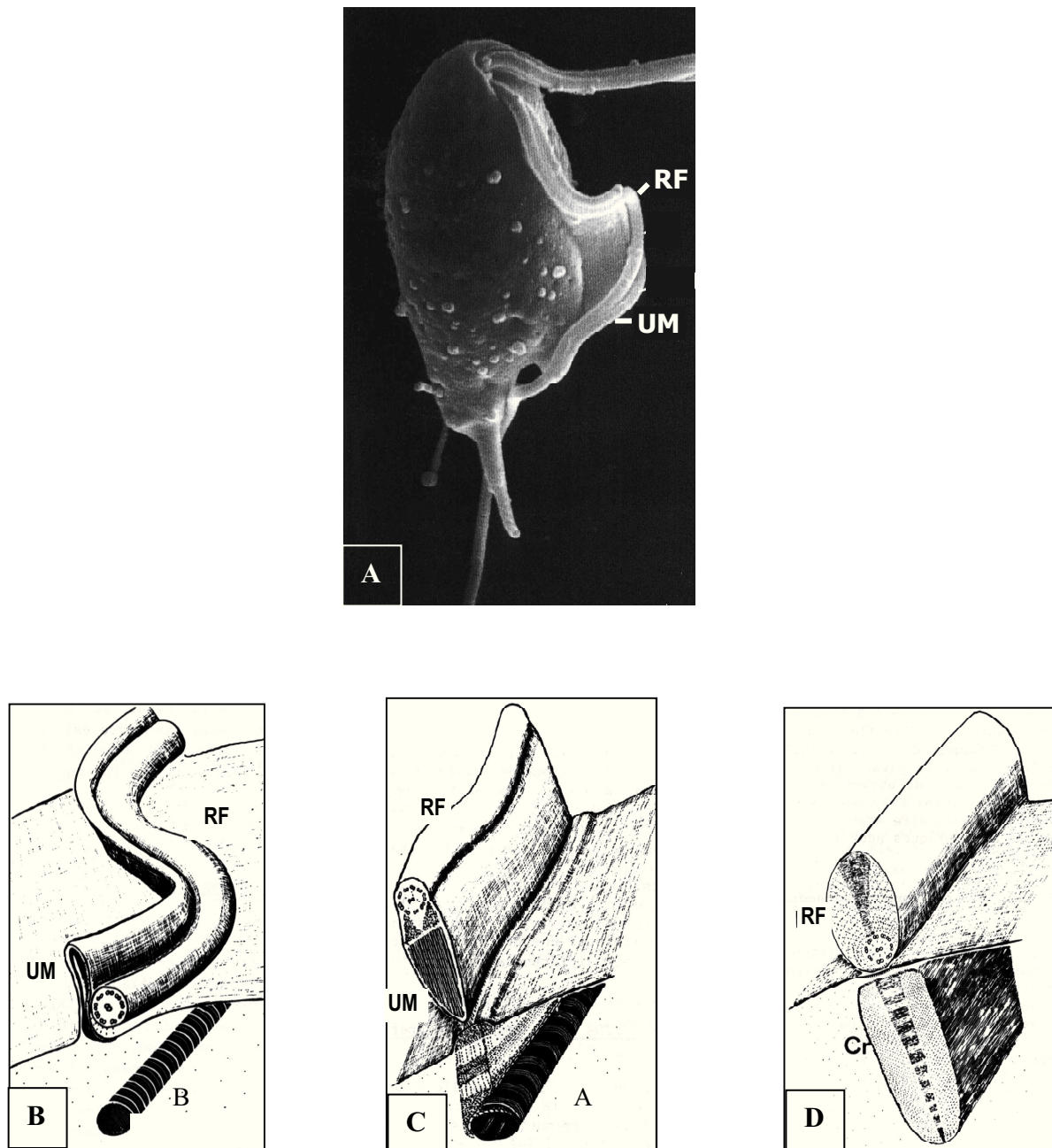
Figure 7: Different types of undulating membrane in trichomonads. **A:** Undulating membrane of *Trichomitus batrachorum* observed by scanning microscopy (Germot et al. 1996); **B:** *Trichomonas*-type undulating membrane; **C:** *Tritrichomonas*-type undulating membrane; **D:** *Devescovina*-type undulating membrane; **B, C, D**: from Brugerolle 1976. **Abbreviations:** B: costa of *Trichomonas*; A: costa of *Tritrichomonas*; Cr: cresta; RF: recurrent flagella; UM: undulating membrane.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
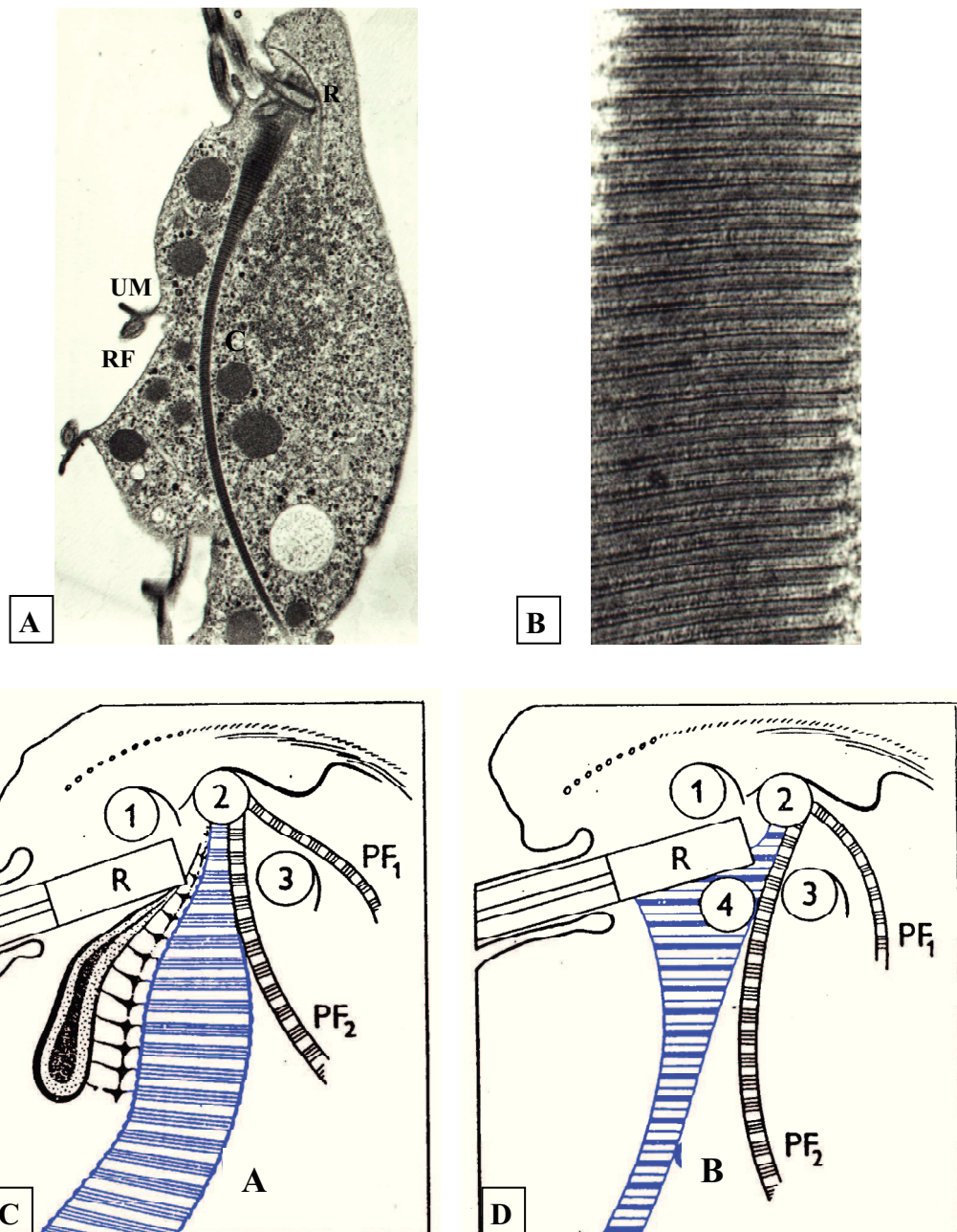
Figure 8: Costa of trichomonads. **A:** Section of *Trichomonas vaginalis* observed by transmission electron microscopy showing the costa attached to the basal bodies of the flagella and situated under the undulating membrane adhering to the recurrent flagellum (from Viscogliosi and Brugerolle 1994); **B:** Longitudinal section of the A-type costa of *Tritrichomonas foetus* (from Viscogliosi and Brugerolle 1994); **C, D:** Schematic representation of A-type (**C**) and B-type costae (**D**) (from Kulda et al. 1986).

Abbreviations: PF1 and PF2: parabasal fibers; RF: recurrent flagella; R, 1, 2, 3 and 4: kinetosome.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Cleveland 1961) and electron microscopy (Hollande 1974; Brugerolle 1975, 1976; Honigberg and Brugerolle 1990), and by immunofluorescence using anti-tubulin antibodies (Delgado-Viscogliosi et al. 1996; Noël et al. 2003).

Parabasalia are also typified by the absence of mitochondria and peroxisomes but possess spherical or rod-shaped organelles, measuring about 1 µm, and named hydrogenosomes (Lindmark and Muller 1973) **(Figure 10)** corresponding to anaerobic energy-generating organelles and involving in metabolic processes that extend glycolysis (Kulda 1999). The hydrogenosomes are constituted of a dense matrix surrounded by a double membrane like mitochondria (Benchimol and De Souza 1983) and divide by segmentation and partition (Benchimol 1996, 2009). In parabasalids, DNA was not detected in this organelle (Turner and Müller 1983; Clemens and Johnson 2000). The question of the origin of hydrogenosomes is still a matter of debate. However, all the recent data strongly support a common origin of hydrogenosomes and mitochondria (see Carlton et al. 2007; Mentel 2008; Hjort et al. 2010; Lithgow and Schneider 2010; Shiflett and Johnson 2010).

II-4 <u>Habitat, pathogenicity and treatment</u>

Globally the distinction of parabasalids into two classes on the basis of morphological characters also corresponds to a clear separation in terms of habitat. Hypermastigids are found exclusively in the hindgut of termites and wood-eating cockroaches of the genus *Cryptocercus* (Yamin 1979) and play a pivotal role in digestion of wood in their mutual symbiotic relationship with the host insect. The microbial community in the gut of termites is composed of both bacteria and protozoa. The supposed role of hypermastigids (and symbiotic trichomonads found in the same host) is to digest cellulose ingested by these insects (Breznak 1982; Grosovsky et Margulis 1982; Ohkuma 2003; Ohkuma 2008). In the gut of lower termites, the presence of cellulolytic protozoa is crucial for highly efficient degradation of cellulose through their cellulases.

Most of trichomonads are symbionts or parasites of a wide range of animals including insects, mollusks, fishes, amphibians, reptiles, birds and mammals, and colonize the gastrointestinal or urogenital tract of their respective hosts. Four trichomonad species are currently found in humans (Honigberg 1990): *Trichomonas vaginalis* in the genitourinary tract, *Trichomonas tenax* in the oral cavity, and *Pentatrichomonas hominis* and *Dientamoeba fragilis* in the intestinal tract. Only *Dientamoeba fragilis* and *Trichomonas vaginalis* are considered pathogenic. *Dientamoeba fragilis* is the causative agent of a common form of

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
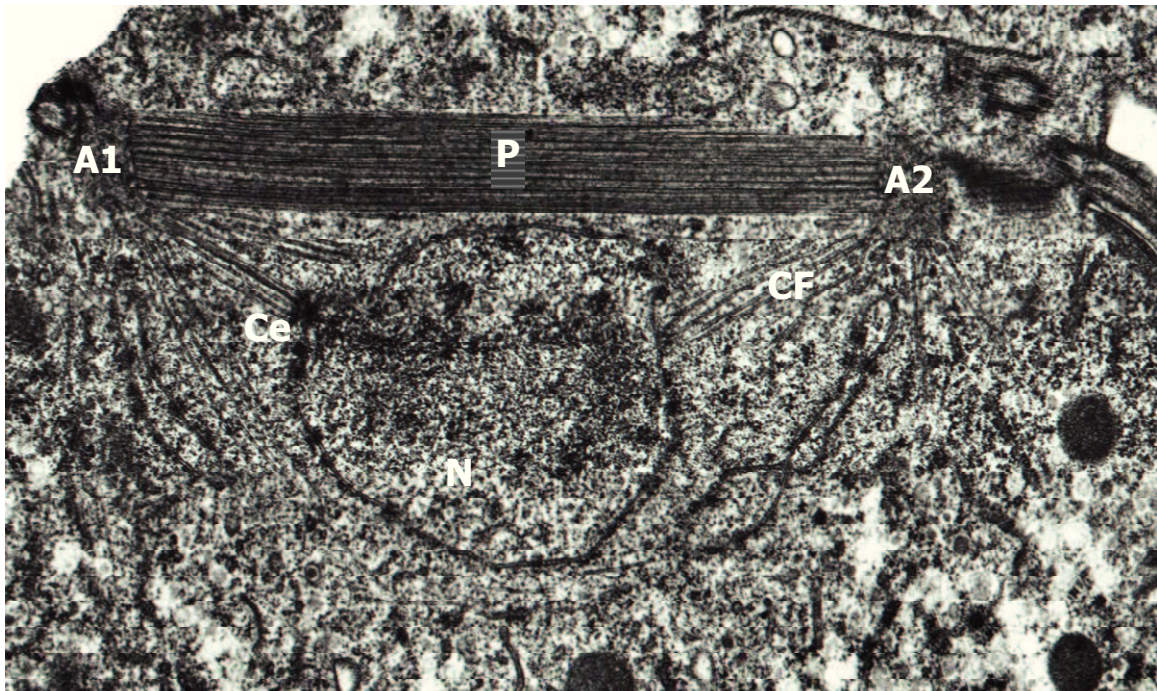
Figure 9: Cryptopleuromitosis in *Trichomonas vaginalis* (from Brugerolle 1975). **Abbreviations:** A1, A2: atractophores; Ce: centromeres; CF: chromosomal fibers; N: nucleus; P: paradesmosis.
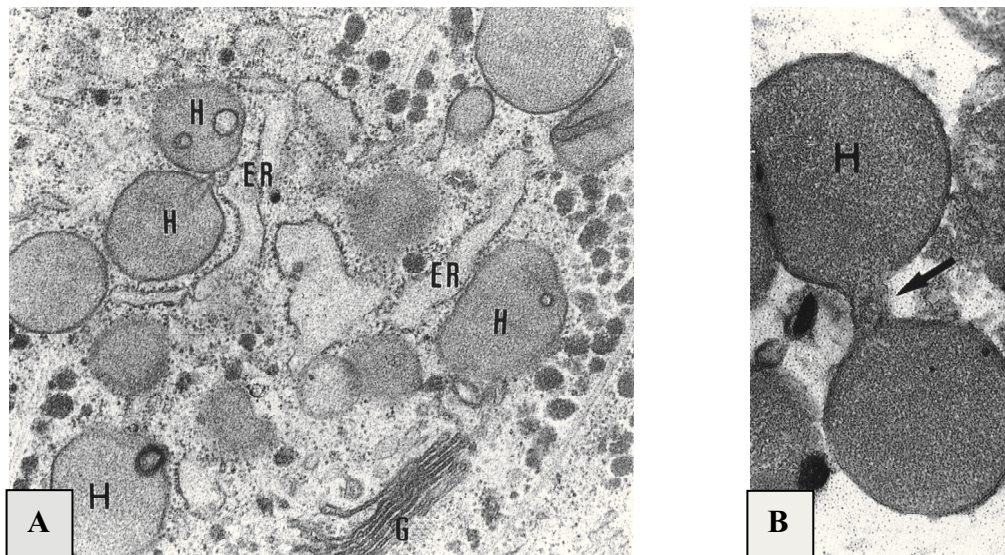
Figure 10: Hydrogenosomes observed by transmission electron microscopy (from Benchimol et al. 1996). **A**: in *Tritrichomonas foetus*; **B**: in *Trichomonas vaginalis* (division of organelles by segmentation).

Abbreviations: ER : endoplasmic reticulum; G: dictyosome; H : hydrogenosomes.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

chronic diarrhea (Stark et al. 2006) while *Trichomonas vaginalis* **(Figure 11)** is the causative agent of human trichomonosis, which has been the most prevalent nonviral sexually transmitted disease worldwide in recent years (for reviews see Gerbase et al. 1998; Petrin et al. 1998; Lehka and Alderete 2000, Soper 2004; Miller et al. 2005; Van der Pol 2007; Krashin et al. 2010). According to the World Health Organization (WHO), the annual incidence of trichomonosis is more than 200 million cases worldwide, making it a major public health problem (Global Prevalence and Incidence of Selected Curable Sexually Transmitted Infections, World Health Organization, Geneva, 2001).

Although *Trichomonas vaginalis* infection is primarily considered as a women disease, it also occurs in men. Trichomonosis in men is largely asymptomatic, which naturally favors the transmission of the parasite. However complications associated with trichomonosis include prostatitis, urethritis, and infertility (Gardner et al. 1986; Krieger et al. 1995; Schwebke et al. 2003). In women, this parasite is the causative agent of vaginitis, urethritis and cystitis. Women who are infected during pregnancy are predisposed to premature rupture of the placental membranes, premature labor, and low-birth-weight infants (Cotch et al. 1997; Sutton 1999; Smith et al. 2002). This parasite is also the potential vector of human viruses such as papilloma viruses known to be the cause of cervical cancers (Pindak et al. 1989; Zhang and Begg 1994; Viikki et al. 2000) and pathogenic bacteria (*Neisseria, Mycoplasma, Ureaplasma urealyticum*) attached to its surface (Rappelli et al. 1998; Dessi et al. 2005, 2006; Diaz et al. 2010). As with other sexually transmitted diseases, *Trichomonas vaginalis* infection can increase the predisposition of individuals to human immunodeficiency virus (HIV) (Mason et al. 2005; McClelland et al. 2007; Van Der Pol et al. 2007).

Many studies have been performed on parasitic trichomonads from animals that have a significant economic impact. This is the case of *Tritrichomonas foetus* **(Figure 11)** which colonizes the genital tract of bovids (Yule et al. 1989; Corbeil 1994) and may be the cause of endometritis, abortions, and temporary or permanent infertility. *Histomonas meleagridis* is the agent of histomonosis of poultries or "blackhead disease", for which the intestinal and liver damages can be fatal in a few weeks (Huber et al. 2006). Almost all of trichomonads are commensals, but few free-living species have been isolated from lake sediments such as *Ditrichomonas honigbergii* (Farmer 1993), *Pseudotrichomonas keilini* (Bishop 1939) and *Honigbergiella* (Hampl et al. 2007), and marine sediments, such as *Monotrichomonas carabina* and *Monotrichomonas* sp. (Edgcomb et al. 1998).
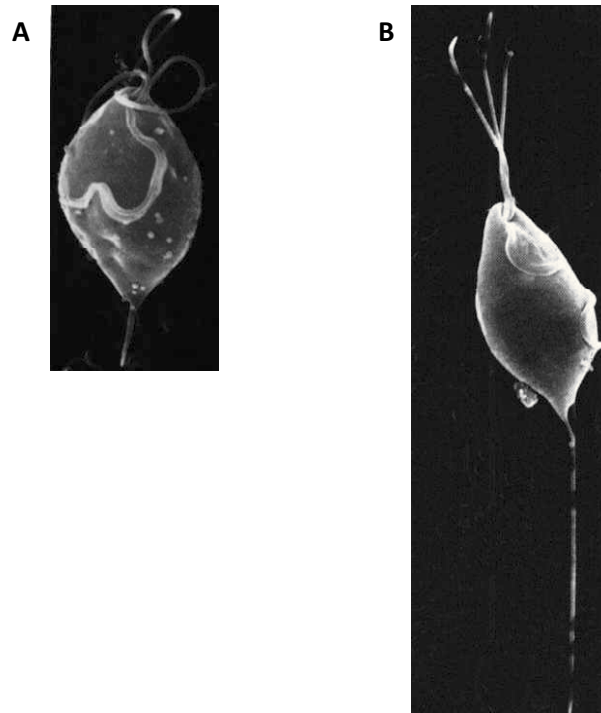
Figure 11: Pathogenic trichomonads observed by scanning electron microscopy (from Warton and Honigberg 1979).

A: *Trichomonas vaginalis*, a human parasite; **B:** *Tritrichomonas foetus*, a parasite of bovids.
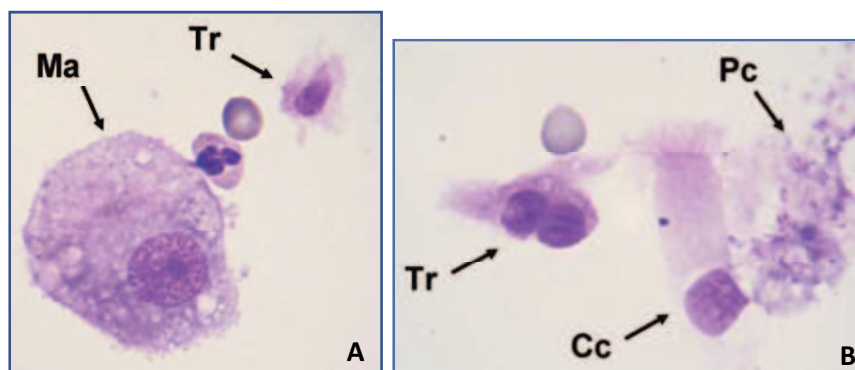
Figure 12: Cytological appearance of trichomonad cells in a bronchoalveolar lavage sample (May-Grünwald-Giemsa staining) (from Duboucher et al. 2006). **A:** A trichomonad cell in the vicinity of a macrophage cell; **B:** An amoeboid trichomonad exhibiting two nuclei in the vicinity of a bronchial ciliated cell and an aggregrate of *Pneumocystis* organisms. **Abbreviations:** Tr: trichomonad cell; Ma: macrophage cell; Cc: bronchial ciliated cell; Pc: *Pneumocystis* organisms.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Diseases caused by trichomonads can be cured by 5-nitroimidazole drugs, such as metronidazole (commercially known as Flagyl) (Upcroft and Upcroft 2001; Cudmore et al. 2004), effective against a variety of anaerobic protozoa and bacteria. Metronidazole enters the trichomonad cell and the hydrogenosome by simple diffusion. In the organelle, it acts as a preferential electron acceptor, competing efficiently for electrons with a hydrogenosomal key enzyme (Kulda 1999). The reduction of the metronidazole results in production of toxic intermediates that break the DNA strands and causes death of the parasite within few hours. Although the cure rate is excellent, treatment fealure is problematic. Indeed, the many published reports of clinically resistant trichomonosis undoubtedly show that resistance of *Trichomonas vaginalis* to metronidazole is on the rise (Cudmore et al. 2004; Goldman et al. 2009; Upcroft et al. 2009). Clearly new antitrichomonal agents are needed to treat resistant organisms. In this context, El-Sherbini et al. (2009) have obtained promising results with an herbal treatment of resistant strains.

II-5 <u>Unusual location of trichomonads and pulmonary trichomonosis</u>

It was thought that each trichomonad species found in humans had a specific tropism for its site of infection. However it has been recently shown that these microorganisms could be found outside their natural habitats, such as in the lungs (pulmonary trichomonosis). For instance, *Trichomonas tenax* has been identified in numerous cases in the upper or lower respiratory tract of humans (Lewis et al. 2003; Mahmoud and Rahman 2004; Mallat et al. 2004; Bellanger et al. 2008). This organism is usually regarded as a harmless commensal of the human mouth associated with poor dentition and oral hygiene and is thought to enter the respiratory tract by aspiration of oropharyngeal secretions. Interestingly, other trichomonad species have been identified in human lungs using immunological and molecular tools. These species include *Trichomonas vaginalis* (Duboucher et al. 2003), *Pentatrichomonas hominis* (Jongwutiwes et al. 2000), *Tritrichomonas foetus* (Duboucher et al. 2006), a genital trichomonad found in bovids, and avian species belonging to the genus *Tetratrichomonas* (Kutisova et al. 2005; Mantini et al. 2009). The identification of strains belonging to the genera *Tritrichomonas* and *Tetratrichomonas* in humans clearly raised the question of the as yet poorly understood zoonotic potential of trichomonads, since these taxa were thought to be of animal origin.

Regarding the recent literature, the presence of trichomonads in the human respiratory tract cannot be hereafter considered unusual. Indeed these microorganisms are found frequently in the course of human *Pneumocystis* pneumonia (PCP) (Duboucher et al. 2005,

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

2007a) and of acute respiratory distress syndrome (ARDS) (Duboucher et al. 2007b). More precisely, trichomonads were detected as coinfecting agents with *Pneumocystis* in 60% of PCP patients and were found in 30% of ARDS patients. Because these protozoa are microaerophilic, they are likely not able to cause pulmonary disease by themselves and require favorable conditions for their development. With regard to PCP and ARDS, it was hypothesized that the development of these protozoa in the alveolar lumens was linked to local hypoxic conditions rather than immunodepression. Indeed, in these two clinical contexts, the alveolar lumens are obliterated by fungi in the case of PCP or by fibrin and cellular debris in the case of ARDS. *Trichomonas vaginalis* was also identified in the lungs during neonatal infections suggesting a maternal infection during birth (Szarka et al. 2002). The occurrence of trichomonads in the respiratory tract of humans is likely overlooked, due mainly to the amoeboid transformation of these microorganisms **(Figure 12)**. Such pleomorphism renders trichomonad identification on cytologic slides difficult as long as the conception of trichomonad morphology is restricted to that of flagellated stage. In addition to the lung, such trichomonads have been identified in other unusual locations. *Trichomonas vaginalis* was also found in the prostate (Gardner et al. 1986), in the uterus (Gardner et al. 1987) in a renal cyst (O'Hara et al. 1980) in the urine (Bellanger et al. 2008), and in a case of peritonitis (Straube 1991) while *Pentatrichomonas hominis* has been reported in a liver abscess (Jakobsen et al. 1987).

II-6 Genome of *Trichomonas vaginalis*

To date, *Trichomonas vaginalis* is the only parabasalid species whose genome was sequenced (Carlton et al. 2007). Its genome is haploid and composed of six monocentric chromosomes. The size of this genome is estimated to be approximately 160 Mb. A core set of around 60,000 protein-coding genes was identified endowing this parasite with one of the highest coding capacities among eukaryotes (for comparison, 35,845 genes for *Homo sapiens* and 9,649 for *Giardia intestinalis*). Only 65 genes present introns and 20 amino acids present Transfer RNAs (tRNAs). In addition, around 250 ribosomal DNA (rDNA) units were identified and localized into one of the six chromosomes. Many gene families in the *Trichomonas vaginalis* genome have undergone expansion on a scale unprecedented in unicellular eukaryotes. At least 65% of the genome is repetitive reflecting a recent massive expansion of genetic material. This expansion is in conjunction with the shaping of metabolic pathways that likely transpired through lateral gene transfer from bacteria. Probably the ancestor of *Trichomonas tenax* and *Trichomonas vaginalis* was a parasite of the digestive

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

tract. After speciation *Trichomonas tenax* has colonized the oral cavity and *Trichomonas vaginalis* the urogenital tract. The settlement required an adaptation of *Trichomonas vaginalis* to its environment and was therefore accompanied by an expansion of several gene families especially those involved in endocytosis, phagocytosis and membrane trafficking and related to lifestyle (Vogel et al. 2006). This expansion has been accompanied by an increase in the size of the parasite facilitating phagocytosis of bacteria and increasing its area of contact with the vaginal mucosa while limiting its own phagocytosis by host cells. The 59 families of repeated sequences occupy the most represented 39 Mb of the genome and are grouped into several categories including virus-like, transposons-like and retrotransposons-like. About 150 cases of possible lateral gene transfer from prokaryotic origin have been identified and involve several pathways. In parallel, genes involved in the mechanisms of RNA interference and meiosis have been identified.

III <u>Emergence of Parabasalia in eukaryogenesis</u>

Regarding the emergence of parabasalids in eukaryogenesis, Sogin (1989, 1991) was the first to propose a comprehensive tree of life, including the SSU rRNA gene sequence of a representative of this group of protozoa, *Trichomonas vaginalis*. In the phylogenetic trees, diplomonads, such as *Giardia*, represented the earliest diverging lineage followed by the microsporidia (*Vairimorpha*) and parabasalids. Leipe et al. (1993) showed that the order of emergence of these three groups of protozoa was unsettled and influenced by outlying prokaryotic taxa with different G+C compositions in their rRNA coding regions. Additional phylogenies based on this molecular indicator (Gunderson et al. 1995; Sogin et al. 1996; Sogin 1997; Sogin and Silberman 1998) confirmed the early divergence of Parabasalia. However it was suggested that the early emergence of this group could be an artifact linked to the phenomenon called "long-branch attraction" or LBA (Philippe and Laurent 1998; Philippe and Germot 2000). In simple terms, molecules such as rRNA of some "primitive" protists such as parabasalids evolve rapidly. Consequently, in the phylogenetic reconstructions, these organisms are "attracted" by bacterial outgroups and are located at the base of the eukaryotic tree although this is not their true phylogenetic position. Molecular phylogeny of eukaryotes was also inferred from the comparison of numerous protein sequences such as such as elongation factor EF-1α (Roger et al. 1999), initiation facteur of transcription eIF-2γ (Keeling et al. 1998), terminator facteur of transcription eRF1 and eRF2 (Moreira et al. 2002), ubiquitine (Keeling and Doolittle 1995), chaperonines CCT alpha and CCT delta (Archibald et al. 2000, 2002) and enolase (Keeling and Palmer 2000). Most of

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

these analyses reinforced the hypothesis of the primitive emergence of Parabasalia. For Bapteste and Philippe (2002), the basal position of Parabasalia in the phylogenies inferred from protein sequences was also likely due to LBA phenomenon or to the bacterial origin of some genes in these protozoa acquiring through horizontal gene transfer. In parallel, Philippe et al. (2000) developed the hypothesis of the "eukaryotic big bang" that suggested a rapid diversification of all eukaryotic lineages. In 2002, Cavalier-Smith proposed a new classification of protozoa and included Parabasalia in the subphylum Excavata. This grouping including around 10 groups of protozoa (Preaxostyla, Euglenozoa, Heterolobosea…) was based on common morphological traits (organization of the flagellar apparatus and presence in most of the taxa of a "feeding groove" or ventral mouth to capture prey) and molecular phylogenies. Therefore, a majority of eukaryotes were included into six groups (Simpson and Roger 2004): Opisthokonta, Amoebozoa, Plantae, Chromalveolata, Rhizaria, and Excavata **(Figure 13).**

The monophyly of Excavata was initially confirmed by Hampl et al. (2005) by reconstructing phylogenies based on concatenated sequences of eight different markers. These authors also identified a group called Metamonada that included Parabasalia, diplomonads such as *Giardia* and oxymonads. Simpson et al. (2006) confirmed that Parabasalia represented the sister group of diplomonads and of the free-living amitochondriate protozoa *Carpediomonas*. Subsequently, Keeling et al. (2005) and Keeling (2007) proposed a tree of eukaryotes with five "supergroups" instead of 6 with Parabasalia within Excavata. Recently, Hampl et al. (2009) maintained the classification with six supergroups on the basis of phylogenomic analyses and were able to determine their relationships **(Figure 14).** In this latter study, the monophyly of Excavata was reconfirmed as well as the position of Parabasalia as sister group of diplomonads. In the tree, the two groups Opisthokonta and Amoebozoa were included in the Unikonts, the other 4 groups forming Bikonts. Although the position of the root of eukaryotes is still controversial, it is probably somewhere between Unikonts and Bikonts. As a result, this implies that Excavata is the most primitive group within Bikonts.

IV Systematic and evolution of Parabasalia
IV-1 Based on analysis of morphological characters

Taxonomic classifications of parabasalids have been first proposed on the basis of
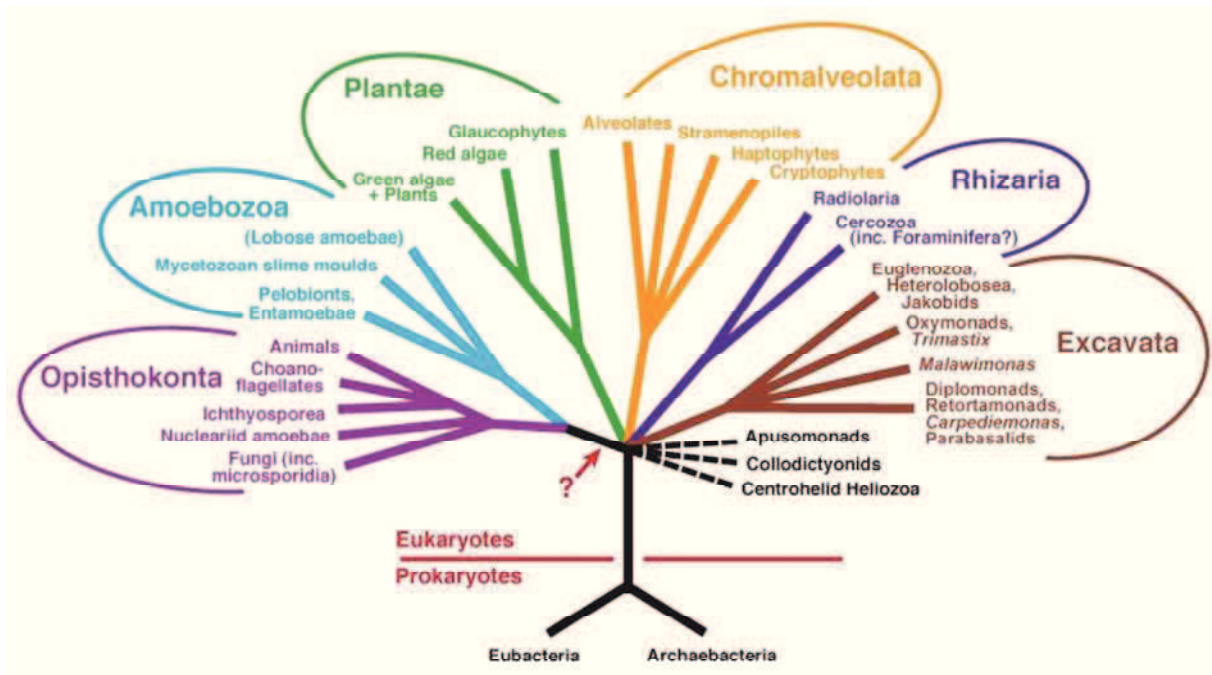
Figure 13: Relationships among the six supergroups of Eukaryotes (from Simpson and Roger 2004).
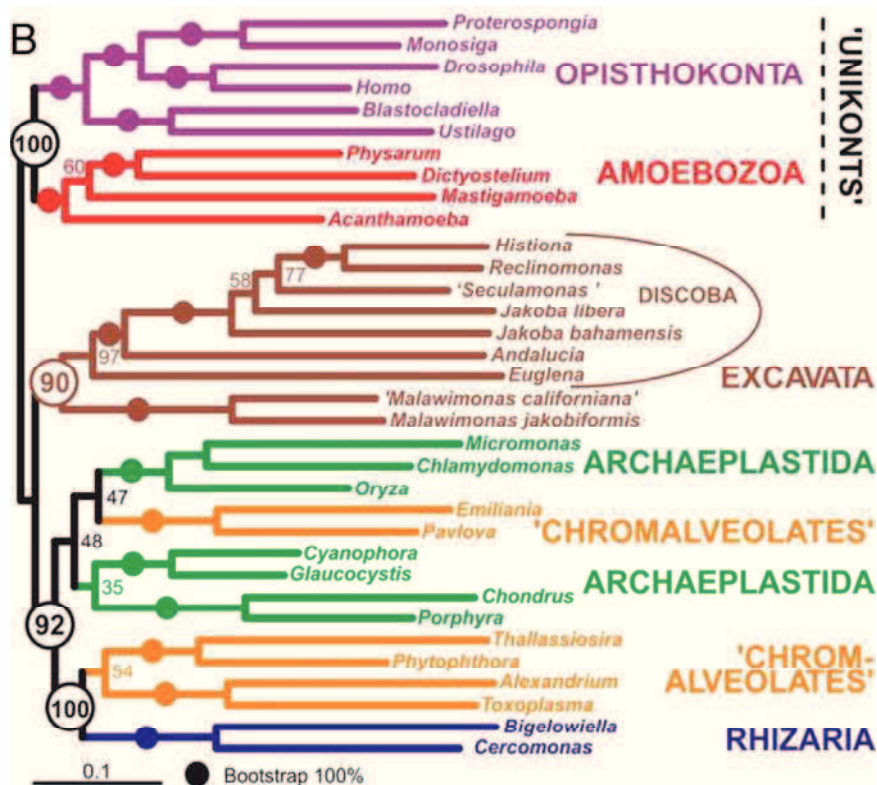
Figure 14: Phylogenomic analysis of the 6 supergroups of Eukaryotes (from Hampl et al. 2009)

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

marked morphological differences (Honigberg 1963; Brugerolle 1976, 1986; Yamin 1979, Levine et al. 1980, Brugerolle and Lee 2001). These evolutionary schemes were primarily based on a restricted number of characters mostly linked to the structure and development of cytoskeleton. According to these studies, Parabasalia were traditionally divided into two classes (Brugerolle and Lee 2001) **(Figure 15)**:

1) Trichomonada including a single order, Trichomonadida and four main families: Trichomonadidae, Monocercomonadidae, Devescovinidae, and Calonymphidae, each consisting of several subfamilies such as Trichomonadinae (*Trichomonas*) and Tritrichomonadinae (*Tritrichomonas*)

2) Hypermastigea divided into three orders: Lophomonadida, Trichonymphida and Spirotrichonymphida each consisting, respectively, of six (Joeniidae, Lophomonadidae, Kofoidiidae, Rhizonymphidae, Microjoenidae and Deltotrichonymphidae), six (Hoplonymphidae, Staurojoenidae, Trichonymphidae, Eucomonymphidae, Teranymphidae and Spirotrichosomidae), and three (Spirotrichonymphidae, Holomastigotoididae and Holomastigotidae) families.

The presumed evolution of parabasalids reflected the traditional view regarding polarization of cytoskeletal development from simple to complex. A progressive increase in the complexity of the cytoskeleton was observed within the trichomonads from the Monocercomonadidae (exhibiting a typical mastigont that comprises a parabasal apparatus, the axial axostyle-pelta complex, three anterior flafella, and one recurrent flagellum) to the Trichomonadidae (development of costa and undulating membrane), then the Devescovinidae/Calonymphidae (development of the cresta and polymastigonte state), and finally the hypermastigids (hyperdevelopment of their cytoskeleton and multiplication of flagella).

IV-2 Based on the analysis of SSU rRNA gene sequences

As an alternative to traditional phenotypic markers, phylogenetic relationships can be established through comparison of SSU rRNA sequences. Many trichomonad members have been cultivated and thus well-studied phylogenetically (Berchtold and König 1995; Berchtold et al. 1995; Gunderson et al. 1995; Fukura et al. 1996; Silberman et al. 1996; Edgcomb et al. 1998; Keeling et al. 1998; Delgado-Viscogliosi et al. 2000; Gerbod et al. 2000, 2001; Tachezy et al. 2002; Hampl et al. 2004, 2006; Cepicka et al. 2006; Dufernez et al. 2007; Mantini et al. 2009). In contrast, almost all the gut dwellers are resistant to cultivation,

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

HYPERMASTIGEA

Joenia

DEVESCO
VINIDAE

Devescovina

CALONYM
PHIDAE

Calonympha

Tritrichomonadinae

Trichomonadinae

TRICHOMO
NADIDAE

Tritrichomonas

Trichomonas

MONOCERCO
MONADIDAE

Monocercomonas

T
R
I
C
H
O
M
O
N
A
D
E
A

Figure 15: Systematic of Parabasalia based on morphological characters (modified from Dyer 1989).

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

presenting a serious obstacle for taxonomic studies of these microorganisms using molecular means. However SSU rRNA gene sequences from parabasalid symbionts were obtained by PCR amplification of whole gut fauna (Berchtold and König 1995; Gunderson et al. 1995; Keeling et al. 1998; Ohkuma et al. 1998, 2000; Gerbod et al. 2000). This has led to the accumulation of a large number of as yet unidentified sequences, but without organismal identification, little can be conducted concerning their taxonomic position and evolution. In some studies, the corresponding organisms from which the SSU rRNA gene sequences were derived have been identified by whole-cell *in situ* hybridizations with sequence-specific probes (Berchtold and König 1995; Ohkuma et al. 1998, 2000; Gerbod et al. 2002; Noël et al. 2007). Alternatively, sequence data were also obtained by amplifying gene sequences from a small population of some uncultivated symbionts physically isolated under microscopy (Dacks and Redfield 1998; Keeling et al. 1998; Fröhlich and König 1999; Gerbod et al. 2002; Keeling 2002; Ohkuma et al. 2005; Carpenter and Keeling 2007; Carpenter et al. 2009; Noda et al. 2009). The SSU rRNA analyses provided new and sometimes unexpected insights into the evolution of this group of protozoa and globally were incongruent with the morphology-based classification of Parabasalia. Briefly, all the trichomonad families as well as the hypermastigids did not form monophyletic groups **(Figure 16)**. In addition, Monocercomonadidae did not form basal branches in the parabasalid tree suggesting that their simplicity is not a primitive state. In contrast, hypermastigids exhibiting a hyperdeveloped cytoskeleton and classically considered to be late-evolving lineages represented the deepest branches of Parabasalia. In sum, the molecular data indicated that the cytoskeletal complexity of parabasalids, while useful as a taxonomic trait, was not a reliable phylogenetic indicator, and that their traditional polarization, from "simple" to "complex" should be turned almost upside-down.

According to these data, many researchers point out the need to revise parabasalid systematic. Brugerolle and Patterson (2001) were the first to propose a new classification of parabasalids at the ordinal level and divided this group into three orders: Trichonymphida, Trichomonadida, and the Cristamonadida. The newly created order Cristamonadida included the trichomonad families Devescovinidae and Calonymphidae, as well the hypermastigid order Lophomonadida on the basis of sharing similarities at the ultrastructural level. The same authors also united the remaining hypermastigid orders Trichonymphida and Spirotrichonymphida into a unique order Trichonymphida, on the sole basis of morphological characters. Yet, this latter grouping was incongruent with additional molecular data reported

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
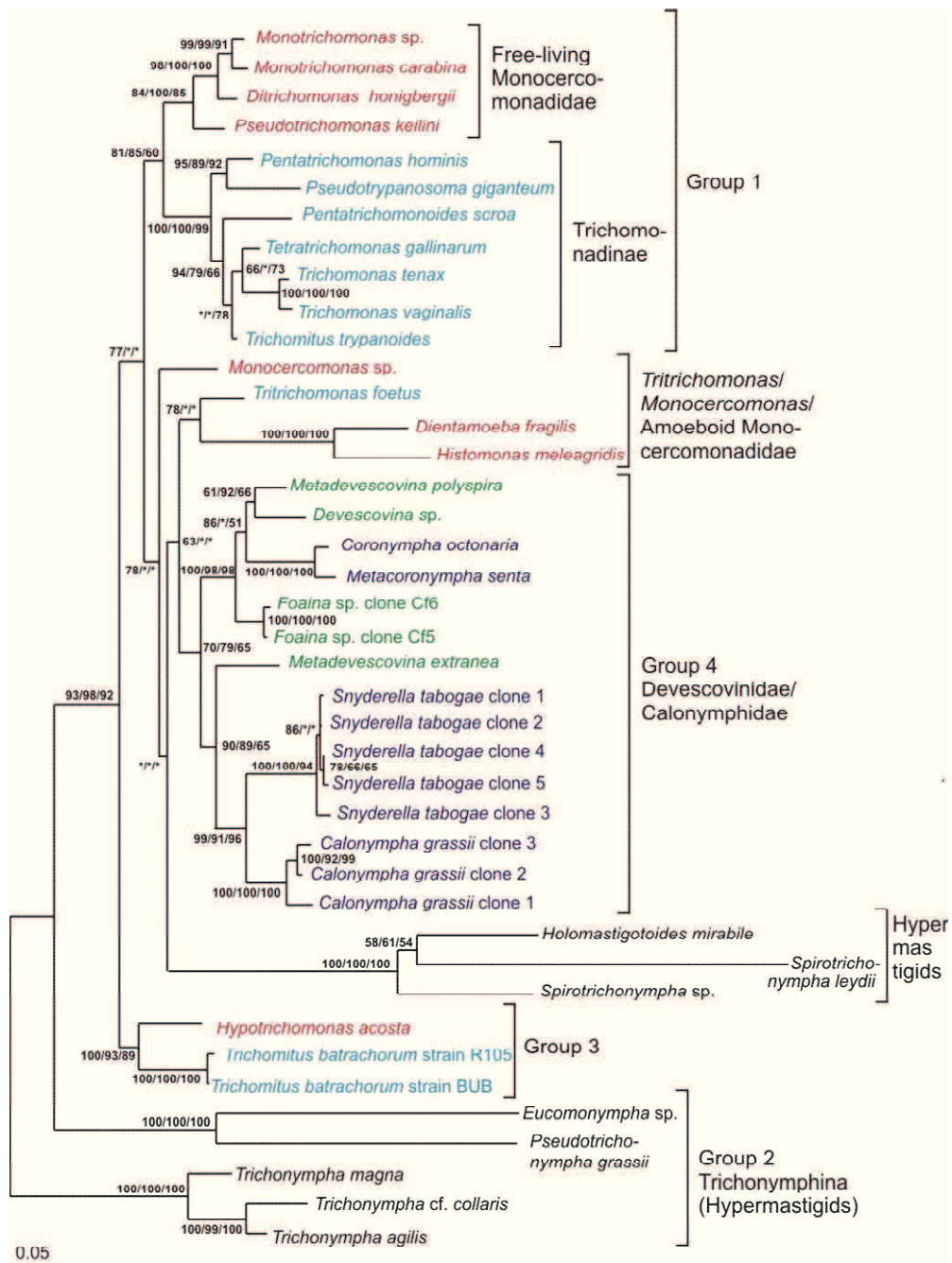
Figure 16: Molecular phylogeny of Parabasalids inferred from SSU rRNA gene sequences (from Gerbod et al. 2002). Trichomonad families are in blue (Trichomonadidae), in red (Monocercomonadidae), in green (Devescovinidae), and in purple (Calonymphidae) whereas hypermastigids are in black.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

by several authors (Hampl et al. 2004; Ohkuma et al. 2005). Consequently, based on morphological and molecular data, Adl et al. (2005) proposed classifying the parabasalids into four clades: Trichomonadida, Cristamonadida, Spirotrichonymphida, and Trichonymphida. However, as stated by others (Hampl et al. 2004, 2006; Carpenter and Keeling 2007; Noël et al. 2007; Noda et al. 2009), although molecular data supported the monophyly of the Cristamonadida, the creation of this order caused the paraphyly or polyphyly of the order Trichomonadida. Thus, the revision proposed by Adl et al. (2005) did not consider the paraphyly or polyphyly of some higher parabasalid taxa. More recently, Cepicka et al. (2010) have proposed dividing the parabasalids into six classes: Trichonymphea, Spirotrichonymphea, Cristamonadea, Tritrichomonadea, Hypotrichomonadea, and Trichomonadea **(Figure 17)**. In this classification, some taxa did not form monophyletic groups suggesting that this revision did not solve all the problems with the systematic of parabasalids.

IV-3 **Based on the analysis of protein sequences**

Comparative phylogenetic studies have demonstrated that single gene phylogenies based on either RNA or protein can, depending on the group in question, be very misleading. It is therefore critical to compare and assess the parabasalid SSU rRNA tree with those of multiple protein genes. Certain proteins including iron-containing superoxide dismutase, class II fumarase and glyceraldehydes-3-phosphate dehydrogenase (GAPDH) have first been used as phylogenetic markers for trichomonads, albeit with limited taxonomic sampling (Viscogliosi et al. 1996; Viscogliosi and Müller 1998; Gerbod et al. 2001). In a further study, new sequences from seven diverse parabasalids including one hypermastigid sequence were added to data sets of GAPDH, enolase, α-tubulin and β-tubulin and used to construct phylogenetic trees (Gerbod et al. 2004). In particular, the GAPDH tree was well resolved and identical in topology to the SSU rRNA tree (polyphyly of Trichomonadidae and Monocercomonadidae and basal position of the hypermastigid *Trichonympha*) suggesting that GAPDH should be a valuable tool in phylogenetic studies of parabasalids. Ohkuma et al. (2007) extended this previous study on the protein phylogeny of GAPDH, enolase, and tubulins by identifying the sequences of these four proteins from five genera of termite gut dwellers representative of the Pseudotrichonymphida, Trichonymphida, and Cristamonadida. GAPDH-based tree yielded a good resolution while the trees of enolase and tubulins were poorly resolved particularly in the branching orders of higher taxonomic groups. Interestingly

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
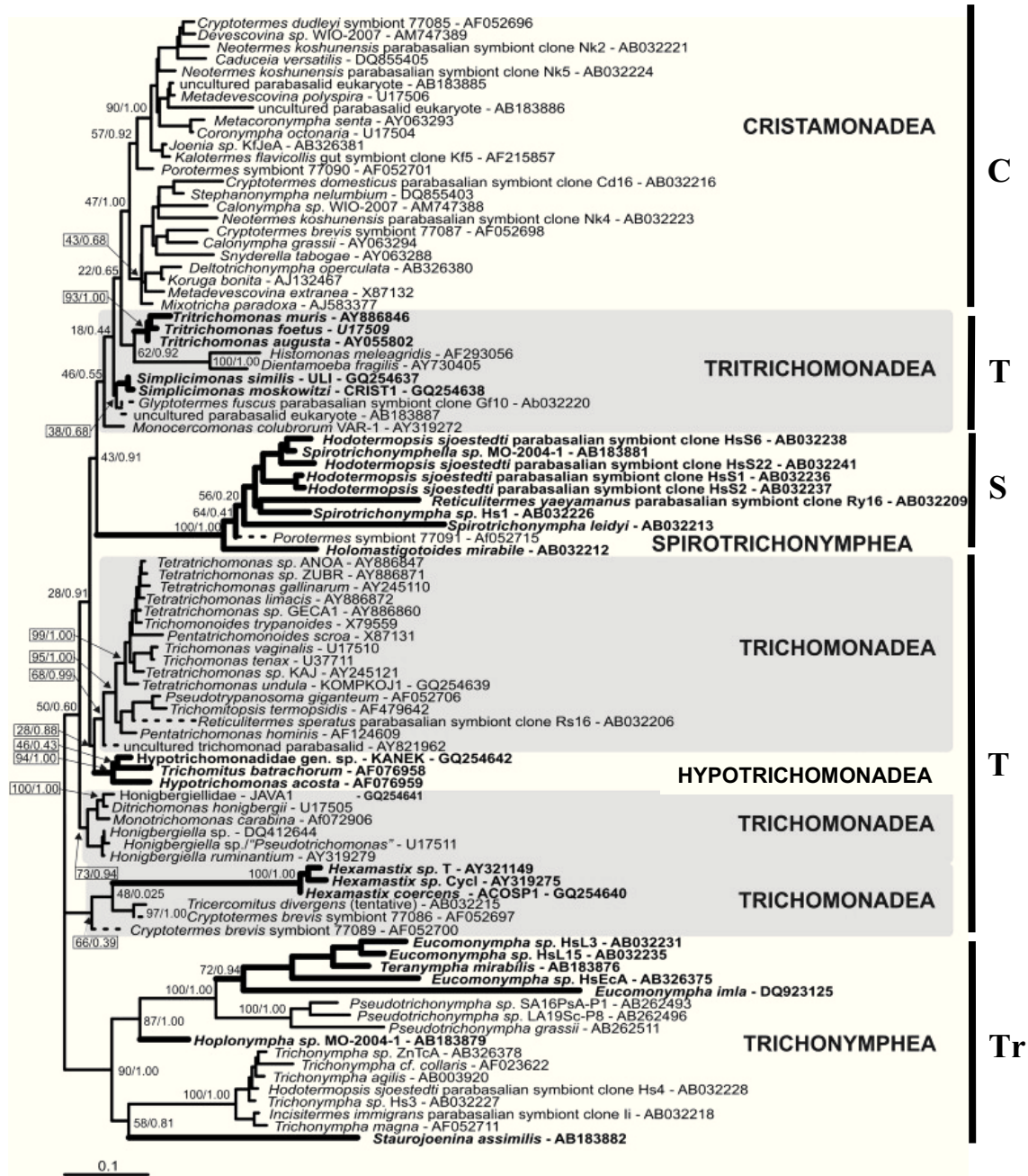
Figure 17: Molecular phylogeny of parabasalids inferred from SSU rRNA gene sequences (from Cepicka et al. 2010). On the right are indicated the parabasalid orders from Adl et al. (2005). **Abbreviations:** **C:** Cristamonadida ; **T:** Trichomonadida ; **S:** Spirotrichonymphida ; **Tr:** Trichonymphida.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

when the four protein sequences were combined (concatenated data set), the phylogenetic relationships among the parabasalids were identical to those inferred from the GAPDH data. In this tree, the Spirotrichonymphida, Trichonymphida, and Cristamonadida were monophyletic while the clade Trichomonadida was paraphyletic. More recently, a conserved single-copy gene encoding the largest subunit of RNA polymerase II (Rbp1) has been suggested as a useful marker (Malik et al. 2011). However only cultured parabasalids have been investigated so far and preliminary attempts found that the gene sequence was hardly obtained from termite-gut symbionts.

IV-4 The quest of the root of Parabasalia

As stated above, Trichonymphea, the most morphologically complex group of parabasalids, frequently occurs as the most basal lineage of Parabasalia in phylogenetic analyses based on SSU rRNA gene (Dacks and Redfield 1998; Ohkuma et al. 1998, 2000) or concatenated data of protein sequences (Ohkuma et al. 2007). Support for this rooting has often been poor, however, and the differences among several alternative positions in parabasalid lineages have been insignificant. This uncertainty has been explained by the absence of any close outgroup species to Parabasalia (Hampl et al. 2004). The parabasalid lineage usually branches out deeply in eukaryote phylogenies, and the branch leading to Parabasalia is very long. Such a long branch may attract the fast-evolving Trichonymphea, causing artificial rooting. Furthermore, the gene encoding GAPDH of Parabasalia has been derived from a bacterium *via* lateral gene transfer (Viscogliosi and Müller 1998), which disturbs root inference in comparison to other eukaryotes.

V *Blastocystis sp.*: an enigmatic Stramenopile
V-1 Classification, morphology and life cycle

Blastocystis sp. is an anaerobic parasite that inhabits the gastro-intestinal tract of humans and a wide range of animals, such as mammals, primates, birds, amphibians and less frequently rodents, reptiles, insects and mollusks (Stenzel and Boreham 1996; Abe et al. 2002; Tan 2004, 2008). It was only recently unambiguously placed among the Stramenopiles (or heterokonts), a complex and heterogeneous evolutionary assemblage of heterotrophic and photosynthetic protozoa (Silberman et al. 1996; Arisue et al. 2002; Riisberg et al. 2009). Interestingly, *Blastocystis* sp. is the only Stramenopile known to commonly cause infection in humans. At the morphological level, four major forms have been described in stools and/or *in*

vitro cultures **(Figure 18)**: vacuolar, granular, amoeboid, and cyst forms (Stenzel and Boreham 1996; Tan 2008; Suresh et al. 2009).

Vacuolar and granular forms are the most easily recognizable and frequently observed in laboratory culture and stool samples. Although rarely reported, the irregular amoeboid form has been postulated to play a role in pathogenesis (Tan and Suresh 2006; Katsarou-Katsari et al. 2008) but this hypothesis was clearly contradicted in a recent study (Souppart et al. 2009). Numerous data obtained from experimental infectivity studies in animals with the cyst form demonstrate that the water- and environmentally-resistant infective cyst undoubtedly represents the transmissible stage of the parasite (Suresh et al. 1993, 2005; Moe et al. 1996; Yoshikawa et al. 2004a). Taking into account these observations, a life cycle for *Blastocystis* sp. was proposed with the cyst as the infectious stage (Tan 2008) **(Figure 19)**.

Upon ingestion of cysts, the parasite undergoes excystation in the large intestine and develops into vacuolar forms. In humans, vacuolar forms divide by binary fission and may develop into amoeboid or granular forms. Then, the encystation may occur during the crossing along the colon before cyst excretion in the feces. Therefore, *Blastocystis* sp. lives in oxygen poor environments and is characterized by the presence of some double-membrane surrounded-organelles named MLOs for mitochondria-like organelles (Nasirudeen and Tan 2004) **(Figure 20)**. Sequencing of complete circular DNA of the MLOs of this parasite by different authors (Stechmann et al. 2008; Pérez-Brocal and Clark 2008; Wawrzyniak et al. 2008) show that these cellular compartments have metabolic properties of both aerobic and anaerobic mitochondria.

V-2 <u>Identification</u>

Most frequently, *Blastocystis* sp. is diagnosed by light microscopic examination of fecal material. The most common approaches for the detection of *Blastocystis* sp. (Stenzel and Boreham 1996; Stensvold et al. 2007b; Tan 2008) consist of direct smear and xenic *in vitro* culture. However, culturing this parasite is time consuming and can bias subsequent genotyping. Therefore, to overcome these limitations, several molecular polymerase chain reaction (PCR)-based diagnostic approaches directly from feces have been recently described (Abe et al. 2003a, b; Yoshikawa et al. 2003, 2004b, c; Scicluna et al. 2006; Stensvold et al. 2006; Santin et al. 2011). Recent studies comparing the relative performances of these various diagnostic methods (Suresh and Smith 2004; Stensvold et al. 2007b; Parkar et al. 2007; René et al. 2009) have shown that the PCR approach was as sensitive as the culture.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
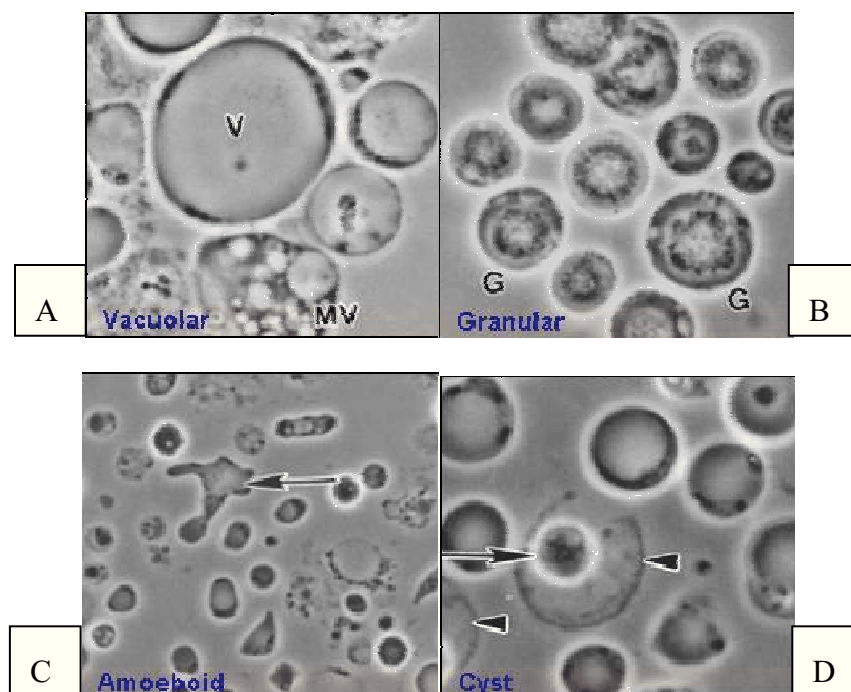*Università degli studi di Sassari*

Figure 18: Morphological forms of *Blastocystis* sp. viewed under light microscopy (from Tan 2004). **A:** Vacuolar; **B:** Granular; **C:** Amoeboid; **D:** Cyst.
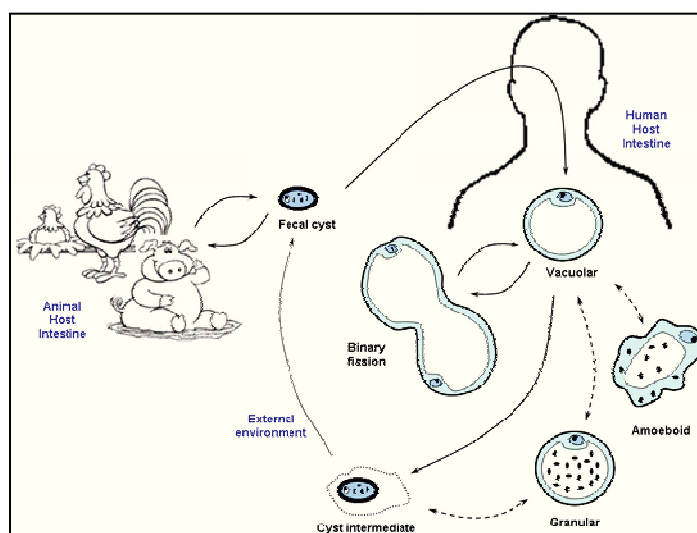

Figure 19: Proposed life cycle for *Blastocystis* sp. (modified from Tan 2004).

More recently, Poirier et al. (2011) have reported a highly-sensitive real-time PCR (qPCR) assay developed to detect *Blastocystis* sp. in stool samples. This assay also allows subtyping of isolates by direct sequencing of qPCR products.

V-3 Genetic diversity

Blastocystis sp. isolates from humans and other animals have been reported to be morphologically indistinguishable. However, extensive genetic variation among numerous *Blastocystis* sp. isolates from both humans and animals has been mainly observed by PCR-Restriction Fragment Length Polymorphism (RFLP) (Hoevers et al. 2000; Kaneda et al. 2001; Abe et al. 2003a and b; Rivera and Tan 2005) and PCR using Sequenced-Tagged-Site (STS) primers (Yoshikawa et al. 2004b, c; Yan et al. 2006; Li et al. 2007a, b; Yoshikawa et al. 2009). This considerable genetic divergence among isolates was subsequently confirmed by molecular phylogenies including those mainly inferred from small subunit (SSU) rRNA gene sequences (Arisue et al. 2003; Noël et al. 2003, 2005; Abe 2004; Scicluna et al. 2006; Özyurt et al. 2008; Rivera 2008; Jones et al. 2009; Souppart et al. 2009, 2010; Stensvold et al. 2009a; Yoshikawa et al. 2009; Parkar et al. 2010; Whipps et al. 2010) **(Figure 21)**.

From these molecular analyses, a consensus on *Blastocystis* sp. terminology was proposed (Stensvold et al. 2007a). In this new classification, all human, mammalian and avian isolates should be designated *Blastocystis* sp. and assigned to one of nine subtypes (STs: ST1 to ST9), each of the STs exhibiting sufficient genetic diversity to be classified as separate species. Thereafter, new STs (ST10 to ST13) were identified from zoo animals (Stensvold et al. 2009b; Parkar et al. 2010). All the phylogenetic analyses cited above were based on the comparison of full-length SSU rRNA gene sequences (approximately 1,800 bp). Scicluna et al (2006) have shown that a 600 bp-fragment of the SSU rRNA coding region provided sufficient information for accurate subtyping of *Blastocystis* sp. isolates.

V-4 Epidemiology

Numerous epidemiological surveys carried out in different countries identify *Blastocystis* sp. as the most common eukaryotic organism reported in human fecal samples. Overall the prevalence of *Blastocystis* sp. is higher than those of other intestinal parasitic protozoa such as *Giardia*, *Entamoeba* and *Cryptosporidium* (Boorom et al. 2008) **(Figure 22)**. An increasing trend in identification of *Blastocystis* sp. suggests that it is an emerging parasite with a worldwide distribution. Prevalence varies widely from country to country and

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
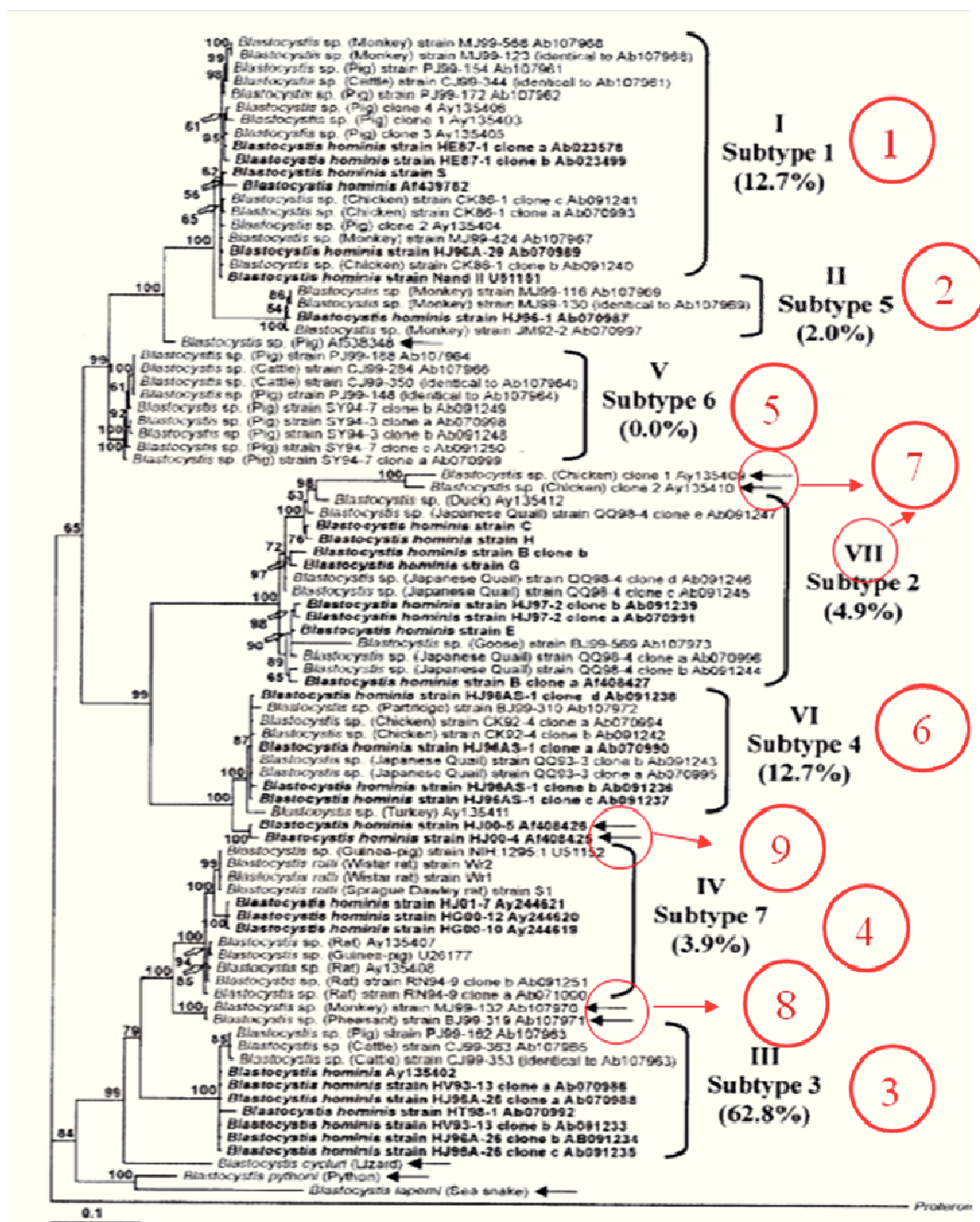
Figure 21: Identification of *Blastocystis* sp. subtypes based on the comparison of SSU rRNA gene sequences (from Noël et al. 2005). Groups proposed by Noël et al. (2005) (I to VII) are indicated in black. The corresponding ST1 to ST9 proposed in the consensus terminology by Stensvold et al. (2007a) are indicated in red.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
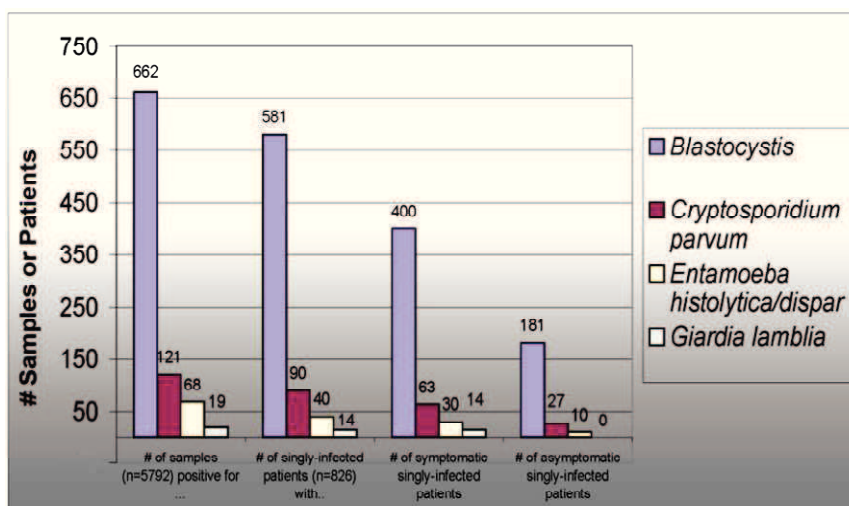
Figure 22: Prevalence of intestinal parasitic protozoa in the American population (from Boorom et al. 2008).

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

within various communities of the same country (Tan 2008; Souppart et al. 2009). In general, developing countries have higher prevalence of the parasite than developed countries. This difference can be explained by poor hygiene practices, close animal contact, and consumption of contaminated food or water (Li et al. 2007a; Leelayoova et al. 2008; Eroglu et al. 2010) since the fecal-oral route is considered to be the main mode of transmission of this parasite. Prevalence can be low in countries such as Japan (0.5 to 1%) (Hirata et al. 2007) and Singapore (3.3%) (Wong et al. 2008) and high in developing nations including Brazil (40.9%) (Aguiar et al. 2007), Egypt (33.3%) (Rayan et al. 2007), and Indonesia (60%) (Pegelow et al. 1997). In some countries, the prevalence can be rather variable and ranges of 1.9 to 32.6% in China (Li et al. 2007b) and of 0.9% to 45.2% in Thailand (Saksirisampant et al. 2003, 2006) depending on the subpopulations studied. Such variations within the same country could reflect true differences between communities or the use of different diagnostic approaches.

V-5 **Predominance of ST3 in the human population and zoonotic potential and transmission of this parasite**

Most of the samples included in published epidemiological surveys represented simple infections. Available data on the prevalence of mixed infections i.e. infections by at least two different STs and resulting of probable multiple sources of infection, indicate that it is similar in different countries and roughly comprised between 2.6% and 14.3% (Yan et al. 2006; Li et al. 2007a; Dogruman et al. 2008; Souppart et al. 2009). However, the true distribution of mixed infections remains difficult to ascertain in a particular individual and likely underestimated as this depends on the method employed for subtyping. In almost all the studies reported so far, a large majority of human infections with *Blastocystis* sp. were attributable to ST3 isolates (worldwide average exceeding 60% according to Souppart et al. 2009). Only few exceptions showed the higher prevalence of ST4 and ST1 in Spain (Dominguez-Marquez et al. 2009) and Thailand (Thathaisong et al. 2003), respectively. Collectively, these studies suggest that the dominant ST3 was the only ST of human origin as was first proposed by Noël et al. 2005 even if it can also be found in some animal groups. Consequently, the predominance of this ST might be mainly explained by large-scale human-to-human transmission (Yoshikawa et al. 2000). Proportions of ST1 to ST4 differ between locations. ST6 and ST7 are common in Asia but rarely observed in European countries. ST8 has only been found in the UK and Italia while ST9 was only identified in Japan. Regarding ST5 and ST11 to ST13, they are found episodically in humans (Parkar et al. 2007, 2010; Yan

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

et al. 2007). Comparison of SSU rRNA gene sequences, cross-transmission experiments and respective prevalence of different STs in human population indicate that almost all of the known STs of supposed animal origin are likely zoonotic and able to infect humans (Noël et al. 2005; Parkar et al. 2007, 2010; Yan et al. 2007; Yoshikawa et al. 2009). Therefore, a higher risk of *Blastocystis* sp. infection has been found in people with close animal contact including zoo-keepers and abattoir workers.

V-6 <u>Clinical aspects of blastocystosis</u>

During the last two decades, it was widely debated in the literature whether *Blastocystis* sp. is a truly pathogenic or commensal organism. However, accumulating recent *in vitro* and *in vivo* studies shed new light on the pathogenic power of this parasite suggesting that *Blastocystis* sp. infection is associated with a variety of gastrointestinal disorders. This parasite has been implicated in various gastrointestinal aspecific symptoms such as diarrhea, abdominal pain, nausea, fatigue, constipation, flatulence, and vomiting (Boorom et al. 2008; Tan 2008; Stensvold et al. 2009c). *Blastocystis* sp. may also play a significant role in chronic gastrointestinal illness such as irritable bowel syndrome (IBS) (Giacometti et al. 1999; Yakoob et al. 2004a, 2010a, b; Dogruman-Al et al. 2009; Jones et al. 2009; Tai et al. 2010; Yamamoto-Furusho and Torijano-Carrera 2010). Indeed, in a recent study excluding patients with other intestinal pathogenic protozoa and bacteria, Dogruman-Al et al. (2010) reported that most IBS patients (76%) were infected with *Blastocystis* sp. Numerous cases were also reported regarding the association of *Blastocystis* sp. infection and urticaria (Gupta and Parsi 2006; Katsarou-Katsari et al. 2008; Hameed et al. 2010). In addition, *Blastocystis* sp. has increasingly been implicated for diarrheal illness in immunocompromised individuals including HIV/AIDS and cancer (Cirioni et al. 1999; Prasad et al. 2000; Tasova et al. 2000; Kurniawan et al. 2009; Tan et al. 2009; Idris et al. 2010). The few studies that have addressed ST-dependent differences in pathogenicity in isolates from symptomatic and asymptomatic individuals (Kaneda et al. 2001; Yan et al. 2006; Dogruman-Al et al. 2008, 2009; Özyurt et al. 2008; Tan et al. 2008; Eroglu et al. 2009; Souppart et al. 2009; Stensvold et al. 2009c) have provided conflicting results mainly due to the limited number of patients examined in each epidemiological survey. However, in patients with IBS and chronic diarrhea, either ST3 (Dogruman-Al et al. 2009; Jones et al. 2009) or ST1 (Yakoob et al. 2010a) were the most commonly found. Similarly, either ST3 (Katsarou-Katsari et al. 2008; Hameed et al. 2010) or ST2 (Vogelberg et al. 2010) were shown to be predominant in patients with urticaria.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

V-7 <u>Development of animal models of blastocystosis</u>

A variety of experimental infection studies involving different animal species have been described (Tan 2008). In pionner studies using rodents (Suresh et al. 1993; Moe et al. 1997), animals became infected after inoculation of fecal cysts orally and of *in vitro* axenic-culture intracecally. Infections were generally self-limiting, although some mice showed weight loss and lethargy. Histological examination of the cecum and colon revealed intense inflammatory-cell infiltration and oedematous lamina propria. These authors also show that there is an age-related susceptibility to *Blastocystis* sp. in mice. Indeed, juvenile BALB/c mice are more susceptible than adult mice, and 8-week-old adult BALB/c mice are totally resistant to *Blastocystis* sp. One of the subsequent experimental infection studies showed that as few as 10 to 100 cysts were able to establish an infection *via* intracecal or oral inoculations (Yoshikawa et al. 2004a). More recently, Hussein et al. (2008) tested the infectivity of human ST1, ST3, ST6 and ST7 isolates obtained from both asymptomatic and symptomatic patients in rats. Interestingly, the moderate and severe degrees of pathological changes observed at 6 weeks postinfection (PI) were found only in symptomatic isolates infected rats while mild degree was found only in asymptomatic isolates infected rats. ST1 symptomatic isolates induced 25% mortality in rats concluding to its pathogenesis, while pathogenic and non-pathogenic variants exist among ST3 and ST6. In parallel, the same authors described intense inflammatory reaction and sloughing mucosa, oedema, and precancerous polyps in cecum and proximal colon tissues in symptomatic infected rats. Finally, 4-weeks-old mice were orally inoculated with *in vitro* culture of symptomatic human isolates (Elwakil and Hewedi 2010). In the highly infected mice showing slow locomotion, lethargy, body weight decrease and for some of them, mucus feces, histopathological examination of large intestine revealed at 2 weeks PI that the parasite infiltrated the lamina propria, the submucosa, and the muscle layers. This was accompanied by active colitis with infiltration of mixed inflammatory cells namely lymphocytes, histiocytes and few eosinophils.

V-8 *In vitro* studies and identification of potential virulence factors

In complement of the development of animal models, few *in vitro* studies were performed to investigate the cytopathic effects of *Blastocystis* sp. on mammalian cell cultures and identify virulence factors. In a first study, Long et al. (2001) showed that neither *Blastocystis* sp. ST1 cells nor culture filtrates induced cytopathic effects in HT-29 and T-84 human colonic epithelial cells. However, it was observed that 24 hours incubation induced the production of proinflammatory cytokines interleukin (IL)-8 and granulocyte-macrophage

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

colony stimulating factor (GM-CSF) suggesting that the parasite was able to modulate the host immune response. Recently, the production of IL-8 in a nuclear factor kappa-light-chain enhancer (NF-kB)-dependant manner from T-84 cells was demonstrated and shown to be induced by cysteine proteases of *Blastocystis* sp. ST4 (Puthia et al. 2008). The same authors have also shown that *Blastocystis* sp. ST4 induced apoptosis in the rat epithelial cell line IEC-6 in a contact-independent manner, decreased transepithelial resistance, and increased epithelial permeability (Puthia et al. 2006). These data suggested that *Blastocystis* sp.-induced apoptosis in host cells and altered epithelial barrier function might play an important role in the pathogenesis of this parasitic infection. As in other parasitic protozoa (Sajid and McKerrow 2002), cysteine proteases of *Blastocystis* sp. which are mainly localized in the central vacuole (Puthia et al. 2008) and show high activity in parasitic lysates (Sio et al. 2006) are undoubtedly involved in parasite survival *in vivo* and represent virulence factors. Variations in cysteine protease activity were observed between ST4 and ST7 isolates which may be attributable to the likely existence of virulent and non-virulent *Blastocystis* sp. strains (Mirza and Tan 2009). These proteases present in *Blastocystis* sp. lysates and spent medium are able to cleave human secretory IgA, the prevalent immunoglobulin defence at the mucosal surface (Puthia et al. 2005). Recently, a surface-located unusual cysteine protease, legumain was described in *Blastocystis* sp. ST7 (Wu et al. 2010). In contrast to plants, legumain appears to display a pro-survival role in the parasite and may function to activate other proteases.

V-9 <u>Treatment of blastocystosis</u>

Treatment of blastocystosis is usually considered if diarrhea is persistent and no other pathogen apart from *Blastocystis* sp. is identified in fecal specimens. In this case, metronidazole is considered first line therapy for *Blastocystis* sp. infection (Nigro et al. 2003; Cassano et al. 2005; Moghaddam et al. 2005). In a first evaluation of efficacy of this drug, Nigro et al. (2003) showed that immunocompetent individuals with *Blastocystis* sp. infection as the only evident cause of diarrhea responded to metronidazole treatment and consequently suggested that the parasite induced intestinal disease. However, there have been accumulating reports of treatment failure, particularly in patients with severe *Blastocystis* sp. infections (Haresh et al. 1999; Yakoob et al. 2004b; Moghaddam et al. 2005; Stensvold et al. 2008, 2010) suggesting the existence of drug resistant isolates. Accordingly, some genes coding for multi drug resistance (MDR) pump proteins (ABC transporters) and know to confer resistance to drugs in protozoan parasites such as *Leishmania, Trypanosoma,* and

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Plasmodium (Sauvage et al. 2009) have also been discovered in the *Blastocystis* sp. ST7 genome (Denoeud et al. 2011). Therefore, metronidazole-resistant strains of *Blastocystis* sp. could have arisen through the action of these MDR proteins. Moreover, some studies have shown that *Blastocystis* sp. strains isolated from patients may exhibit large differences in sensitivity to metronidazole (Haresh et al. 1999) and that the cyst form is highly resistant to the cytotoxic effect of the drug (Zaman and Zaki 1996). In such circumstances, several standard antimicrobials (cotrimoxazole, ornidazole, nitazoxanide, paromomycin, chloroquine and others) may be considered as second-choice drugs (Armentia et al. 1993; Ok et al. 1999; Cimerman et al. 2003; Diaz et al. 2003; Moghaddam et al. 2005; Rossignol et al. 2005; Stensvold et al. 2008; Mirza et al. 2010). Even if some of these drugs have been shown to be globally effective against *Blastocystis* sp., treatment failures were also largely reported. Moreover, extensive variations in drugs susceptibilities have been demonstrated among STs (Mirza et al. 2010).

V-10 <u>Genome sequence</u>

Recently, Denoeud et al. (2011) revealed the genome sequence of a *Blastocystis* sp. ST7 isolate, which is the smallest Stramenopile genome sequenced so far (18.8 megabases). Briefly, analyses of the predicted proteome and secretome have identified proteins potentially involved in adhesion and pathogenicity (cysteine proteases) of the parasite. This parasite possesses also enzymatic machinery to fight against oxidative burst resulting from its own metabolism and induced by the host immune system. Moreover, the authors have established the *in silico* proteome of *Blastocystis* sp. MLOs, described their metabolic characteristics, and confirmed the mitochondrial nature of these puzzling organelles. These data together with those of recent *in vivo* and *in vitro* studies (see above) allowed to proposing a model for pathogenesis of this parasite (Denoeud et al. 2011; Tan 2008) **(Figure 23)**.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari
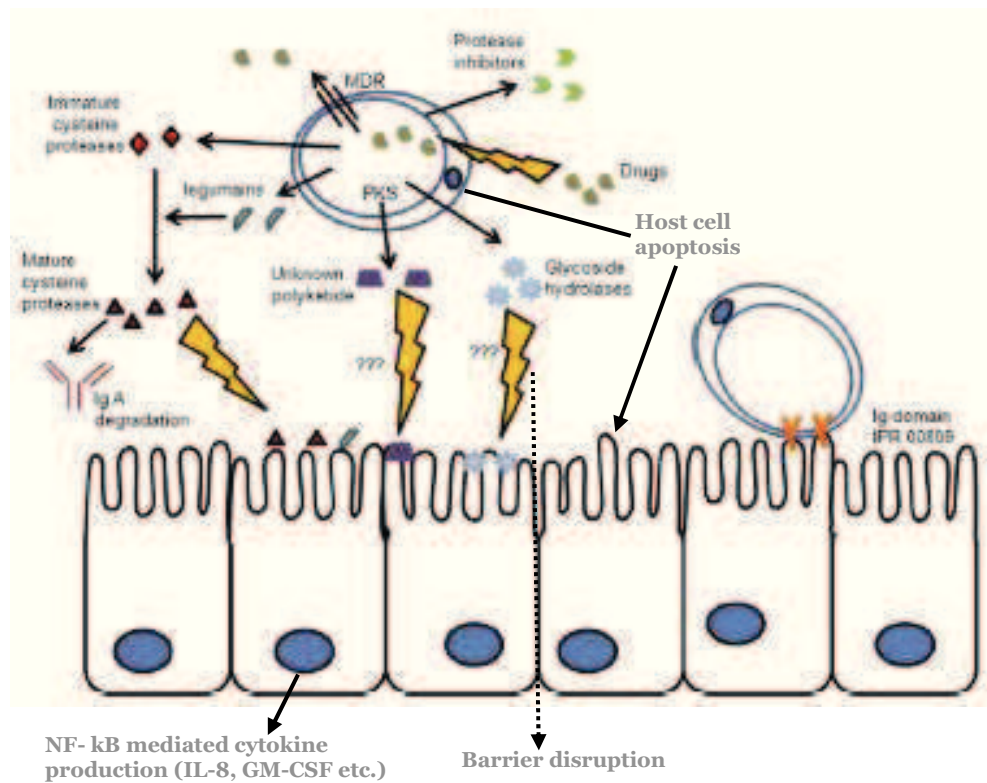
Figure 23: Model for pathogenesis of *Blastocystis* sp. (from Denoeud et al. 2011). **Abbreviations:** PKS, polyketide synthase; MDR, multi drug resistance.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Results

_____

Part I

_____

IDENTIFICATION AND PATHOGENICITY OF TRICHOMONADS

I-1 <u>Intoduction</u>

As stated above, four trichomonad species are mainly found in humans: *Trichomonas vaginalis* in the urogenital tract, *Trichomonas tenax*, in the oral cavity; *Pentatrichomonas hominis*, in the large intestine and *Dientamoeba fragilis* in the colon. Among them, only *Trichomonas vaginalis* and *Dientamoeba fragilis* are considered pathogens. The pathogenicity of *Pentatrichomonas hominis* in humans remains largely debated since some studies have recognized an association between diarrhea and this parasite in dogs and cats (Gookin et al. 2007; Kim et al. 2010).

Until the last years, trichomonad parasites of humans were thought to be site-specific. Recently, several studies revealed the large prevalence of trichomonads in the respiratory tract of patients with underlying pulmonary pathologies including PCP and ARDS (for review see Duboucher et al. 2008). However, their morphological identification is difficult since most trichomonads detected in unusual locations do not present the universally known flagellated form.

Sequences of genes coding for SSU rRNA and the internal transcribed spacer 1 (ITS1)-5.8 S-ITS2 region are available in databases for a large set of trichomonad species including those found in humans. These molecules represent molecular tools for the identification at the species level of trichomonads particularly in some unusual locations. In collaboration with the Centre Hospitalier Régional Universitaire de Tours, we have identified trichomonad species observed i) in the stools of two patients with gastrointestinal troubles **(Article 1)** and ii) in samples of five patients presenting different pathologies and symptoms **(Article 2)**.

I-2 <u>Results</u>

Article I (see above)

<u>Meloni D</u>, Mantini C, Goustille J, Desoubeaux G, Maakaroun-Vermesse Z, Chandenier J, Gantois N, Duboucher C, Fiori PL, Dei-Cas E, Duong TH, Viscogliosi E. Molecular identification of *Pentatrichomonas hominis* in patients with gastrointestinal symptoms. Journal of Clinical Pathology, 2011, 64:933-935

Article 2

<u>Meloni D</u>, Desoubeaux G, Mantini C, gantois N, Noda S, Ohkuma M, Delhaes L, Chandenier J, Duboucher C, Fiori PL, Dei-Cas E, Duong TH, Viscogliosi E. Unusual occurrences of

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Trichomonas tenax and *Trichomonas vaginalis* in humans. Journal of Clinical Pathology, 2011, in preparation

<u>Brief description Article 2:</u> This report includes five patients presenting different pathologies and symptoms **(Table 1)**. According to their pathologies/symptoms/treatment and follow-up at hospital, different biological samples from these patients **(Table 1)** have been examined by microscopic observation of stained cells and revealed numerous microorganisms assumed to be trichomonads **(Figure 24)**. DNA extraction from the samples, PCR amplification of the ITS1-5.8S-ITS2 region and cloning of the PCR products were performed as described elsewhere (Duboucher et al. 2006 and Article 1). Following cloning of the PCR products, 5 clones were arbitrarily selected for each sample and sequenced on both strands **(Figure 25)**. According to these sequences, trichomonad cells were unequivocally identified as *Trichomonas tenax* or *Trichomonas vaginalis* in the samples. In the colic biopsy of one patient (JF), we identified a coinfection with both *Trichomonas* species. Globally we confirm that trichomonads are able to colonize different unusual sites under favourable conditions. For instance *Trichomonas tenax* is a known commensal of the oral cavity but this microorganism is able to colonize the lungs (patient XR). More interestingly, to our knowledge, this is the first report of *Trichomonas tenax* in the feces as well as in the colon of patients (patients VJ and JF) that raises the question of its transit through the intestinal tract. Moreover, this parasite, together with *Trichomonas vaginalis* has been found in the colon of an immunocompetent individual with diarrhea (JF).

I-3 <u>Conclusions</u>

In the first paper, we identified *Pentatrichomonas hominis* in the stools of two patients with intestinal pain and diarrhea. According to our data and those available in the literature, *Pentatrichomonas hominis* could be the causative agent of gastrointestinal troubles especially in children. Interestingly, the same species could possibly colonize many mammal species including dogs, bovid and humans that raised the question of the as-yet-unknown zoonotic potential of this trichomonad as well as of its "human" origin. Moreover, this parasite is able to survive in water for a relatively long period of time (at least 10 days) as pseudocysts suggesting a significant potential of this species to use waterborne routes of transmission.

In the second paper (in preparation), we identified *Trichomonas vaginalis* and/or *Trichomonas tenax* in various samples from patients presenting different pathologies. In one of these patients we confirm that *Trichomonas tenax* is able to colonize lungs. Moreover, this

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

species was also identified in the oral cavity and stools of the same patient suggesting for the first time that this parasite is able to transit through the entire gastrointestinal tract. In addition, *Trichomonas vaginalis* and *Trichomonas tenax* were found in the colic biopsy of a patient presenting diarrhea. Strikingly this patient did not present other intestinal parasites or pathogenic bacteria suggesting that *Trichomonas tenax* or *Trichomonas vaginalis* could be responsible for gastrointestinal disorders. Globally we confirm that trichomonads are able to colonize different unusual sites under favourable conditions. It is clear that these cases together with those that will be described in further studies, will better define the incidence, pathogenicity and mechanisms of invasion of these microorganisms as well as their possible links with diseases or immunodeficiency.
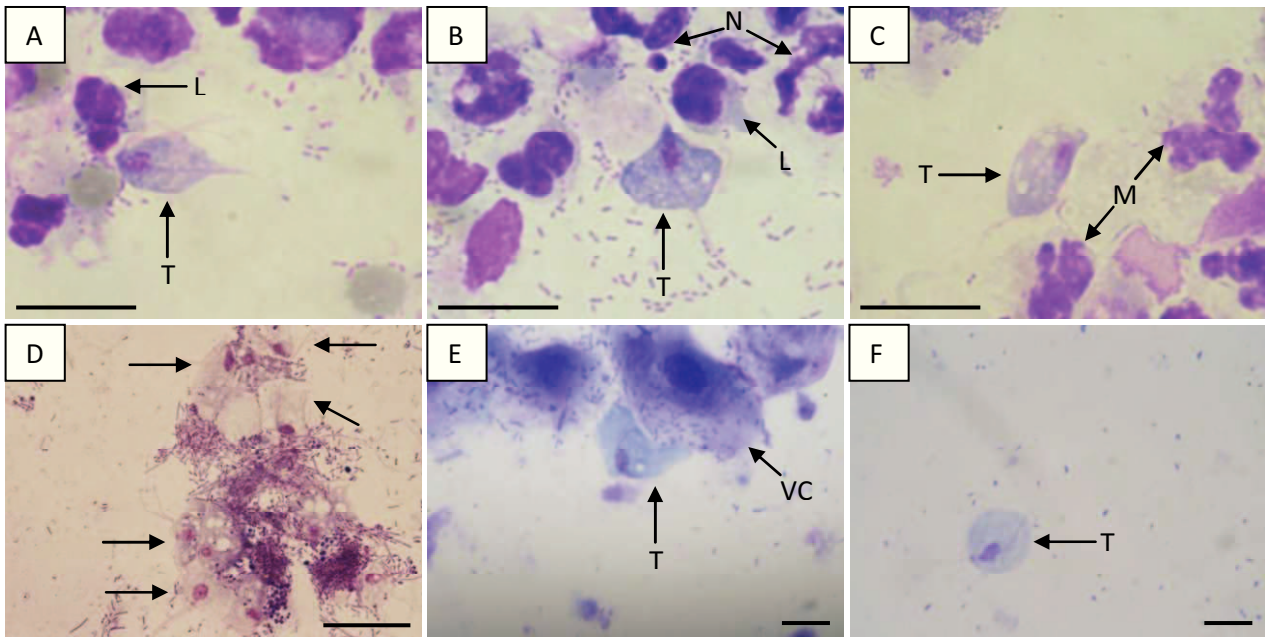
Figure 24: Cytological appearance of trichomonad cells in samples (May-Grünwald-Giemsa staining). **A**, **B**, **C:** patient XR; Trichomonads recognizable in the pleural fluid in the vicinity of lymphocytes, neutrophile polymorphonuclear and macrophage cells. **D:** patient VJ; Trichomonads in the oral wash specimen (arrows). **E**, **F:** patient FC; Trichomonads observed in the urine in the vicinity of vaginal cells.

Abbreviations: T: trichomonads; L: lymphocytes; N: neutrophile polymorphonuclear; M: macrophage; VC: vaginal cell. The scale bar represents 10 µm.

```
Trichomonas vaginalis Tv1:MCP L29561   CGGTAGGTGAACCTGCCGTTGGATCAGTTCTAGTT*TTAATAACTAACACCAA*CTTC*TTTTATT**AAACAAAAACCAATACAAAATTAAA*AACTAA  94
FC clones 1 to 5                               TRICHO-F   GGATCAGTTCTAGTT*TTAATAACTAACACCAA*CTTC*TTTTATT**AAACAAAAACCAATACAAAATTAAA*AACTAA  74
JF clone 2                                                GGATCAGTTCTAGTT*TTAATAACTAACACCAA*CTTC*TTTTATT**AAACAAAAACCAATACAAAATTAAA*AACTAA  74
JF clones 3 and 4                                         GGATCAGTTCTAGTT*TTAATAACTAACACCAA*CTTC*TTTTATT**AAACAAAAACCAATACAAAATTAAA*AACTAA  74
JF clone 5                                                GGATCAGTTCTAGTT*TTGATAACTAACACCAA*CTTC*TTTTATT**AAACAAAAACCAATACAAAATTAAA*AACTAA  74
Trichomonas tenax 30207 U86615                             ATCAGTTCTAGTT*TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  73
MOC clones 1 to 5                                         GGATCAGTTCTAGTT**TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  75
XR clones 1 to 5                                          GGATCAGTTCTAGTT**TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  75
JF clone 1                                               GGATCAGTTCTAGTT**TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  75
VJB clones 1 to 5                                        GGATCAGTTCTAGTT**TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  75
VJF clones 1 to 5                                        GGATCAGTTCTAGTT**TAATAACTAATACCAA*CTTCTTTTTATT**AATCAAAAACTAATACAATTATAAATAACTAA  75
Tetratrichomonas sp. TXB3 AY244461                        ATCAGTTCTAGTT*TTAATAACTAACACCAA*CTTCTTTTAATTAA*TTTTTAAAACCACATA**CAACAACAAAAAA  73
Tritrichomonas foetus DQ243910                           AGTT*TCGTTAATAATTACAAACATATTTTTTTAATTTCTA**TAACTATTT***ATACAAAATTAAACACATAA  69
Pentatrichomonas hominis 30098 U86616                     ATCAGTTCTAGTT*********TAACACCAAACTTCTTTTT**T**GAT**AAAACCAATACAAT*ATAAAC**CTAA  60
                                       ----------SSU rRNA----------|-------------------------------ITS1-----------------------------

Trichomonas vaginalis Tv1:MCP L29561   CTTCATCAAAAA**CCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCAAACAT 191
FC clones 1 to 5                       CTTCATCAAAAA**CCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCAAACAT 171
JF clone 2                             CTTCATCAAAAA**CCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGCGGCATAATGTGTTAAGTAACCGGAGTTGCAAACAT 171
JF clones 3 and 4                      CTTCATCAAAAA**CCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCAAACAT 171
JF clone 5                             CTTCATCAAAAA**CCAAGTCTCTAAGCAAT*GGATGTCTTGGCTTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCAAACAT 171
Trichomonas tenax 30207 U86615         CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 170
MOC clones 1 to 5                      CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 172
XR clones 1 to 5                       CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 172
JF clone 1                             CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 172
VJB clones 1 to 5                      CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 172
VJF clones 1 to 5                      CTTCATCAAAAA**TCAAGTCTCTAAGCAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGGCATAATGTGTTAAGTAACCGGAGTTGCATACAT 172
Tetratrichomonas sp. TXB3 AY244461     CATAATAAAAAA*GTTAGGACTCTAATCAAC*GGATGTCTTGGCTCCTTACACGATGAAGAACGTTGCATAATGTGATAAGTAACTGGAGTTGCGAACAT 171
Tritrichomonas foetus DQ243910         TCTAA***AAAA**TTTAGACCTTAGGCAAT*GGATGTCTTGGCTTCTTACACGATGAAGAACGTTGCATAATGCGATAAGCGGCTGGATTAGCTTTCTT 163
Pentatrichomonas hominis 30098 U86616  CTT*****AATG**TAAGGTCTCTAAACAAT*GGATGTCTTGGCTCCTCACACGATGAAGAACGTGCATAATGTGATAAGTAACTGGAGTTGCGAACAT 152
                                       -----------------|--------------------------------5.8S rRNA---------------------------------

Trichomonas vaginalis Tv1:MCP L29561   CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTCGATC*TCGGTCG*AGAA**GCATGGGTGTGACAGTACTACATCTTT*TATAAT****AA 281
FC clones 1 to 5                       CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTCGATC*TCGGTCG*AGAA**GCATGGGTGTGACAGTACTACATCTTT*TATAAT****AA 261
JF clone 2                             CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTCGATC*TCGGTCG*AGAA**GCATGGGTGTGACAGTACTACATCTTT*TATAAT****AA 261
JF clones 3 and 4                      CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTCGATC*TCGGTCG*AGAA**GCATGGGTGTGACAGTACTACATCTTT*TATAAT****AA 261
JF clone 5                             CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTCGATC*TCGGTCG*AGAA**GCATGGGTGTGACAGTACTACATCTTT*TATAAT****AA 261
Trichomonas tenax 30207 U86615         CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 260
MOC clones 1 to 5                      CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 262
XR clones 1 to 5                       CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 262
JF clone 1                             CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 262
VJB clones 1 to 5                      CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 262
VJF clones 1 to 5                      CATGACAGGTT*AATCTTTGAATGCAAATTGCGCTAAACTTGGCT*TCGGCTG*AGAA**GCATGCGTGTAACAGTACAACATAATT*TATAAT****AA 262
Tetratrichomonas sp. TXB3 AY244461     TGTGACAAGTT*AATCTTTGAATGCAAATTGCGCTACCTTTGCTCCTTTGCTCTCGGGCA*TTGGTTGCATGTATGCCACAGTACTGCATAAAA*TAATAT****AT 264
Tritrichomonas foetus DQ243910         TGCGCACAAGTTCGATCTTTGAATGCACATTGCGCGCGTTTTAGC*TTGCTAGAACAC**GCATATATGTTACAGTAACCCATATTAATTTAACGCCAAA 260
Pentatrichomonas hominis 30098 U86616  TGTGACAAGTT*AATCTTTGAATGCAAATTGCGCATACCTGAATT*T*ATTC**AGAT**GCATATATATTCTAGTACTACATCTTT***TAAT****AA 238
                                       ----------------------------------5.8S rRNA------------------------------------|------------------

Trichomonas vaginalis Tv1:MCP L29561   TTCTTATTCTAAGCGAATAAGTAAATAATATATAAG*ACAAACAACACGTAGTCTGCCATACGCAGGAAGACCCGCTGAACTGAAGCA 368
FC clones 1 to 5                       TTCTTATTCTAAGCGAATAAGTAAATAATATATAAG*ACAAACAACACGTAGTCTGCCATACGCAGGA       TRICHO-R       328
JF clone 2                             TTCTTATTCTAAGCGAATAAGTAAATAATATATAAG*ACAAACAACACGTAGTCTGCCATACGCAGGA                       328
JF clones 3 and 4                      TTCTTATTCTAAGCGAATAAGTAAATAATATATAAG*ACAAACAACACGTAGTCTGCCATACGCAGGA                       328
JF clone 5                             TTCTTATTCTAAGCGAATAAGTAAATAATATATAAG*ACAAACAACACGTAGTCTGCCATACGCAGGA                       328
Trichomonas tenax 30207 U86615         TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCA                          323
MOC clones 1 to 5                      TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCAGGA                       328
XR clones 1 to 5                       TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCAGGA                       328
JF clone 1                             TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCAGGA                       328
VJB clones 1 to 5                      TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCAGGA                       328
VJF clones 1 to 5                      TTCTTATTCTAAGCGAATAAGTAAATAAATTATAAG*ACAAACTT*ACGTTGTCTGCTATACGCAGGA                       328
Tetratrichomonas sp. TXB3 AY244461     ATTCTTCTTTTCGCAAAGAAGCAAATATTATTTAAGCAAAAAGAT*ACGTAGTCTGCGACATACA                          328
Tritrichomonas foetus DQ243910         TTCTCTTTTTAAGCAAAGAAGCGAAAAACAAATATGTATTAACAAAA*GGGTTCTGTCTCATAT                           323
Pentatrichomonas hominis 30098 U86616  TTCTTCTTATAA***AAGAAGCAAATAACT*AAAAG**CAAACA***CGTAGTCTGGGATATGTA                          294
                                       -----------ITS2----------------------------|------------------LSU rRNA---------------
```

Figure 25: Alignment of sequences of the ITS1-5.8S-ITS2 region obtained from five patients in this study and those of trichomonad species of interest found in humans. Differences between the four of the JF clones obtained in this study and the homologous sequence of *Trichomonas vaginalis* are shaded. Sequences of the primers used in this study are underlined. In case of patient VJ, five clones were obtained from both oral was specimen (VJB) and stools (VJF). Gaps are represented by asterisks. **Abbreviations:** SSU: small subunit; LSU: large subunit.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Patients	Pathology/symptoms/treatment	Samples analyzed	Identification of trichomonads
XR	Gastrointestinal haemorrhage Pleural effusion	Pleural fluid	*Tritrichomonas tenax*
VJ	Lymphoma/chemotherapy	Oral wash specimen	*Tritrichomonas tenax*
		Stool	*Tritrichomonas tenax*
JF	Chronic entheropaty/diarrhea	Colic biopsy	*Tritrichomonas tenax*
			Trichomonas vaginalis
MOC	Myeloma/chemotherapy	Oral wash specimen	*Tritrichomonas tenax*
FC	Nervous breakdown Cephalea/faintness	Urine	*Trichomonas vaginalis*

Table 1: Pathologies/symptoms/treatment of patients and molecular identification of trichomonads found in the analyzed samples.

CORRESPONDENCE

Molecular identification of *Pentatrichomonas hominis* in two patients with gastrointestinal symptoms

INTRODUCTION

The trichomonad species *Pentatrichomonas hominis* colonises the gastrointestinal tract and is generally considered as a commensal organism in humans. However, some studies have recognised an association between diarrhoea and *P hominis* infection in dogs and cats.[1][2] In the present report, we have identified this species using molecular tools in two patients with gastrointestinal troubles. Our data suggest that *P hominis* is a possible zoonotic species with a significant potential of transmission by water and could be the causative agent of intestinal symptoms in children.

CASE REPORTS

An adult (case 1) was followed up for different pathologies including irritable bowel syndrome (IBS). Diarrhoeic stools of the patient were examined and were negative for intestinal parasites. Filter paper/slant culture technique for the recovery of *Strongyloides stercoralis* larval-stage nematodes from fresh faeces was performed. After 5 days, microscopic examination of the stool culture using merthiolate-iodine-formalin and RAL555 stains did not detect larval nematodes but numerous flagellates, provisionally identified as trichomonads (figure 1A,B). Although no treatment against trichomonads was administered to the patient, the stool examinations performed afterwards did not reveal trichomonads or other parasites.

A young child (case 2) presented with abdominal pain and loose stools without fever. A stool sample was examined by direct light microscopy and showed numerous cells assumed to be trichomonad organisms. Additional investigations did not reveal the presence of coinfecting intestinal parasites or pathogenic bacteria. The child was treated with metronidazole (30 mg/kg body weight/day) for 5 days. Treatment was rapidly followed by the clearance of the parasite in subsequent stool tests and the absence of gastrointestinal symptoms.

DISCUSSION

Two trichomonad species have been isolated from the human intestinal tract: *Dientamoeba fragilis* and *P hominis*. Only the former species is considered pathogenic. Regarding *P hominis*, its host range appears to be very wide since it colonises the gastrointestinal tract of a number of mammalian hosts. To our knowledge, there are only a handful of reports on *P hominis* mean
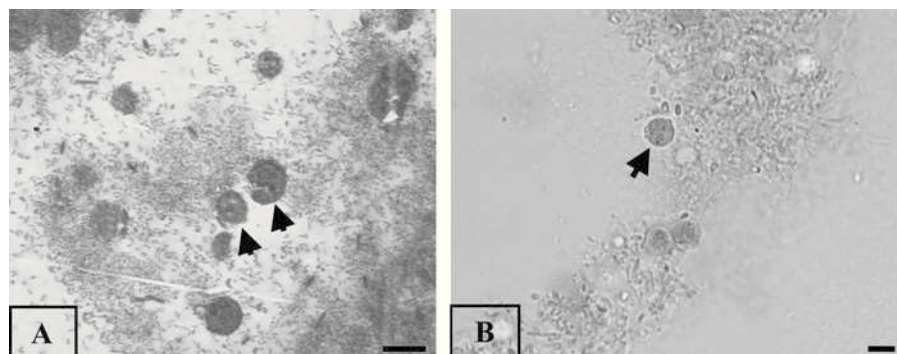

Figure 1 Cytological appearance of trichomonad cells in stool culture (case 1). (A) Merthiolate-iodine-formalin- and (B) RAL555-stained smears showing numerous trichomonad cells (arrows). Note the round shape of the micro-organisms. Typical microtubular cytoskeletal structures of trichomonads including flagella and axostyle—pelta complex are not visible. Bar=15 μm.

prevalence in humans. It does not exceed 1—1.5% in Australian Aboriginal communities[3] or in adolescent girls, pregnant women and commercial sex workers in Zambia.[4] It was suggested that this trichomonad is, in most instances, a harmless commensal organism of the digestive tract. In other mammals, several investigators have recognised a possible association between feline and canine large-bowel diarrhoea and *P hominis* infection.[1][2]

In order to identify the trichomonad species found in stool culture (case 1) or stool samples (case 2) from the two patients, the internal transcribed spacer 1 (ITS1)-5.8S rRNA-ITS2 region was amplified from DNA extracted from these samples as described elsewhere.[5] Following cloning of the PCR products, 10 clones were arbitrarily selected for each sample and sequenced on both strands. The clones from strains PM (case 1) and KGB (case 2) exhibited 98.3—100% sequence identity strongly suggesting that they were all derived from the same species (figure 2). These sequences also exhibited 98.9—99.6% identity to homologous sequences from *P hominis* strains available from different hosts. These results unequivocally identified trichomonads found in the digestive tract of the two patients as *P hominis* and suggested that the same *P hominis* species could possibly colonise many mammal species, including dogs, bovid and humans. Consequently, our data raised the crucial question of the as-yet-unknown zoonotic potential of *P hominis*, which is frequently found in pets. Other trichomonad species such as *Tritrichomonas foetus*[5] and *Tetratrichomonas* sp.[6] have an animal reservoir and are also able to colonise humans.

In the stool culture of the first patient, *P hominis* cells were observed exhibiting a round shape in merthiolate-iodine-formalin- and RAL555-stained smears (figure 1A,B). Typical microtubular structures of trichomonads (flagella, axostyle—pelta) were not visible, suggesting that the observed cells represented pseudocysts.[7] The main mode of transmission of *P hominis* is the likely

consumption of contaminated food and water through the standard faecal—oral route. Since this species does not form true cysts, vegetative cells have to survive for some time in the environment. In the experimental conditions of filter paper/slant cultures performed in the present work, trichomonads were growing in distilled water at 25°C and remained viable for at least 10 days, suggesting a relatively prolonged *P hominis* survival in water and, consequently, a significant potential of this species to use waterborne routes of transmission.

As described above, *P hominis* was the probable causative agent of gastrointestinal troubles in the child (case 2). Indeed, the child did not present other gastrointestinal pathogens and metronidazole treatment led to clearance of trichomonads and clinical recovery. By contrast, in the first patient, IBS was likely to be the main source of intestinal symptoms and probably represented favourable conditions for the development of *P hominis* infection as a secondary event, since we observed the clearance of trichomonad cells after IBS treatment without antiparasitic treatment. Until now, this trichomonad has been described as a potential aetiological agent of child diarrhoea[3][8] (as in case 2), and only rarely in adults. Evidence suggests that more children than adults appear to harbour *P hominis*, possibly due to children having less strict personal hygiene and the as-yet-unknown increased child susceptibility to the pathogenic effects of this parasite.

Dionigia Meloni,[1,2] **Cléa Mantini,**[1]
Julien Goustille,[3] **Guillaume Desoubeaux,**[3]
Zoha Maakaroun-Vermesse,[4]
Jacques Chandenier,[3] **Nausicaa Gantois,**[1]
Christophe Duboucher,[5] **Pier Luigi Fiori,**[2]
Eduardo Dei-Cas,[1] **Thanh Hai Duong,**[3]
Eric Viscogliosi[1]

[1]Center for Infection and Immunity of Lille, Institut Pasteur de Lille, Inserm U1019, CNRS UMR 8204, University Lille Nord de France, Lille, France; [2]Department of Biomedical Sciences, Division of Experimental and Clinical Microbiology, University of Sassari, Sassari, Italy; [3]Hôpital Bretonneau, Service de Parasitologie-Mycologie-Médecine Tropicale, Tours,
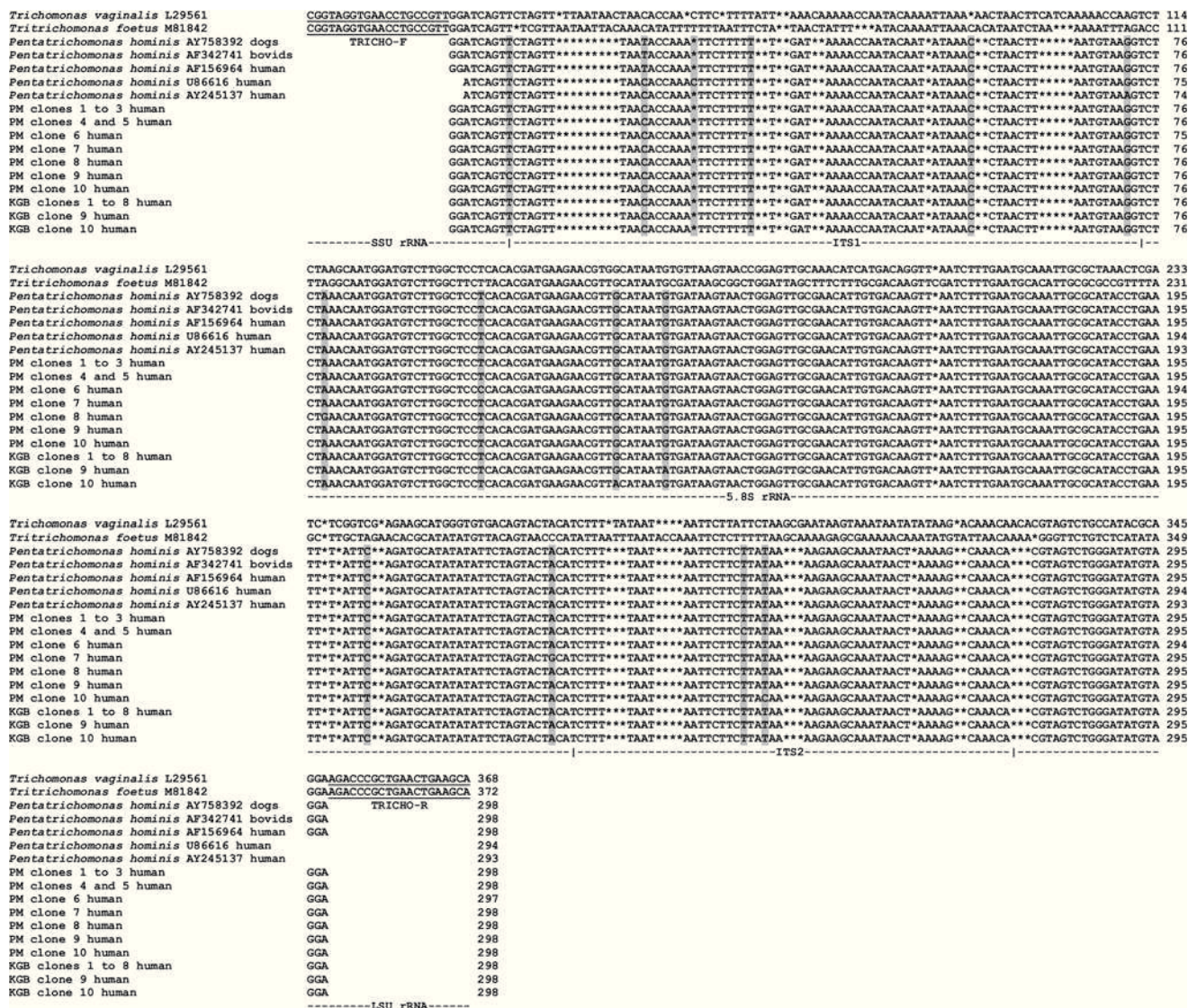
PostScript

Figure 2 Alignment of sequences of the ITS1-5.8S rRNA-ITS2 region of clones obtained in this study (PM and KGB strains from case 1 and case 2, respectively) and those of trichomonad species of interest found in humans (*Trichomonas vaginalis*) and animals (*Tritrichomonas foetus*). The low number of differences (up to four) observed between the clones of the same *Pentatrichomonas hominis* strain was likely due to the expected variation within the multiple copies of the RNA genes in any given genome. In the common part of our alignment (335 positions including gaps), the sequences of strains PM and KGB exhibited 74.8—75.4%, 59.6—60.2% and 98.9—99.6% identity to homologous sequences from *T vaginalis*, *T foetus* and *P hominis* strains isolated from dogs, bovid and humans. Differences between the clones obtained in this study and differences between these clones and the homologous sequences of *P hominis* strains isolated from different hosts are shaded (14 variable positions in the common part of the alignment). The nucleotide sequences obtained in this study have been deposited in GenBank under accession numbers HM853978 to HM853987. The sequences of the primers used in this study are underlined. Gaps are represented by asterisks. SSU, small subunit; LSU, large subunit.

Take-home messages

▶ *Pentatrichomonas hominis* is a possible zoonotic species with a significant potential of transmission by the waterborne route and could be the aetiological agent of gastrointestinal troubles in children.

▶ The difficulty in identifying pseudo-cysts of *P hominis* in stool samples could explain the underestimation of this protozoan infection in humans.

France; [4]Hôpital Clocheville, Service de Pédiatrie, Tours, France; [5]CHI de Poissy/Saint-Germain-en-Laye, Service de Pathologie, Saint-Germain-en-Laye, France

Correspondence to Eric Viscogliosi, Center for Infection and Immunity of Lille (CIIL), Institut Pasteur de Lille, Inserm U1019, CNRS UMR 8204, University Lille-Nord de France, Biology and Diversity of Emerging Eukaryotic Pathogens, 1 rue du Professeur Calmette, BP 245, 59019 Lille cedex, France; eric.viscogliosi@pasteur-lille.fr

Details have been removed from these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

REFERENCES

1. **Kim YA,** Kim HY, Cho SH, *et al*. PCR detection and molecular characterization of Pentatrichomonas

hominis from feces of dogs with diarrhea in the Republic of Korea. *Korean J Parasitol* 2010;**48**:9—13.

2. **Gookin JL,** Stauffer SH, Levy MG. Identification of *Pentatrichomonas hominis* in feline fecal samples by polymerase chain reaction assay. *Vet Parasitol* 2007;**145**:11—15.

3. **Meloni BP,** Thompson RC, Hopkins RM, et al. The prevalence of Giardia and other intestinal parasites in children, dogs and cats from aboriginal communities in the Kimberley. *Med J Aust* 1993;**158**:157—9.

4. **Crucitti T,** Jespers V, Mulenga C, et al. *Trichomonas vaginalis* is highly prevalent in adolescent girls, pregnant women, and commercial sex workers in Ndola, Zambia. *Sex Transm Dis* 2010;**37**:223—7.

5. **Duboucher C,** Caby S, Dufernez F, et al. Molecular identification of Tritrichomonas foetus-like organisms as coinfecting agents of human Pneumocystis pneumonia. *J Clin Microbiol* 2007;**44**:1165—8.

6. **Mantini C,** Souppart L, Noël C, et al. Molecular characterization of a new Tetratrichomonas species in a patient with empyema. *J Clin Microbiol* 2009;**47**:2336—9.

7. **Brugerolle G.** Cytologie ultrastructurale, systématique et evolution des Trichomonadida. *Ann Station Biologique de Besse-en-Chandesse* 1976;**10**:1—90.

8. **Yang CR,** Meng ZD, Wang X, et al. Diarrhoea surveillance in children aged under 5 years in a rural area of Hubei Province, China. *J Diar Dis Res* 1990;**8**:155—9.

Ochronosis and calcification in the mediastinal mass of a patient with alkaptonuria

Alkaptonuria (AKU) is a rare autosomal recessive condition with a prevalence of 1 in 250 000—1 000 000[1] caused by the deficiency of homogentisate 1,2 dioxygenase, the enzyme responsible for breaking down homogentisic acid (HGA).[2] Without this enzyme, levels of HGA within the tissues increase, even with highly efficient renal clearance of HGA. Urinary excretion of HGA is the only manifestation of AKU in children, and symptoms appear in adults due to joint damage. HGA has a high affinity for collagenous tissues, primarily the articular cartilages where it deposits as a pigmented polymer. Deposition also occurs at non-collagenous sites, suggesting that factors other than the presence of HGA in the tissues may favour or promote ochronosis.[2 3] Ochronotic pigment does not stain with Perl's reaction and can be highlighted in histological sections by Schmorl's stain.[4]

We report the case of a 51-year-old woman with shoulder joint arthritis who had been diagnosed previously with AKU elsewhere at the age of 50 years; the diagnosis was confirmed by urinary measurement of HGA. She attended the Liverpool AKU clinical evaluation unit where routine assessment by whole-body MRI revealed a 9 cm diameter, partly calcified mass in the anterior mediastinum (figure 1). In 1995, she was diagnosed with hyperthyroidism and had a reaction to carbimazole, requiring propylthiouracil and radioiodine therapy on three occasions. She became hypothyroid and was on thyroxine 175 μg daily. At the time of presentation, she was euthyroid with thyroid stimulating hormone and T4 levels within the normal ranges; autoantibodies were not measured. The anterior mediastinum was accessed by median sternotomy. The mass was located posterior to the manubrium and there were inflammatory adhesions between the mass and the posterior table of the sternum, the mediastinal pleura bilaterally (including the left phrenic nerve and the right lung), the pericardium and the left brachiocephalic vein. The mass was resected en bloc with the pericardium, mediastinal pleura and part of the right upper lobe of the lung. The patient made an uneventful recovery. The outer aspect of the specimen showed distinct areas of ochronotic pigmentation (figure 2). Dissection revealed a unilocular cyst containing yellow acellular material with a partly calcified wall showing pigmentation on the internal aspect of the cyst and the cut surfaces. Microscopically, the cyst was lined by mononuclear and multinucleate histiocytes, some containing ochronotic pigment. The fibrous wall of the cyst showed patchy lymphocytic infiltration and variably dense ochronotic pigmentation (figure 3). Granules of the ochronotic pigment were deposited
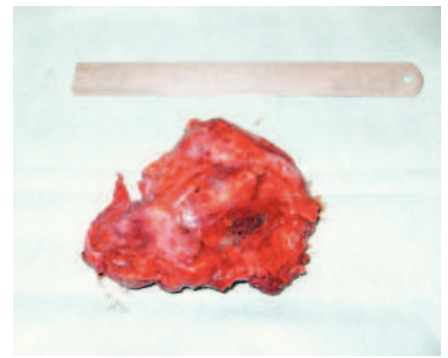

Figure 2 The excised mass shows patches of dense black ochronotic pigmentation. Adhesions between the mediastinal pleura bilaterally and the pericardium can be seen on the outer aspect of the mass.

between and surrounding matrix fibres, parallel to the matrix fibres in the tissues. In regions of dense pigmentation, numerous fibroblast-like cells contained cytoplasmic masses of pigment (figure 4).

This is the first report of ochronosis associated with a mediastinal cyst, or any form of cyst. Mediastinal cysts are uncommon and usually benign. Cysts arising in the anterior mediastinum are most commonly of thymic origin or mature cystic teratomas.[5] Although there are no particular features to indicate the nature of the cyst in this patient, the presence of ochronotic pigment in the cyst wall but not in other mediastinal tissues suggests that inflammation and fibrosis in the cyst wall have promoted pigment deposition. Thymic cysts have been observed following radiotherapy and exhibit areas of inflammation and fibrosis.[6] We have previously suggested that local tissue factors must promote the deposition of pigment in specific tissue.[3] The detection of calcified deposits within the mass closely associated with the ochronotic pigmentation suggests that polymeric HGA is not a homogeneous mix of the HGA
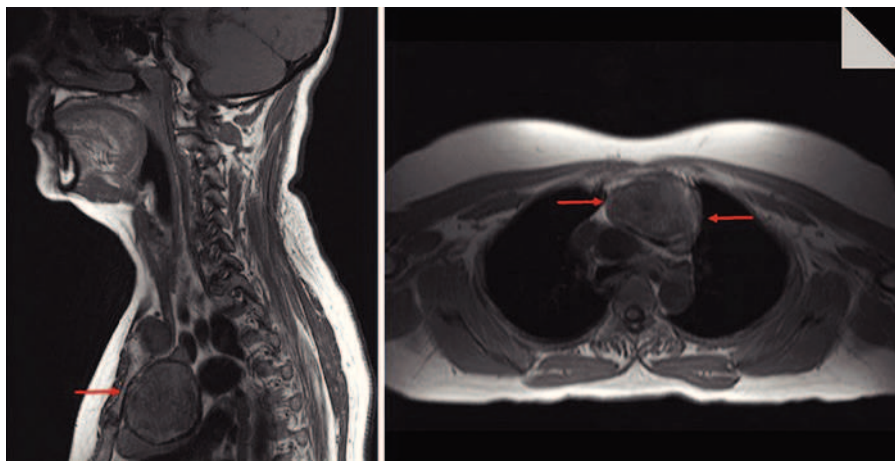

Figure 1 MRI scan of the thorax. Sagittal (left) and axial (right) T1 weighted images demonstrate a large anterior mediastinal cyst (arrows) in close relation to the great vessels.
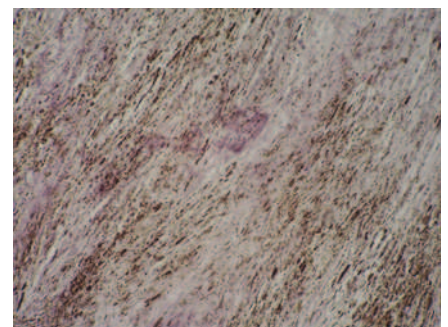

Figure 3 Granules of ochronotic pigment are present among collagen fibres. A range of sizes of pigment deposits can be seen, some of which appear to completely surround collagen fibres while others appear as individual deposits between the collagen fibres.

Molecular identification of *Pentatrichomonas hominis* in two patients with gastrointestinal symptoms

Dionigia Meloni, Cléa Mantini, Julien Goustille, et al.

Updated information and services can be found at:

http://jcp.bmj.com/content/64/10/933.full.html

These include:

References	This article cites 8 articles, 1 of which can be accessed free at: http://jcp.bmj.com/content/64/10/933.full.html#ref-list-1
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:

http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:

http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:

http://group.bmj.com/subscribe/

Part II

_____

EVOLUTION AND SYSTEMATICS OF PARABASALIDS

II-1 <u>Intoduction</u>

The SSU rRNA molecule is commonly used in the reconstruction of molecular phylogenies of Parabasalia. Recent molecular phylogenies inferred from sequences of this marker and including a large sampling of parabasalid taxa globally conflict with the systematic of this group of protozoa based on morphological characters. Consequently, Cepicka et al. (2010) have proposed a new taxonomy of parabasalids based on morphological and molecular data. In this classification, parabasalids are divided into six classes. Nevertheless, this revision does not solve all the problems with the systematic of parabasalids, and great uncertainty remains with respect to the phylogenetic relationships among and within these classes and to the root (most basal lineage) of this group. A lack of resolution and possible tree-construction artefacts have occurred because of a high level of sequence divergence of the SSU rRNA gene across parabasalid lineages. To overcome this drawback, approaches using the sequence data of multiple genes increase the number of informative characters for phylogenetic inference. In parallel, increasing the number of sampling taxa is also crucial. Consequently, to establish a reliable framework of the phylogeny of Parabasalia and to investigate the root of this group of protozoa, sequences of genes encoding actin, EF1a, GAPDH and SSU rRNA were obtained from a large sampling of species covering the large diversity of the parabasalids. We inferred the phylogenetic relationships among parabasalids and examined the root position based on the concatenation of the sequences of these markers.

II-2 <u>Results</u>

Article 3

Noda S, Mantini C, <u>Meloni D</u>, Inoue J-I, Kitade O, Viscogliosi E, Ohkuma M. Molecular phylogeny and evolution of Parabasalia with new protein markers and improved taxon sampling. PLoS ONE, 2011, in revision

II-3 <u>Conclusions</u>

In this study, the analysis of many important taxa with multi-gene sequences, greatly expand our understanding of the evolution of parabasalid biodiversity. For the first time, the concatenated dataset for a large number of common taxa outlines the robust relationships of most of the major parabasalid lineages and a more plausible new root position. Globally, our phylogenetic inference supports the new classification by Cepicka et al. (2010) even if we

propose uniting the Cristamonadea and Tritrichomonadea classes. In addition, the new root position observed in our analysis suggests that the ancient, most primitive parabasalid has a trichomonad-like character. This also indicates that the evolution of parabasalids is principally simple-to-complex, but the complexity has emerged independently in multiple lineages in different modes of flagellar system multiplication. In parallel, secondary reduction of cellular complexity seems to have occurred in several of the six classes.

Molecular phylogeny and evolution of Parabasalia with new protein markers and improved taxon sampling

Satoko Noda[1,2]*, Cléa Mantini[3], Dionigia Meloni[3,4], Jun-Ichi Inoue[1], Osamu Kitade[5], Eric Viscogliosi[3], and Moriya Ohkuma[1]*

[1]Microbe Division/Japan Collection of Microorganisms, RIKEN BioResource Center, Wako, Saitama 351-0198, Japan
[2]Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi 400-8510, Japan
[3]Center for Infection and Immunity of Lille, Institut Pasteur de Lille, 59019, Lille cedex, France, and Inserm U1019, CNRS UMR 8204, and University Lille – Nord de France, 59000 Lille cedex, France
[4]Department of Biomedical Sciences, Division of Experimental and Clinical Microbiology, University of Sassari, 07100 Sassari, Italy
[5]Natural History Laboratory, College of Science, Ibaraki University, Mito, Ibaraki 310-8512, Japan
*Corresponding authors.

Abstract

Background: Inferring the evolutionary history of phylogenetically isolated, deep-branching groups of taxa—in particular determining the root—is often extraordinarily difficult because their close relatives are unavailable as suitable outgroups. One of these taxonomic groups is the phylum Parabasalia, which comprises morphologically diverse species of flagellated protists of ecological, medical, and biological significance. Indeed, previous molecular phylogenetic analyses of members of this phylum have yielded conflicting and possibly erroneous inferences. Furthermore, many species of Parabasalia are symbionts in the gut of termites and therefore formidably difficult to cultivate, rendering available data insufficient. Increasing the numbers of examined taxa and informative characters (e.g., genes) is likely to produce more reliable inferences.

Principal Findings: Actin and elongation factor-1α genes were newly identified from 22 species of termite-gut symbionts through careful manipulations and seven cultured species, which covered major lineages of Parabasalia. Their protein sequences were concatenated and analyzed with sequences of previously and newly identified glyceraldehyde-3-phosphate dehydrogenase and small-subunit rRNA gene. For the first time, significant resolution of the phylogenetic relationships among major groups of Parabasalia and a more plausible new root position were obtained from this concatenated dataset.

Conclusions/Significance: We consider that increasing the number of sampled taxa as well as the addition of new sequences greatly improves the accuracy and robustness of the phylogenetic inference. A morphologically simple cell is likely the ancient form in Parabasalia as opposed to a cell with elaborate flagellar and cytoskeletal structures, which was defined as the most basal in previous inferences. Nevertheless, the evolution of Parabasalia is complex owing to several independent multiplication and simplification events in these structures. Therefore, systematics based solely on morphology can no longer reflect its true evolution.

Introduction

The phylum Parabasalia, or commonly parabasalids, comprises a monophyletic but complex assemblage of diverse species of flagellated protists typified by the presence of a characteristic parabasal apparatus (Golgi complex associated with striated fibers), closed mitosis with an external spindle (pleuromitosis), and anaerobic-energy-generating organelles called hydrogenosomes [1,2]. Based on morphological characters, more than 80 genera and 400 parabasalid species have been described thus far [1,3]. Most parabasalids inhabit the digestive tract of animal hosts as commensals, parasites, or symbionts. In particular, symbiotic parabasalids found in the gut of termites and wood-eating cockroaches play a central role in the digestion of cellulose [4]. This symbiotic relationship is considered a key element in the evolution of social behavior in the hosts [5] and has ecological significance for the decomposition of plant litter in terrestrial ecosystems [6]. Several parabasalids are also of considerable medical and veterinary importance as pathogens—i.e., *Trichomonas vaginalis* and *Tritrichomonas foetus* [7,8].

In addition to the functional host interactions of parabasalids, their biodiversity, conspicuous morphology, unique anaerobic biochemistry, and potentially crucial position in various schemes of eukaryotic evolution have long fascinated researchers. Indeed, Parabasalia is included in the supergroup Excavata [9] and in Metamonada with Fornicata (e.g., *Giardia*) and Preaxostyla (e.g., *Monocercomonoides* and *Trimastix*) [10]. The monophyly of Excavata or Metamonada is still intensely debated, but multigene analyses interestingly suggest that Excavata stems from a very deep branching event within the history of Eukaryotes [11–13].

Molecular phylogenetic studies based on small-subunit (SSU) rRNA gene sequences have mainly focused on cultured representatives of parabasalids [14–25]. In contrast, molecular studies of the parabasalids found in the gut of termites have been impeded because these organisms live in complex microbial communities and are very difficult to culture [26,27]. SSU rRNA gene sequences from parabasalid symbionts have, however, been subsequently obtained from the whole-gut communities of various termite species using polymerase chain reaction (PCR) [14,28–37]. Several of these sequences have been assigned to the corresponding termite symbionts in the mixed population of the gut community using whole-cell in situ hybridization with specific oligonucleotide probes [28,29,31,33,34,38–41]. Alternatively, sequence data have also been obtained by amplifying gene sequences from a small population of morphologically recognizable symbionts isolated under microscopy through micromanipulation [30, 33–48].

Because of a high level of sequence divergence of the SSU rRNA gene across parabasalid lineages, a lack of resolution and possible tree-construction artifacts have occurred, even though a large number of taxa have been investigated [24,35,39,46]. To overcome this drawback, multiple protein-encoding genes such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH), enolase, and tubulins have been examined [49–54]. In these phylogenetic reconstructions, however, taxon sampling was limited, and some conflicting results were obtained depending on the genes. Worse, in the case of enolase, paralogous copies resulting from ancient duplications might confound phylogenetic interpretation. Among these indicators, GAPDH sequences that contain a greater phylogenetic signal have given well-resolved trees largely congruent with the SSU rRNA gene phylogeny. Consequently, the GAPDH sequences have been gradually used in phylogenetic reconstructions, and the data of many important taxa have accumulated to cover a wide range of parabasalid diversity [34,46].

Taxonomic classifications of parabasalids have been proposed on the basis of marked morphological differences, particularly in the arrangement of the basal bodies of the flagella and associated cytoskeletal elements [1,55,56]. Parabasalids have historically been divided into two classes: Trichomonada and Hypermastigia [1]. Hypermastigia (or hypermastigids)

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

comprises species of typically large-cell forms equipped with numerous flagella, whereas Trichomonada cells are usually smaller and simpler than those of hypermastigids, with up to six flagella. Molecular studies provide critical and sometimes unexpected insights into the evolution of parabasalids and globally conflict with established systematics. For instance, hypermastigids have been considered to have a polyphyletic origin; therefore, many authors have pointed out the need to revise parabasalid systematics on the basis of these molecular data [16,17,19,38], and indeed some systematic revisions have been started [57,58]. Consequently, Cepicka *et al*. [24] have proposed dividing the parabasalids into six classes: Trichonymphea, Spirotrichonymphea, Cristamonadea, Tritrichomonadea, Hypotrichomonadea, and Trichomonadea. The traditional hypermastigids are assigned to the former three classes, which almost exclusively comprise species occurring in the gut of termites and wood-eating cockroaches. Nevertheless, these revisions do not solve all the problems with the systematics of parabasalids, and great uncertainty remains with respect to the phylogenetic relationships among and within these classes [20,22,24,33,39,46].

In phylogenetic analyses based on SSU rRNA gene or concatenated data of protein sequences, Trichonymphea, the most morphologically complex group of parabasalids, frequently occurs as the most basal lineage of Parabasalia [29–31,42,52,53]. Support for this rooting has often been poor, however, and the differences among several alternative positions in parabasalid lineages have been insignificant. This uncertainty has been explained by the absence of any close outgroup species to Parabasalia [19]. The parabasalid lineage usually branches out deeply in eukaryote phylogenies, and the branch leading to Parabasalia is very long. Such a long branch may attract the fast-evolving Trichonymphea, causing artificial rooting. Furthermore, the gene encoding GAPDH,like many other glycolytic enzymes, of Parabasalia has been derived from a bacterium *via* lateral gene transfer (LGT) [49,59,60], which disturbs root inference in comparisons to other eukaryotes.

Approaches using the sequence data of multiple genes increase the number of informative characters for phylogenetic inference. In parallel, increasing the number of sampling taxa (or species) is also important. Assembling datasets rich in both genes and taxa is likely to produce more accurate and robust results, although controversy exists about which strategy (increasing the number of genes or taxa) contributes most to the accurate inference [61–63]. In general, recent genome sequencing efforts have increased the number of genes for some species, but the genome analyses of either cultured microbial species or yet-uncultivated species are still poor. In terms of taxon sampling, data availability is also restricted, particularly for yet-uncultivated species. Therefore, to establish a reliable framework of the phylogenetic relationships among parabasalid groups as well as of the evolution of parabasalid biodiversity, we examined two additional genes encoding actin and elongation factor (EF)-1α in diverse parabasalid taxa of both cultivated and yet-uncultivated species. These two proteins are functionally independent from each other and from GAPDH. We inferred the phylogenetic relationships among parabasalids and examined the root position of this protistan group based on the concatenation of SSU rRNA, actin, EF-1α, or GAPDH gene sequences.

Results

Actin, EF-1α, and GAPDH sequences

In this study, sequences of genes encoding actin and EF-1α were determined from 29 species of parabasalids (Table 1). Among these species, 22 are symbionts in the gut of termites, whereas seven were cultured representatives from other animals. In each gene, several sequences showing more than 98% amino acid identity were obtained for most of these species. These differences can be explained by sequence variations of duplicated gene copies in the genome, at least in the cases of single-cell analyses, or by intra-species variations of

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

the gene in the population when multiple cells were used for the analysis. The GAPDH sequences were also obtained in two previously unstudied species, *Hoplonympha* sp. and *Staurojoenina assimilis*. From the latter species, two distinct sequences showing 89% amino acid identity were obtained as usual for the GAPDH gene in parabasalids [42, 45]. Maximum protein sequence differences between parabasalian species were 23%, 33%, and 45% for actin, EF-1α, and GAPDH, respectively.

Figure 1 shows the phylogenetic trees inferred from the sequences of the three proteins. The sequences of the multiple gene copies identified in this study as well as those found in the genome of *T. vaginalis* [64] showed close phylogenetic relatedness, and thus did not confound the phylogenetic reconstruction—at least the reconstruction needed to infer the relationships among major parabasalian groups, although some inter- or intra-species relationships seemed to be difficult to clarify. GAPDH yielded a relatively good resolution, with many of the branches receiving high statistical supports, whereas the actin and EF-1α trees were rather poorly resolved, particularly within Cristamonadea in both cases and within Trichonymphea in the case of actin. Each of Spirotrichonymphea, Trichonymphea, Cristamonadea, and Hypotrichomonadea was monophyletic, with significant supports for the former two in the three trees. Each of Tritrichomonadea and Trichomonadea was paraphyletic in both the actin and EF-1α trees. A remarkable difference among the three protein-based trees was the position of Spirotrichonymphea. Indeed, it was sister to Trichonymphea in the GAPDH tree and to Cristamonadea in the EF-1α tree (both were supported but only fairly), whereas it was sister to neither Trichonymphea nor Cristamonadea in the actin tree. Depending on the placement of Spirotrichonymphea, the relationships of the six parabasalian classes were considerably different among the three proteins, except that Cristamonadea and Tritrichomonadea were grouped together in the GAPDH and actin trees and the group of Trichomonadea and Trichonymphea was separated from the other classes in the actin and EF-1α trees. The results indicated that the analysis of single proteins gave poor resolution and would cause incorrect inferences of relationships.

Sequence concatenation and relationships among parabasalian classes
To overcome the poor resolution with single markers, one representative of the multiple gene copies in each of the 28 common parabasalian species was used for the sequence concatenation of the three proteins and SSU rRNA gene and subsequent analysis (Figure 2). A significant resolution—particularly at the level of higher taxonomic groups—was obtained from this concatenated dataset. The monophyly of each parabasalian class was completely supported except in Tritrichomonadea. In the apical part of the tree, Cristamonadea and Tritrichomonadea formed a well-supported clade, which formed a sister group with Spirotrichonymphea. These three classes were further grouped with Hypotrichomonadea, and the four were completely separated from Trichonymphea and Trichomonadea.

Alternative phylogenetic relationships of parabasalian classes were examined with the concatenated dataset using the Shimodaira-Hasegawa (SH) test, in which all possible pairs of the classes that did not appear in the tree of the concatenated dataset were compared (Table 2). The test did not reject the pairs from Cristamonadea, Tritrichomonadea, Spirotrichonymphea, and Hypotrichomonadea, suggesting that some ambiguities remained in the relationships of these four classes.

Congruency of the inferred relationships of the six parabasalian classes was examined using SH tests. The branching order of the six classes obtained by the sequence concatenation was not significantly worse in all the datasets of single protein or SSU rRNA gene ($P > 0.05$), whereas the relationships obtained with the GAPDH and SSU rRNA gene sequences were rejected in the concatenated dataset ($P = 0.006$ and 0.004, respectively), suggesting some phylogenetic noise in these two datasets. Nevertheless, the removal of either the GAPDH or the SSU rRNA gene sequences from the concatenated dataset did not

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

seriously change the overall relationship of the six classes except that Tritrichomonadea became paraphyletic in both cases and that Spirotrichonymphea and Hypotrichomonadea formed a sister group with only a weak support in the case of SSU rRNA gene sequence removal (see additional file 1: Figures S1 and S2).

Root of Parabasalia

The root of Parabasalia was investigated through analyses with outgroup taxa. Because the GAPDH gene of Parabasalia has likely been acquired from a bacterium via LGT [49], the sequence concatenation of EF-1α, actin, and SSU rRNA gene of 30 common parabasalian species was analyzed with representatives of diverse eukaryotic lineages as outgroup taxa (Figure 3). The root of Parabasalia was located at the position dividing the Trichonymphea plus Trichomonadea group from the others (position k in Figure 3). The monophyly of Parabasalia was fully supported. Except for the paraphyly of Tritrichomonadea, the relationships among the parabasalian classes were quite similar to those revealed in the analyses without outgroup taxa, although the supporting values of the branching orders of Cristamonadea, Tritrichomonadea, Spirotrichonymphea, and Hypotrichomonadea were decreased. Notably, each of the grouping of these four classes and the sister-group relationship between Trichomonadea and Trichonymphea was considerably supported (94/1 and 74/1, respectively), indicating the significance of the parabasalian root position.

Because some outgroup taxa were very long branch, which might violate the inference of the parabasalian root, these long-branch outgroup taxa were removed stepwise from the analyses (see Table S1). A series of analyses removing long-branch outgroup taxa from 23 to 13 did not substantially affect the root position, the relationships among the parabasalian classes, and their support values, indicating that these long-branch outgroup taxa did not disturb the inference.

We compared 11 possible root positions (a–k in Figure 3) using the SH test (Table 3). The root positions at the nodes leading to Trichonymphea (position f), Trichomonadea (g), and Hypotrichomonadea (e), the node dividing the group of these three classes from the others (j), and the clade of *Histomonas* plus *Dientamoeba* (b) were not rejected ($P > 0.05$). The results indicated that although the inferred root position k was the most likely, some ambiguities remained.

When GAPDH was included in the analysis of the sequence concatenation with a number of eukaryotic taxa as outgroups (data not shown), the identical root position was obtained with considerable support values. Because Preaxostyla such as *Trimastix* is the only known eukaryotic group that shares a common origin of the GAPDH gene of Parabasalia (LGT in their common ancestor or LGTs from closely related bacteria) [59,60], *Trimastix* was used as an outgroup taxon for parabasalian root analysis of the sequence concatenation including GAPDH (Figure 4). Again, the identical root position was inferred, and significant support values were obtained both at the node uniting Trichonymphea and Trichomonadea and at the node grouping the other four classes. The relationships among the six classes were identical to those shown in Figures 2 and 3.

Addition of α-tubulin and β-tubulin sequences to the concatenated dataset, which reduced the number of the parabasalian taxa available for the analysis to 12, demonstrated that Trichonymphea was the most basal class (see Figure S3), which is consistent with the conclusions of previous studies [24,53,54]. To investigate the effects of the number of parabasalian taxa on the inference of their root position, a fixed number of randomly chosen parabasalian taxa were excluded from the concatenated dataset of EF-1α, actin, and SSU rRNA gene sequences and analyzed repeatedly (Table 4). As the number of examined taxa became small, the root position gradually shifted to Trichonymphea, suggesting that the limited numbers of taxon sampling caused the wrong rooting at Trichonymphea.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Discussion

Importance of taxon sampling and new phylogenetic markers

The analyses of many important genera with multi-gene sequences in this study greatly expand our understanding of the evolution of parabasalian biodiversity. Actin and EF-1α sequences are newly obtained in all the examined taxa except for *Trichomonas* and *Pentatrichomonas*. For the first time, the concatenated dataset for 28 or 30 common taxa outlines the robust relationship of most of the major parabasalian groups and a more plausible new root position, thereby largely overcoming the problems encountered in previous molecular phylogenetic studies.

The increased number of parabasalian taxa sampled for the analyses is likely a key parameter in the improvement of Parabasalia rooting. A long-standing debate exists in phylogenetics about whether improved accuracy results from increasing the number of examined taxa (species) or the number of genes (informative characters) [61–63,65]. Studies of empirical data often emphasize the importance of the number of sampling taxa. Particularly, if a small number of taxa that tend to cause long-branch attraction are evaluated with a large number of characters, some slight systematic biases can become magnified and misinterpreted as phylogenetic signals and may cause unfortunately well-resolved, but wrong inferences [65]. In this situation, the addition of taxa in the analysis is an efficient approach. Indeed, multiple changes in an alignment site are more easily detected, and the model parameters for the inference are more precisely optimized when many taxa are analyzed [62,65].

Previous molecular phylogenetic studies have typically favored rooting at the branch leading to Trichonymphea. Hampl *et al*. [19], however, have considered this rooting as an artificially generated wrong inference, even though they did not suggest any robust alternative position. The taxa exclusion analyses described in this study (see Table 4) strongly support the importance of the number of sampling taxa for the inference of the parabasalian root. As the number of examined taxa was reduced, the root gradually shifted to Trichonymphea. The addition of the protein sequences of α- and β-tubulins to the concatenated dataset for the reduced number of taxa (12) unfortunately resulted in rooting at the branch leading Trichonymphea (additional file 1: Figure S3), probably because some systematic biases caused the wrong inference in the limited taxon sampling. Altogether, we consider that the increased number of sampling taxa in parabasalids improves the accuracy of their rooting, although a closely related species to parabasalids as a suitable outgroup is still lacking.

This study provides a significant resolution of the relationship of the major groups of parabasalids. Except for GAPDH, previously examined protein sequences (tubulins and enolase) have been demonstrated to generate only low levels of phylogenetic signals [52,53], and indeed, studies of the concatenation of these protein sequences have yielded conflicting results with poor resolution [24,53,54]. Each protein also demonstrated a poorly resolved and conflicting relationship of major parabasalian groups (see Figure 1). Therefore, the addition of the actin and EF-1α protein sequences to the sequence concatenation is important for resolution. Nevertheless, the phylogenetic relationship within some parabasalian classes such as Cristamonadea was still poorly resolved, probably owing to the limited phylogenetic information of actin and EF-1α sequences (maximum sequence differences of 17% and 15% between Cristamonadea members, respectively). Moreover, the possibility of several alternative phylogenetic relationships could not be rejected (see Table 2). Further study of other protein markers with sufficient taxon sampling is still needed. Recently, a conserved single-copy gene-encoding large subunit of RNA polymerase II has been suggested as a useful marker [66]; however, only cultured parabasalids have been investigated so far and our preliminary attempts at PCR amplification as well as a survey in our ongoing EST data found that the gene sequence was hardly obtained from termite-gut symbionts.

Phylogenetic relationships among parabasalids and their morphology

These improvements in molecular phylogeny of parabasalids provide us new insights about their evolutionary relationships. A salient point in the present study is the placement of Spirotrichonymphea, which branches out distantly from Trichonymphea and Cristamonadea in the tree based on the sequence concatenation, although the possibility of the sister-group relationship with the latter class cannot be excluded completely. In previous studies, Spirotrichonymphea has been ambiguously placed somewhere among parabasalids. The branch leading to Spirotrichonymphea is very long in the phylogenetic trees based on the SSU rRNA gene sequence, suggesting its artificial placement [19, 31, 33]. Some studies have shown its affinity to Cristamonadea, and others have located it in a more basal position, near Trichonymphea. In the present study, such conflicting results were obtained in the single-gene phylogenies (see Figure 1).

Historically, Spirotrichonymphea and Trichonymphea have been considered evolutionary closely related to each other because of the similarity of their morphogenesis. Indeed, in both groups, two symmetrical sets of basal bodies, basal fibers, and flagellar bands are separated equally at cell division and then completed in the sister cells [1]. In Spirotrichonymphea, however, flagella are uniquely arranged in left-handed spiral bands originating at the cell apex. Brugerolle [67] has emphasized some common ultrastructural features between Spirotrichonymphea and trichomonads, such as the organization of a privileged basal body (#2 in his description) bearing preaxostylar fibers connected to the pelta-axostyle junction. According to our molecular phylogeny based on the concatenated dataset, Spirotrichonymphea and Trichonymphea are neither sister nor sequentially branching lineages. Therefore, we suggest that their common characteristics in morphogenesis have evolved convergently.

In addition to Spirotrichonymphea and Trichonymphea, lophomonads (*Joenia*, *Joenoides*, *Joenina*, and *Deltotrichonympha*) included within Cristamonadea are parabasalids exhibiting a hypermastigid nature. In our trees, they are usually distantly related to Spirotrichonymphea and Trichonymphea. During the cell division, lophomonads retain only four privileged basal bodies, and the flagella are reconstructed in the daughter cells, whereas Spirotrichonymphea and Trichonymphea permanently maintain the multiflagellar state [1]. This feature specific to lophomonads supports their independent emergence from the other two groups. As shown in the present study and previously [46], these lophomonad genera branch out basally in Cristamonadea, whereas simpler devescovinids and multinucleated calonymphids likely emerged later. In addition, Tritrichomonadea, which is sister to Cristamonadea, contains the most rudimentary monocercomonads (*Histomonas* and *Dientamoeba*). Therefore, the apical group of parabasalids comprising Cristamonadea and Tritrichomonadea have undergone dynamic morphological transitions of multiplication and reduction of flagellar and cytoskeletal systems as well as transitions to multinucleated status.

In Trichonymphea, several morphological peculiarities have distinguished the families Hoplonymphidae (*Hoplonympha*) and Staurojoeninidae (*Staurojoenina*) from the other members of this class. Their flagellar areas are restricted to the anterior rostrum and form two and four symmetrical longitudinal rows in Hoplonymphidae and Staurojoeninidae, respectively [1,68–70]. In the other Trichonymphea members, flagella in the rostrum form a so-called rostral tube, which is composed of two half-round plates of parabasal fibers. Furthermore, the similarity of Hoplonymphidae and Staurojoeninidae with trichomonads or Spirotrichonymphea members has sometimes been argued [70–72]. In our study, relationships within Trichonymphea were fully resolved and Trichonymphea was divided into two robust groups (see Figures 2 and 3). Each of the two groups contains either Hoplonymphidae or Staurojoeninidae as the basal lineage. If the flagellar organization in the

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

limited number of longitudinal rows is primitive, the rostral tube and flagella in the pos-rostral area likely developed convergently in these two groups.

Our phylogenetic inference mostly supports the new classification proposed by Cepicka *et al*. [24]. Former classifications that elect a group corresponding to Cristamonadea as an order level [57,58] have resulted in the paraphyly or polyphyly of members outside Cristamonadea, Spirotrichonymphea, and Trichonymphea, as previous studies have repeatedly stated [19,24,39,46]. These members are now divided into the three classes—Tritrichomonadea, Hypotrichomonadea, and Trichomonadea—which comprise trichomonads and monocercomonads (members in the former families Trichomonadidae and Moncercomonadidae, respectively). These three classes clearly form distinct lineages in the trees reported in this study. Therefore, this reclassification indeed marks significant progress for the parabasalian systematics, presumably reflecting their true evolution. Nevertheless, some uncertainties remain unsolved. In our analyses, Tritrichomonadea was either monophyletic or paraphyletic at the base of Cristamonadea, but neither relationship was significantly supported. Because most systematists suppose that formal taxonomic units should be monophyletic, the revision of the taxonomic status of Tritrichomonadea (and also Cristamonadea) may be necessary as discussed previously [24]; for instance, these two classes can be united when more data for important taxa become available. Moreover, several newly proposed classes have not definitively received monophyletic confidence owing to the lack of robust molecular phylogeny, and molecular phylogenetics often denies morphology-based classifications from family to genus or species levels [36,46]. A large number of species examined with the SSU rRNA gene sequences, thus providing only a low level of phylogenetic resolution, have rarely been analyzed using protein sequences. These species are, for example, *Honigbergiella*, *Ditrichomonas*, *Lacusteria*, and *Hexamastix* (Trichomonadea) or *Simplicimonas* (Tritrichomonadea).

Evolutionary implications

The new root position that we have uncovered suggests that the ancient, most primitive parabasalid has a trichomonad-like character (Figure 5), although other possibilities cannot be excluded. Supporting this conclusion, all the possible alternative basal lineages of parabasalids (root positions b, e, f, and j in Figure 3; see also Table 3) are close to or representative of the classes including trichomonads except for the Trichonymphea lineage, which likely results from an incorrect inference as discussed above. Trichomonads in Tritrichomonadea, Hypotrichomonadea, and Trichomonadea are, in particular, specified by the presence of a costa and undulating membrane. Their recurrent flagellum is associated with the cell body, forming an undulating membrane underlain by a striated fiber, the costa. The structures of the costa and undulating membrane exhibit variations among these three classes. The costa has a similar banding pattern in members of Tritrichomonadea and Hypotrichomonadea (A-type striation according to Honigberg *et al*. [72]), whereas the costa of Trichomonadea shows a different banding pattern (B-type striation). The undulating membrane is rail-like (in *Tritrichomonas*) or lamelliform (in *Simplicimonas*) in Tritrichomonadea and lamelliform in both Hypotrichomonadea and Trichomonadea. The primitive parabasalid likely possessed the lamelliform undulating membrane, and the rail-like undulating membrane evolved later. The homologous protein components of both types of costa suggest their common origin [73]. The differentiation of the two types of costa probably occurred very early in parabasalid evolution; however, the A-type striation pattern may be primitive because it also occurs in the parabasal fibers of most parabasalids.

Based solely on comparison of morphological characters, the most ancient lineage of parabasalids has historically been believed to be monocercomonads owing to their cytoskeletal simplicity, and a more complex structure and morphology is considered to have developed later during parabasalian evolution [55,56]. Earlier molecular phylogenetic studies

have challenged this simple-to-complex evolutionary scheme, and have suggested almost the reverse scheme owing to the basal placement of the most elaborate Trichonymphea. The new root position described in this study indicates that the evolution of parabasalids is principally simple-to-complex, but the complexity has emerged independently in multiple lineages in different modes of flagellar system multiplication.

Possible occasional simplifications make the matter more intricate. Monocercomonads, the most rudimentary parabasalids, are polyphyletic, as clearly shown in the present study as well as previous work [16,18,22–24]. Indeed, some monocercomonads such as *Monocercomonas* and *Histomonas* are closely related to *Tritrichomonas*. The monocercomonad *Hypotrichomonas* forms a clade with *Trichomitus*. Other monocercomonads such as *Ditrichomonas* and *Honigbergiella* are likely related to Trichomonadea according to the analyses based on their SSU rRNA gene sequences [16,17,22]. Therefore, secondary reduction of cellular complexity seems to have occurred in each of the three classes—Tritrichomonadea, Hypotrichomonadea, and Trichomonadea (see Figure 5).

A growing number of free-living parabasalids have been recently identified in *Honigbergiella*, *Lacusteria*, and *Pseudotrichomonas* in addition to *Ditrochomonas* and *Monotrichomonas* [24,25], although phylogenetic positions of these free-living species have not been examined any protein sequence. The free-living species are seemingly dispersed in SSU rRNA gene-based phylogenetic trees but many form a paraphyletic assemblage near the origin of Trichomonadea [25]. Considering the new root position of Parabasalia uncovered in this study, it is possible that free-living species represent the most basal lineages of Parabasalia. This possibility, though needs to be examined in future, is of ecological and evolutionary significance for the origin of parabasalids as well as parasitic trichomonads as discussed previously [25].

The hypermastigid nature of flagellar multiplication in a single mastigont has independently evolved three times, and each multiplication has led to Trichonymphea, Spirotrichonymphea, and lophomonads in Cristamonadea (see Figure 5). Because these three classes are found only in the gut of termites and phylogenetically related wood-feeding cockroaches, each ancestor of these three classes (presumably trichomonad-like species) established a symbiotic relationship with an ancestor of these hosts and evolved and diversified in the gut of their hosts as previously described for Trichonymphea members [34]. Adaptation to the gut environment significantly affects their evolution. Indeed, the development of their flagella and associated cytoskeletal system is advantageous for the improvement of fitness in this niche, because it allows vigorous movement that prevents their flowing out of the gut and facilitates their access to food in the gut. Their habitats also likely affect cell size, which associates with cytoskeletal development. The cell of the gut symbionts is large enough to incorporate masticated wood particles by phagocytosis, whereas parasitic and free-living species absorb smaller molecules or tiny bacterial cells as food. Meanwhile, members of the genera *Pseudotrypanosoma* and *Trichomitopsis* in Trichomonadea are also found among the gut symbionts of termites, and their large cell sizes seem to represent a consequence of their adaptation to the gut environment, although they apparently do not develop a flagellar system [30,43].

Conclusions

This study provides a more reliable phylogenetic framework for Parabasalia. We consider both the increasing number of taxa sampled and the use of new protein markers as particularly important factors in the accuracy and robustness of our inferences. Culture-independent analyses of the termite-gut symbionts are critical for collecting data for the large number of examined taxa, and such a technique is powerful for the investigation of additional data and taxa. The evolution of Parabasalia is complex in terms of morphology owing to a

number of independent multiplications and simplifications of flagella and associated cytoskeletal structures. Morphology-based systematics has sometimes hampered the understanding of the true nature of parabasalian evolution. Likely, their ecology greatly affects evolution through adaptation to the niches and co-diversification with their hosts.

Materials and Methods

Cultivation and DNA extraction of trichomonads

Culturable strains used were as follows: *Tetratrichomonas gallinarum* strain A6 (cf. [63]); *T. foetus* strain KV1 (ATCC 30924); *Monocercomonas* sp. strain NS-1PRR (ATCC 50210); *Trichomitus batrachorum* strain G11 (ATCC 30066); *Hypotrichomonas acosta* strain L3 (ATCC 30069). The origins of their isolation are shown in Table 3. All strains were grown axenically at 37 °C or 27 °C in trypticase-yeast extract-maltose (TYM) medium [74] without agar supplemented with 10% (v/v) heat-inactivated horse serum (Gibco-BRL), 100 U/mL of penicillin G, and 50 μg/mL of streptomycin sulfate. Genomic DNA was isolated as described [75]. DNAs of *Dientamoeba fragilis* strain Bi/PA (ATCC 30948) and *Histomonas meleagridis* strain HmzL were provided by C. G. Clark (London School of Hygiene and Tropical Medicine, London, UK) and F. Delbac (LMGE, CNRS UMR 6023, Aubière, France), respectively.

Manipulation of termite symbionts

Table 1 lists the gut symbionts investigated in this study and their host termites. All taxa were stably found in the hindgut flora of the respective termites and were easily recognizable on the basis of their morphological characters [1,3]. The cells of parabasalian symbionts in the hindgut suspension of each termite were manually isolated and washed extensively under a microscope equipped with a micromanipulator (Cell Tram, Eppendorf) as described elsewhere [76,77]. A single cell or a pool of 10-30 cells showing typical morphology were isolated and used as templates for reverse transcriptase (RT)-PCR. In the cases of *Stephanonympha* sp., *Metadevescovina cuspidata*, and *Gigantomonas herculea*, the isolated cells were subjected to isothermal whole-genome amplification as previously described [78,79], and the amplified genome DNA was used as a template for PCR.

Cloning and sequencing of actin, EF-1α, and GAPDH genes

The protein encoding genes were amplified using RT-PCR from the isolated cells of termite-gut symbionts using protein-specific primers for the amino-terminal conserved region and the oligo-dT primer as previously described [53]. The following protein-specific primers were used: actin-F1, 5′-TGGGANGANATRGARAARATYTGG3′ and EF1F1, 5′-AARGCDGARCGNGARCGDGG-3′. The protein-specific primers for the carboxy-terminal conserved region used for PCR were actin-R1, 5′-GAAGCAYTTNCKRTGNACDAT-3′ and EF1R1, 5′-GRAAYTTRCANGCDATRTG-3′. If sufficient amplification product was not obtained during the first PCR, a second amplification was performed using the primer actin-F2, 5′-ATRGARAARATYTGGCAYCA-3′ and the oligo-dT primer for the actin gene, and the primer EF1F2, 5′-CGDGGDATYACNATYGAYAT-3′ and the EF1R1 or oligo-dT primer for the EF-1α gene. The GAPDH gene was amplified with RT-PCR using previously designed primers [53]. The amplification products were separated using agarose gel electrophoresis, purified, and cloned into pCR2.1-TOPO (Invitrogen). Clones containing inserts of the expected size were picked and partially sequenced, and the complete DNA sequence of each representative clone was obtained via primer walking. The sequences obtained in this study have been deposited in the DNA Databank of Japan and the accession numbers are shown in Figure 1.

Phylogenetic analyses

The protein sequences identified in this study and publicly available were used for the analyses. The genome sequence of *T. vaginalis* G3 [64] was searched for homologous sequences with the three proteins. The EST sequences of *Pentatrichomonas hominis* in the public database (FL516063–FL518016) were also searched, and the identified ESTs were assembled to produce *in silico*-translated amino acid sequences. All the non-identical amino acid sequences showing significant identities to sequences of other parabasalids (such as >50% identity) and covering most of the protein region were included in the analyses. In the case of *P. hominis*, only a single unique sequence was found in each protein. The recently reported GAPDH sequences of *H. meleagridis* [54] were also used.

These protein sequences were aligned using Clustal W2 [80] and refined manually. Only unambiguously aligned positions were used for phylogenetic inference. Because any outgroup sequences may cause long-branch attraction owing to their distant relationships to parabasalids, the phylogenetic tree of a single protein was inferred without outgroups. The appropriate model of sequence evolution was selected using the program ProtTest 2.4 [81]. For the tree of each single protein dataset, maximum likelihood (ML) analysis was carried out with RAxML 7.2.6 [82] using the PROTGAMMAWAG model.

The alignments of the protein sequences and previously aligned sequences of the SSU rRNA gene [46] were concatenated manually. When the analyses included GAPDH, the taxa *Deltotrichonympha* sp. and *D. fragilis* were excluded because of the lack of their GAPDH sequences. The outgroup taxa and their sequences used for the analyses are shown in Table S2. The sequence alignments used for the analyses shown in Figures 2, 3, and 4 are available as supplementary data.

For the analyses of only parabasalids, the ML tree was estimated in RAxML using mixed models (GTRGAMMA for the SSU rRNA gene and PROTGAMMAWAG for each protein sequence). Parameters and branch length were optimized for each of the partitions individually and bootstrap values were obtained from 1000 replicates. The ML estimation and the bootstrap were also conducted with the site-heterogeneous CAT model in RAxML. Bayesian analysis was performed in MrBayes 3.1.2 [83] using separate models (GTR+I+Γ for the SSU rRNA gene and WAG+I+Γ for each protein sequence), and parameters and branch lengths were optimized for each partition individually. The starting tree was random, and four simultaneous Markov chains in duplicate were run for 1,000,000 generations. Log likelihoods stabilized well before 100,000 generations, and the remaining generations were used to measure Bayesian posterior probabilities. The homogeneity of sequence composition in each protein or gene was evaluated in the χ^2 test implemented in TREE-PUZZLE 5.2 [84].

For the analyses with outgroup taxa, the ML tree was estimated as described above but using the CAT model (GTRMIX and PROTMIXWAG in RAxML) instead of the site-homogeneous model, because the site-heterogeneous CAT model appears to be more robust than site-homogeneous models against artifacts by long-branch attraction [85]. Bootstrap analyses of 1000 replicates were conducted with the CAT model in RAxML. Bayesian analysis was performed in MrBayes as described above.

Differences in alternative tree topologies were compared with the SH test implemented in CONSEL [86] using the site-wise log-likelihood outputs obtained with RAxML. Using only the data from parabasalids, an alternative tree topology was obtained under the constraint of a given phylogenetic hypothesis by the RAxML analysis with the same substitution model described above. For the root of parabasalids, the outgroup was grafted onto 11 possible root positions of parabasalids and tree topology was obtained under each constraint, and these root positions were evaluated using the SH test. To measure the effect of the number of analyzed taxa on the parabasalian root position, randomly chosen parabasalian taxa were excluded and analyzed repeatedly with different sets of fixed numbers of excluded taxa.

Acknowledgments

Authors Contributions

Conceived and designed the experiments: SN EV MO. Performed the experiments: SN CM DM JI OK. Analyzed the data: SN JI MO. Wrote the paper: SN EV MO.

References

1. Brugerolle G, Lee JJ (2001) Phylum Parabasalia. In: Lee JJ, Leedale GF, Patterson DJ, Bradbury PC, editors. An illustrated guide to the protozoa. 2nd edition. Lawrence, KS: Society of Protozoologists. pp. 1196–1250.
2. Honigberg BM, Brugerolle G (1990) Structure. In: Honigberg BM, editor. Trichomonads parasitic in humans. Heidelberg: Springer-Verlag. pp. 2–35.
3. Yamin MA (1979) Flagellates of the orders Trichomonadida Kirby, Oxymonadida Grassé, and Hypermastigida Grassi & Foà reported from lower termites (Isoptera families Mastotermitidae, Kalotermitidae, Hodotermitidae, Termopsidae, Rhinotermitidae, and Serritermitidae) and from the wood-feeding roach *Cryptocercus* (Dictyoptera: Cryptocercidae). Sociobiology 4: 5–119.
4. Brune A, Ohkuma M (2011) Role of the termite gut microbiota in symbiotic digestion. In: Bignell DE, Roisin Y, Lo N, editors. Biology of termites: a modern synthesis. Dordrecht: Springer. pp. 439–475.
5. Nalepa CA, Bignell DE, Bandi C (2001) Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. Insectes Sociaux 48: 194–201.
6. Ohkuma M (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl Microbiol Biotechnol 61: 1–9.
7. Honigberg BM (1978) Trichomonads of veterinary importance. In: Krieir JP, editor. Parasitic protozoa. Volume 2. New York: Academic Press. pp. 163–173.
8. Honigberg BM (1978) Trichomonads of importance in human medicine. In: Krieir JP, editor. Parasitic protozoa. Volume 2. New York: Academic Press. pp. 275–454.
9. Simpson AG (2003) Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota). Int J Syst Evol Microbiol 53: 1759–1777.
10. Cavalier-Smith T (2003) The excavate protozoan phyla Metamonada Grasse emend. (Anaeromonadea, Parabasalia, Carpediemonas, Eopharyngia) and Loukozoa emend. (Jakobea, Malawimonas): their evolutionary affinities and new higher taxa. Int J Syst Evol Microbiol 53: 1741–1758.
11. Hampl V, Horner DS, Dyal P, Kulda J, Flegr J, et al (2005) Inference of the phylogenetic position of oxymonads based on nine genes: support for metamonada and excavata. Mol Biol Evol 22: 2508–2518.
12. Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, et al (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci U S A 106: 3859–3864.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

13. Simpson AG, Inagaki Y, Roger AJ (2006) Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes. Mol Biol Evol 23: 615–625.

14. Gunderson J, Hinkle G, Leipe D, Morrison HG, Stickel SK, et al (1995) Phylogeny of trichomonads inferred from small-subunit rRNA sequences. J Eukaryot Microbiol 42: 411–415.

15. Silberman JD, Clark CG, Sogin ML (1996) *Dientamoeba fragilis* shares a recent common evolutionary history with the trichomonads. Mol Biochem Parasitol 76: 311–314.

16. Edgcomb V, Viscogliosi E, Simpson AGB, Delgado-Viscogliosi P, Roger AJ, et al (1998) New insights into the phylogeny of trichomonads inferred from small subunit rRNA sequences. Protist 149: 359–366.

17. Delgado-Viscogliosi P, Viscogliosi E, Gerbod D, Kulda J, Sogin ML, et al (2000) Molecular phylogeny of parabasalids based on small subunit rRNA sequences, with emphasis on the Trichomonadinae subfamily. J Eukaryot Microbiol 47: 70–75.

18. Gerbod D, Edgcomb VP, Noël C, Zenner L, Wintjens R, et al (2001) Phylogenetic position of the trichomonad parasite of turkeys, *Histomonas meleagridis* (Smith) Tyzzer, inferred from small subunit rRNA sequence. J Eukaryot Microbiol 48: 498–504.

19. Hampl V, Cepicka I, Flegr J, Tachezy J, Kulda J (2004) Critical analysis of the topology and rooting of the parabasalian 16S rRNA tree. Mol Phylogenet Evol 32: 711–723.

20. Hampl V, Vrlik M, Cepicka I, Pecka Z, Kulda J, et al (2006) Affiliation of *Cochlosoma* to trichomonads confirmed by phylogenetic analysis of the small-subunit rRNA gene and a new family concept of the order Trichomonadida. Int J Syst Evol Microbiol 56: 305–312.

21. Cepicka I, Hampl V, Kulda J, Flegr J (2006) New evolutionary lineages, unexpected diversity, and host specificity in the parabasalid genus *Tetratrichomonas*. Mol Phylogenet Evol 39: 542–551.

22. Hampl V, Cepicka I, Flegr J, Tachezy J, Kulda J (2007) Morphological and molecular diversity of the monocercomonadid genera *Monocercomonas*, *Hexamastix*, and *Honigbergiella* gen. nov. Protist 158: 365–383.

23. Mantini C, Dalia-Cornette J, Noda S, Van Der Heijden HM, Capron M, et al (2009) Molecular identification and phylogenetic relationships of trichomonad isolates of galliform birds inferred from nuclear small subunit rRNA gene sequences. Parasitol Res 106: 163–170.

24. Cepicka I, Hampl V, Kulda J (2010) Critical taxonomic revision of Parabasalids with description of one new genus and three new species. Protist 161: 400–433.

25. Yubuki N, Céza V, Cepicka I, Yabuki A, Inagaki Y, et al (2010) Cryptic diversity of free-living parabasalids, *Pseudotrichomonas keilini* and *Lacusteria cypriaca* n. g., n. sp., as inferred from small subunit rDNA sequences. J Eukaryot Microbiol 57: 554–561.

26. Ohkuma M (2008) Symbioses of flagellates and prokaryotes in the gut of lower termites. Trends in Microbiology 16: 345–352.

27. Ohkuma M, Brune A (2011) Diversity, structure, and evolution of the termite gut microbial community. In: Bignell DE, Roisin Y, Lo N, editors. Biology of termites: a modern synthesis. Dordrecht: Springer. pp. 413–438.

28. Berchtold M, König H (1995) Phylogenetic position of the two uncultivated trichomonads *Pentatrichomonoides scroa* Kirby and *Metadevescovina extranea* Kirby from the hindgut of the termite *Mastotermes darwiniensis* Froggatt. Syst Appl Microbiol 18: 567–573.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

29. Ohkuma M, Ohtoko K, Grunau C, Moriya S, Kudo T (1998) Phylogenetic identification of the symbiotic hypermastigote *Trichonympha agilis* in the hindgut of the termite *Reticulitermes speratus* based on small-subunit rRNA sequence. J Eukaryot Microbiol 45: 439–444.

30. Keeling PJ, Poulsen N, McFadden G (1998) Phylogenetic diversity of parabasalian symbionts from termites, including the phylogenetic position of *Pseudotrypanosoma* and *Trichonympha*. J Eukaryot Microbiol 45: 643–650.

31. Ohkuma M, Ohtoko K, Iida T, Tokura M, Moriya S, et al (2000) Phylogenetic identification of hypermastigotes, *Pseudotrichonympha*, *Spirotrichonympha*, *Holomastigotoides*, and parabasalian symbionts in the hindgut of termites. J Eukaryot Microbiol 47: 249–259.

32. Gerbod D, Edgcomb VP, Noël C, Delgado-Viscogliosi P, Viscogliosi E (2000) Phylogenetic position of parabasalid symbionts from the termite *Calotermes flavicollis* based on small subunit rRNA sequences. Int Microbiol 3: 165–172.

33. Ohkuma M, Iida T, Ohtoko K, Yuzawa H, Noda S, et al (2005) Molecular phylogeny of parabasalids inferred from small subunit rRNA sequences, with emphasis on the Hypermastigea. Mol Phylogenet Evol 35: 646–655.

34. Ohkuma M, Noda S, Hongoh Y, Nalepa CA, Inoue T (2009) Inheritance and diversification of symbiotic trichonymphid flagellates from a common ancestor of termites and the cockroach *Cryptocercus*. Proc Biol Sci 276: 239–245.

35. Carpenter KJ, Chow L, Keeling PJ (2009) Morphology, phylogeny, and diversity of *Trichonympha* (Parabasalia: Hypermastigida) of the wood-feeding cockroach *Cryptocercus punctulatus*. J Eukaryot Microbiol 56: 305–313.

36. Harper JT, Gile GH, James ER, Carpenter KJ, Keeling PJ (2009) The inadequacy of morphology for species and genus delineation in microbial eukaryotes: an example from the parabasalian termite symbiont *Coronympha*. PLoS One 4: e6577.

37. Strassert JFH, Desai MS, Brune A, Radek R (2009) The true diversity of devescovinid flagellates in the termite *Incisitermes marginipennis*. Protist 160: 522–535.

38. Gerbod D, Noël C, Dolan MF, Edgcomb VP, Kitade O, et al (2002) Molecular phylogeny of parabasalids inferred from small subunit rRNA sequences, with emphasis on the Devescovinidae and Calonymphidae (Trichomonadea). Mol Phylogenet Evol 25: 545–556.

39. Noël C, Noda S, Mantini C, Dolan MF, Moriya S, et al (2007) Molecular phylogenetic position of the genera *Stephanonympha* and *Caduceia* (Parabasalia) inferred from nuclear small subunit rRNA gene sequences. J Eukaryot Microbiol 54: 93–99.

40. Ikeda-Ohtsubo W, Desai M, Stingl U, Brune A (2007) Phylogenetic diversity of 'Endomicrobia' and their specific affiliation with termite gut flagellates. Microbiology 153: 3458–3465.

41. Noda S, Hongoh H, Sato T, Ohkuma M (2009) Complex coevolutionary history of symbiotic Bacteroidales bacteria of various protists in the gut of termites. BMC Evol Biol 9: 158.

42. Dacks JB, Redfield RJ (1998) Phylogenetic placement of *Trichonympha*. *J Eukaryot Microbiol* 45: 445–447.

43. Keeling PJ (2002) Molecular phylogenetic position of *Trichomitopsis termopsidis* (Parabasalia) and evidence for the Trichomitopsiinae. Eur J Protistol 38: 279–286.

44. Carpenter KJ, Keeling PJ (2007) Morphology and phylogenetic position of *Eucomonympha imla* (Parabasalia: Hypermastigida). J Eukaryot Microbiol 54: 325–332.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

45. Noda S, Kitade O, Inoue T, Kawai M, Kanuka M et al (2007) Cospeciation in the triplex symbiosis of termite gut protists (*Pseudotrichonympha* spp.), their hosts, and their bacterial endosymbionts. Mol Ecol 16: 1257–1266.

46. Noda S, Mantini C, Bordereau C, Kitade O, Dolan MF, et al (2009) Molecular phylogeny of parabasalids with emphasis on the order Cristamonadida and its complex morphological evolution. Mol Phylogenet Evol 52: 217–224.

47. Carpenter KJ, Horak A, Keeling PJ (2010) Phylogenetic position and morphology of Spirotrichosomidae (Parabasalia): new evidence from *Leptospironympha* of *Cryptocercus punctulatus*. Protist 161: 122–132.

48. Desai MS, Strassert JFH, Meuser K, Hertel H, Ikeda-Ohtsubo W et al (2010) Strict cospeciation of devescovinid flagellates and Bacteroidales ectosymbionts in the gut of dry-wood termites (Kalotermitidae). Environ Microbiol 12: 2120–2132.

49. Viscogliosi E, Müller M (1998) Phylogenetic relationships of the glycolytic enzyme, glyceraldehydes-3-phosphate dehydrogenase, from parabasalid flagellates. J Mol Evol 47: 190–199.

50. Keeling PJ, Palmer JD (2000) Parabasalian flagellates are ancient eukaryotes. Nature 405: 635–637.

51. Keeling PJ (2004) Polymorphic insertions and deletions in parabasalian enolase genes. J Mol Evol 58: 550–556.

52. Gerbod D, Sanders EP, Moriya S, Noël C, Takasu H, et al (2004) Molecular phylogenies of Parabasalia inferred from four protein genes and comparison with rRNA trees. Mol Phylogenet Evol 31: 572–580.

53. Ohkuma M, Saita K, Inoue T, Kudo T (2007) Comparison of four protein phylogeny of parabasalian symbionts in termite guts. Mol Phylogenet Evol 42: 847–853.

54. Hauck R, Hafez HM (2010) Systematic position of *Histomonas meleagridis* based on four protein genes. J Parasitol 96: 396–400.

55. Brugerolle G (1976) Cytologie ultrastructurale, systématique et évolution des Trichomonadida. Ann Stat Biol de Besse-en-Chandesse 10: 1–90.

56. Honigberg BM (1963) Evolutionary and systematic relationships in the flagellate order Trichomonadida Kirby. J Protozool 10: 20–63.

57. Brugerolle G, Patterson DJ (2001) Ultrastructure of Joenina pulchella Grassi, 1917 (Protista, Parabasalia), a reassessment of evolutionary trends in the parabasalids, and a new order Cristamonadida for devescovinid, calonymphid and lophomonad flagellates. Org Divers Evol 1: 147–160.

58. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, et al (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52: 399–451.

59. Liapounova NA, Hampl V, Gordon PMK, Sensen CW, Gedamu L, et al (2006) Reconstructing the mosaic glycolytic pathway of the anaerobic eukaryote *Monocercomonoides*. Eukaryot Cell 5: 2138-2146.

60. Stechmann A, Baumgartner M, Silberman JD, Roger AJ (2006) The glycolytic pathway of *Trimastix pyriformis* is an evolutionary mosaic. BMC Evol Biol 6: 101.

61. Delsuc F, Brinkmann H, Philippe H (2005) Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6: 361–375.

62. Hillis DM, Pollock DD, McGuire JA, Zwickl DJ (2003) Is sparse taxon sampling a problem for phylogenetic inference? Syst Biol 52: 124–126.

63. Rosenberg MS, Kumar S (2003) Taxon sampling, bioinformatics, and phylogenomics. Syst Biol 52: 119–124.

64. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, et al. (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315: 207–212.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

65. Hedtke SM, Townsend TM, Hillis DM (2006) Resolution of phylogenetic conflict in large data sets by increased taxon sampling. Syst Biol 55: 522–529.

66. Malik S-B, Brochu CD, Bilic I, Yuan J, Hess M, et al (2011) Phylogeny of parasitic Parabasalia and free-living relatives inferred from conventional markers vs. *Rpb1*, a single-copy gene. PLoS One 6: e20774.

67. Brugerolle G (2001) Morphological characters of spirotrichonymphids: *Microjoenia*, *Spirotrichonymphella* and *Spirotrichonympha* symbionts of the Australian termite *Porotermes grandis*. Eur J Protistol 37: 103–117.

68. Dolan M, Margulis L (1997) *Staurojoenina* and other symbionts in *Neotermes* from San Salvador Island, Bahamas. Symbiosis 22: 229–239.

69. Hollande A (1986) Les hypermastigines du genre *Staurojoenina*. Précisions sur la structure de leur rostre et description d'espèces nouvelles. Protistologica 22: 331–338.

70. Hollande A, Carruette-Valentin J (1971) Les atractophores, l'induction du fuseau et la division cellulaire chez les hypermastigines. Etude infrastructurale et révision systématique des Trichonymphines et des Spirotrichonymphines. Protistologica 7: 5–100.

71. Kirby H (1947) Flagellate and host relationships of trichomonad flagellates. J Parasitol 33: 214–228.

72. Honigberg BM, Mattern CF, Daniel WA (1971) Fine structure of the mastigont system in *Tritrichomonas foetus* (Riedmuller). J Protozool 18: 183–198.

73. Viscogliosi E, Brugerolle G (1994) Striated fibers in trichomonads: costa proteins represent a new class of proteins forming striated roots. Cell Motil Cytoskeleton 29: 82–93.

74. Diamond LS (1957) The establishment of various trichomonads of animals and man in axenic cultures. J Parasitol 43: 488–490.

75. Viscogliosi E, Philippe H, Baroin A, Perasso R, Brugerolle G (1993) Phylogeny of trichomonads based on partial sequences of large subunit rRNA and on cladistic analysis of morphological data. J Eukaryot Microbiol 40: 411–421.

76. Noda S, Iida T, Kitade O, Nakajima H, Kudo T, et al (2005) Endosymbiotic *Bacteroidales* bacteria of the flagellated protist *Pseudotrichonympha grassii* in the gut of the termite *Coptotermes formosanus*. Appl Environ Microbiol 71: 8811–8817.

77. Noda S, Kawai M, Nakajima H, Kudo T, Ohkuma M (2006) Identification and *in situ* detection of two lineages of *Bacteroidales* ectosymbionts associated with a termite gut protist, *Oxymonas* sp. Microbes Environ 21: 16–22.

78. Hongoh Y, Sato T, Dolan MF, Noda S, Ui S, et al (2007) The motility symbiont of the termite gut flagellate *Caduceia versatilis* is a member of the "*Synergistes*" group. Appl Environ Microbiol 73: 6270–6276.

79. Inoue J-I, Noda S, Hongoh Y, Ui S, Ohkuma M (2008) Identification of endosymbiotic methanogen and ectosymbiotic spirochetes of gut protists of the termite *Coptotermes formosanus*. Microbes Environ 23: 94–97.

80. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 21: 2947–2948.

81. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104–2105.

82. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

83. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755.

84. Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18: 502–504.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

85.   Lartillot N, Brinkmann H, Philippe H (2007) Suppression of long-branch attraction artifacts in the animal phylogeny using a site-heterogeneous model. BMC Evol Biol 7(Suppl 1): S4.

86.   Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17: 1246–1247.

87.   Kitade O, Maeyama T, Matsumoto T (1997) Establishment of symbiotic flagellate fauna of *Hodotermopsis japonica* (Isoptera: Termopsidae). Sociobiology 30: 161–167.

Figure Legends

Figure 1. Maximum likelihood phylogenetic analyses of parabasalids based on GAPDH (A), actin (B), and EF-1α (C) sequences. Unambiguously aligned sequences of 277 (A), 280 (B) and 274 (C) sites were used for phylogenetic inference. The species names of the parabasalids are shown in Table 1. The GAPDH sequences of *Trichonympha acuta* and *Eucomonympha imla* and the EF-1α sequence of *Trichomonas tenax* (not shown in Table 1) were also included in the analyses. The sequence accession number was indicated for each taxon. The sequences used for the concatenation are in bold. The trees were estimated in RAxML and the numbers near the nodes indicate the bootstrap values. Values below 50% are not shown. Vertical bars to the right of the trees represent the parabasalian classes according to [24]: C, Cristamonadea; S, Spirotrichonymphea; Tn, Trichonymphea; Tt, Tritrichomonadea; H, Hypotrichomonadea; and Tm, Trichomonadea. Scale bars correspond to 0.10 substitutions per site.

Figure 2. Phylogenetic relationship of parabasalids inferred from the concatenated dataset. The concatenated dataset comprising 257 sites of GAPDH, 268 sites of actin, 274 sites of EF-1α, and 1338 sites of SSU rRNA gene sequences was analyzed in 28 parabasalian species. The tree was estimated in RAxML using separate models with the parameters and branch length optimized for each gene partitions individually. The supporting values (bootstrap in RAxML/Bayesian posterior probability) are indicated at the nodes. Values below 50% or 0.5 are indicated with hyphens. When the site-heterogeneous CAT model was used in each partition, the identical tree topology with similar bootstrap values was obtained (data not shown). Vertical bars to the right of the tree represent the parabasalian classes. The scale bar corresponds to 0.10 substitutions per site.

Figure 3. A maximum likelihood tree based on concatenation of actin, EF-1α, and SSU rRNA gene sequences and rooted by diverse eukaryotes. Unambiguously aligned sites of actin (268), EF-1α (274), and SSU rRNA (1265) gene sequences were concatenated and analyzed in 30 parabasalian species and 23 diverse eukaryotes as outgroups. The tree was estimated with RAxML using the CAT model (CATMIX). The parameters and branch length were optimized for each gene partition individually. The supporting values (bootstrap in RAxML/Bayesian posterior probability) are indicated at the nodes. Values below 50% or 0.5 are indicated with hyphens. Vertical bars to the right of the tree represent the parabasalian classes. The 11 possible root positions are indicated in red letters. The scale bar corresponds to 0.10 substitutions per site.

Figure 4. A maximum likelihood tree based on concatenation of GAPDH, actin, EF-1α, and SSU rRNA gene sequences and rooted by *Trimastix*. Unambiguously aligned sites of GAPDH (268), actin (257), EF-1α (274), and SSU rRNA (1338) gene sequences were concatenated and analyzed in 28 parabasalian species with *Trimastix* as an outgroup. The tree was estimated in RAxML using the CAT model (CATMIX). The parameters and branch

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

length were optimized for each gene partition individually. The supporting values (bootstrap in RAxML/Bayesian posterior probability) are indicated at the nodes. Values below 50% or 0.5 are indicated by hyphens. Vertical bars to the right of the tree represent the parabasalian classes. The scale bar corresponds to 0.10 substitutions per site.

Figure 5. Proposed evolutionary relationships of parabasalids. Tree showing the relationships of the six parabasalian classes based on the results of the present study. The common ancestor of parabasalids very likely possessed a parabasal apparatus and hydrogenosomes because they are common characters of parabasalids. As opposed to previous molecular phylogenetic studies, the results of our study indicate that the common ancestor possibly shares trichomonad-like characters, although possibilities of other traits such as multiple flagella in a mastigont system cannot be excluded. Flagellar multiplication in a single mastigont system has occurred independently in the classes Trichonymphea, Spirotrichonymphea, and Cristamonadea (boxed), and members of these classes are almost exclusively present in the gut of termites and related insects. These multiplications have occurred ancestrally in the former two classes (marked with filled circles) and probably at least twice within the latter (open circle) because of the polyphyly of lophomonads [46]. In Cristamonadea, multinucleate forms (calonymphids) as well as cells with one nucleus and well-developed cytoskeletal structures (devescovinids) also appear. Simplification of the cytoskeleton, such as secondary loss or reduction of the undulating membrane (UM) and costa, has likely occurred within the classes Tritrichomonadea, Hypotrichomonadea, and Trichomonadea (triangles). Although the Trichomonadea members examined in this study (*Trichomonas*, *Tetratrichomonas*, and *Pentatrichomonas*) do not show such a simplification, species of the simple cell form such as *Pseudotrichomonas* and *Lacusteria* are closely related to them in the SSU rRNA gene analyses [25]. See the text for details.

Tables

Table 1. Parabasalian species used for the gene identification and phylogenetic analyses.

Species[a]	Class[b]	Family[c]	Host animal[d]
Macrotrichomonas sp.	C	L (D)	*Glyptotermes satsumensis*
Metadevescovina cuspidata	C	L (D)	*Incisitermes minor*
Foaina nana	C	L (D)	*Cryptotermes domesticus*
Caduceia versatilis	C	L (D)	*Cryptotermes cavifrons*
Devescovina sp.	C	L (D)	*Neotermes koshunensis*
Gigantomonas herculea[e]	C	L (D)	*Hodotermes mossambicus*
Stephanonympha sp. CcSt	C	L (C)	*Cryptotermes cavifrons*
Stephanonympha sp. NkSt	C	L (C)	*Neotermes koshunensis*
Snyderella tabogae	C	L (C)	*Cryptotermes cavifrons*
Deltotrichonympha sp.[e]	C	L (De)	*Mastotermes darwiniensis*
Joenina pulchella	C	L	*Porotermes adamsoni*
Joenia annectens	C	L	*Kalotermes flavicollis*
Joenoides intermedia	C	L	*Hodotermes mossambicus*
Spirotrichonympha leidyi	S	Hl (Sp)	*Coptotermes formosanus*
Holomastigotoides mirabile	S	Hl	*Coptotermes formosanus*
Staurojoenina assimilis	Tn	St	*Incisitermes minor*
Trichonympha agilis	Tn	Tn	*Reticulitermes speratus*
Trichonympha sp.	Tn	Tn	*Hodotermopsis sjoestedti*
Hoplonympha sp.	Tn	Hp	*Hodotermopsis sjoestedti*
Pseudotrichonympha grassii	Tn	Te (E)	*Coptotermes formosanus*
Eucomonympha sp.	Tn	Te (E)	*Hodotermopsis sjoestedti*
Teranympha mirabilis	Tn	Te	*Reticulitermes speratus*
Histomonas meleagridis	Tt	Di (M)	*Meleagris gallopavo*
Dientamoeba fragilis[e]	Tt	Di (M)	*Homo sapiens*
Tritrichomonas foetus	Tt	Tt (Tm)	*Bos primigenus*
Monocercomonas sp.	Tt	M	*Natrix sipedon*
Hypotrichomonas acosta	H	H (M)	*Drymarchon corais couperi*
Trichomitus batrachorum	H	H (Tm)	*Elaphe obsoleta*
Pentatrichomonas hominis[e]	Tm	Tm	*Homo sapiens*
Tetratrichomonas gallinarum	Tm	Tm	*Anas platyrhynchos*
Trichomonas vaginalis[e]	Tm	Tm	*Homo sapiens*

[a]According to a comprehensive list of flagellate species in the gut of termites [3], cristamonads inhabiting each termite species are as follows: *P. adamsoni* and *I. minor* harbors only *J. pulchella* and *M. cuspidata*, respectively; *H. mossambicus* harbors only *J. intermedia* and *G. herculea*; *K. flavicollis* harbors two *Foaina* spp. in addition to *J. annectens*; *C. cavifrons* harbors species of *Foaina* in addition to the three cristamonads examined in this study; and *C. domesticus* harbors species of *Devescovina* and *Stephanonympha* in addition to *F. nana*. As cristamonad symbionts, *N. koshunensis* harbors species of *Foaina* as well as *Devescovina* and *Stephanonympha* [38], and *G. satsumensis* harbors species of *Foaina* and *Devescovina* in addition to *Macrotrichomonas* sp. [46]. The SSU rRNA and GAPDH genes of the cristamonad species mentioned above (except *Stephanonympha* sp.) were simultaneously analyzed using the same cell preparations [46]. *M. darwiniensis* harbors species of *Koruga*, *Mixotricha*, and *Metadevescovina* in addition to *Deltotrichonympha* [3], and the SSU rRNA gene sequence of *Deltotrichonympha* sp. obtained from our preparation was almost identical to those of *Deltotrichonympha* spp. (AJ583378, AJ583378, and AB326380). *C. formosanus* harbors only three parabasalian species [3], all of which were examined in this study. As species in Trichonymphea that show conspicuous morphology and thus are easily recognizable, *I. minor* harbors only *S. assimilis*

[3], *R. speratus* harbors only *T. agilis* and *T. mirabilis* [3], *H. sjoestedti* harbors only species of *Trichonympha*, *Eucomonympha*, and *Hoplonympha* [87]. All of these genera were examined in this study.

[b]Abbreviations of the classes are: C, Cristamonadea; Tt, Tritrichomonadea; S, Spirotrichonymphea; H, Hypotrichomonadea; Tm, Trichomonadea; and Tn, Trichonymphea.

[c]When the family name has changed in the new parabasalian classification [24], the corresponding former name [1] is also indicated in parenthesis. Abbreviations of the families are the following: L, Lophomonadidae; D, Devescovinidae; C, Calonymphidae; De, Deltotrichonymphidae; Sp, Spirotrichonymphidae; Hl, Holomastigotoididae; St, Staurojoeninidae; Tn, Trichonymphidae; Hp, Hoplonymphidae; E, Eucomonymphidae; Te, Teranymphidae; Di, Dientamoebidae; Tt, Tritrichomonadidae; M, Monocercomonadidae; and Tm, Trichomonadidae.

[d]*M. darwiniensis*, *P. adamsoni*, *K. flavicollis*, and *H. mossambicus* were collected in Australia, Australia, France, and Kenya, respectively, and generously provided by C. Bordereau (Université de Bourgogne, France). *C. cavifrons* collected in the United States was provided by M. F. Dolan (University of Massachusetts, USA). *I. minor* collected in Japan was provided by W. Ohmura (Forestry and Forest Products Research Institute, Japan). *G. satsumensis*, *C. domesticus, C. formosanus, R. speratus*, and *H. sjoestedti* were collected in Japan.

[e]We failed to identify the EF-1α gene in *G. herculea*. The GAPDH genes of *Deltotrichonympha* sp. and *D. fragilis* were unavailable. The database sequences of *P. hominis* and *T. vaginalis* were used for phylogenetic analyses (see Methods).

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Table 2. Shimodaira-Hasegawa test for alternative monophyletic relationships of parabasalian classes.

Monophyletic constraint	P value
C + S	0.614
C + H	0.119
Tt + S	0.426
Tt + H	0.182
S + H	0.866
C + Tm	<0.001*
C + Tn	<0.001*
Tt +Tm	<0.001*
Tt +Tn	<0.001*
S + Tm	0.011*
S + Tn	0.020*
H + Tm	0.016*
H + Tn	0.067

Abbreviations of the classes are shown in the footnote of Table 1 or the legend of Figure 1. Asterisks indicate that the tested monophyly was significantly different from the best ML topology at $P < 0.05$. Each of the monophyletic groupings of C + Tt and Tm + Tn appeared in the best ML topology.

Table 3. Shimodaira-Hasegawa test for parabasalian root positions.

Root position	P value
a	0.012*
b	0.088
c	0.005*
d	0.019*
e	0.367
f	0.378
g	0.491
h	0.009*
i	0.018*
j	0.065
k	Best

Root positions are depicted in Figure 3. Asterisks indicate that the root position was significantly different from the best ML topology at $P < 0.05$.

Table 4. Exclusion of parabasalian taxa and the effect on their root.

No. of parabasalian taxa excluded	23 outgroup taxa	13 outgroup taxa
3	10	NT
6	10	NT
12	10	10
16	4	8
18	5	6

Values represent the number of occurrences of root position k (shown in Figure 3) in 10 replicates of the random taxa exclusion analyses in each defined number of excluded taxa. NT, not tested. The 23 outgroup taxa correspond to the concatenate dataset of EF-1α, actin, and SSU rRNA gene sequences using 23 outgroup taxa as shown in Figure 3, whereas the 13 outgroup taxa correspond to those remained after excluding 10 long-branch outgroup taxa (as investigated in Table S1). Note that in the cases of 16 and 18 taxa exclusions, all other replicates demonstrated the root position at the branch leading to Trichonymphea (position g).

Supporting Information

Figure S1. Exclusion of GAPDH from the sequence concatenation of parabasalids. The GAPDH sequence was excluded from the dataset used for the inference shown in Figure 2, and the sequences of *Deltotrichonympha* sp. and *Dientamoeba* were added to the analysis. The tree was inferred using RAxML with the parameters and branch length optimized for each partition. The bootstrap values above 50% are indicated at the nodes. Note that the relationship of the six parabasalian classes and the root position did not changed after the exclusion except for the paraphyly of Tritrichomonadea.
(PDF)

Figure S2. Exclusion of SSU rRNA gene from the sequence concatenation of parabasalids. The SSU rRNA gene sequence was excluded from the dataset used for the inference shown in Figure 2, and the tree was inferred using RAxML with the parameters and branch length optimized for each partition. The bootstrap values above 50% are indicated at the nodes. Note that the relationship of the six parabasalian classes and the root position did not change after the exclusion except for the paraphyly of Tritrichomonadea and the sister-group relationship of Spirotrichonymphea and Hypotrichomonadea. The removal of the SSU rRNA gene sequence was particularly important because the base composition of this gene was heterogeneous in the two Spirotrichonymphea members and *Histomonas*. All taxa were homogeneous in amino acid composition in the three protein sequences.
(PDF)

Figure S3. A maximum likelihood tree of the sequence concatenation of actin, EF-1α, α-tubulin, β-tubulin, and SSU rRNA gene in 12 parabasalian taxa and outgroup eukaryotes. Unambiguously aligned sites of α-tubulin (351) and β-tubulin (328) were concatenated with the dataset used for the analysis shown in Figure 3. The sequence data for both tubulins used for the concatenation are the same as described previously [53]. The tree was inferred using RAxML with the parameters and branch length optimized for each partition. The bootstrap values above 50% are indicated at the nodes. Because the sequences of α- and β-tubulin available for *Pentatrichomonas*, *Histomonas*, and *Entamoeba* are short, we excluded them from the analysis. Trichonymphea was the most basal parabasalian lineage in this analysis; however, this rooting was likely a wrong inference caused by the limited taxon sampling (see the main text). Spirotrichonymphea instead of Tritrichomonadea was sister to Cristamonadea, but this change was not supported at all. The taxa exclusion analyses (see Table 4) indicated that when the number of parabasalian taxa was reduced, the frequency of the pairing of Cristamonadea and Spirotrichonymphea increased up to four times in 10 repeated analyses, suggesting an artificial pairing owing to the limited number of examined taxa.

Table S1. Removal of long-branch outgroup taxa and the effect on bootstrap values for major nodes within Parabasalia. (DOC)

Table S2. Species used for the outgroup and their sequence accession numbers used for the concatenation. (DOC)

Dataset S1. NEXAS-format alignment of sequence concatenation of GAPDH, actin, EF-1α, and SSU rRNA gene used for the analyses shown in Figure 2.

Dataset S2. NEXAS-format alignment of sequence concatenation of actin, EF-1α, and SSU rRNA gene used for the analyses shown in Figure 3.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
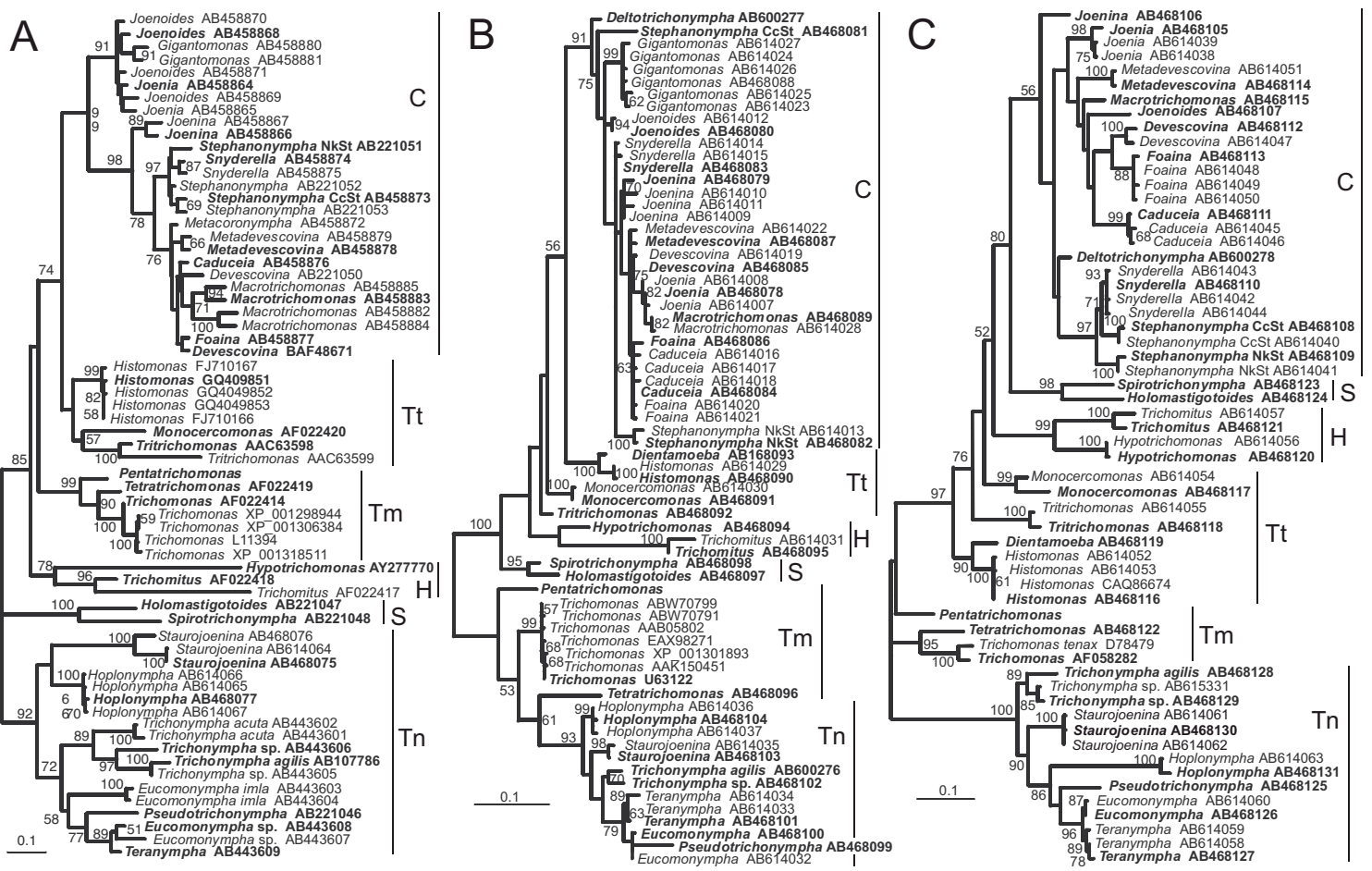*Università degli studi di Sassari*

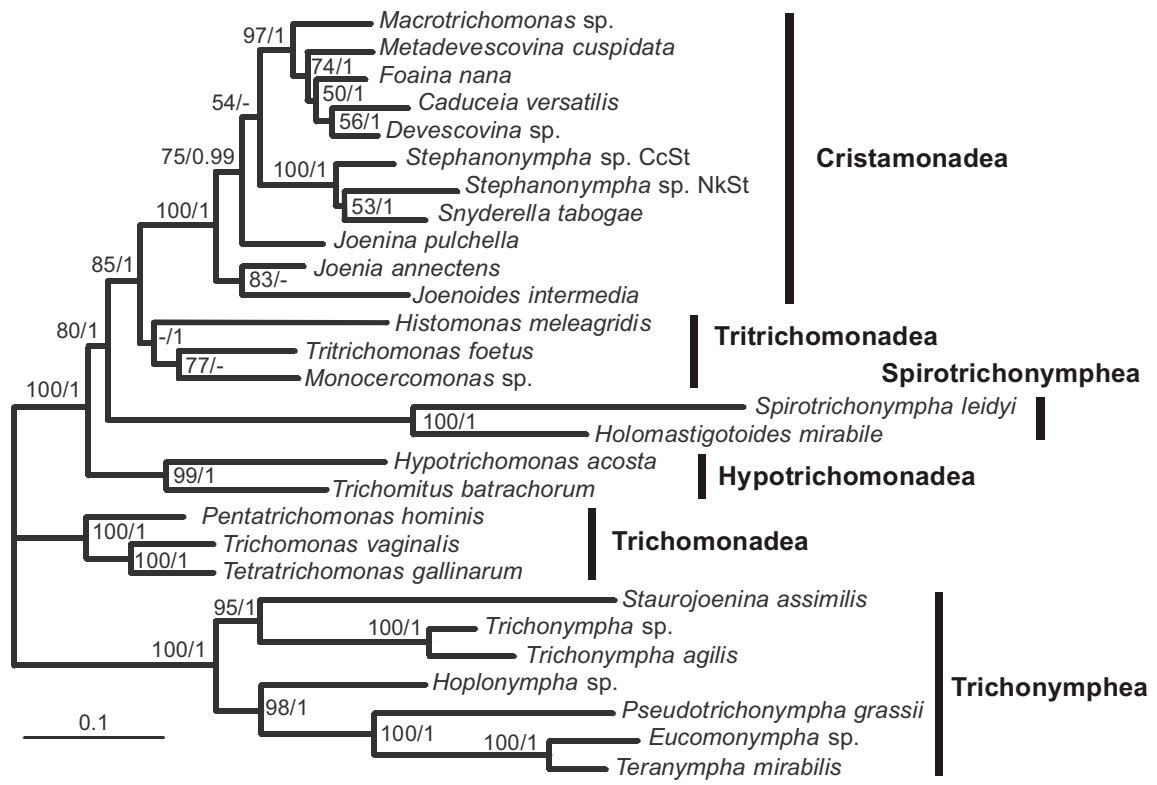Figure 1

Figure 2

Figure 3

Figure 4

Cristamonadea — Numerous flagella regressing at division
or multi-nucleate
or well-developed pelta-axostyle complex
UM absent, costa absent

Tritrichomonadea — UM, rail-like or lameliform or absent
costa, A-type or absent

Spirotrichonymphea — spiraled rows of numerous flagella

Hypotrichomonadea — lamelliform UM
costa, A-type or reduced

Trichomonadea — lamelliform UM
costa, B-type or absent

Trichonymphea — numerous flagella of
bilateral symmetry in the rostrum

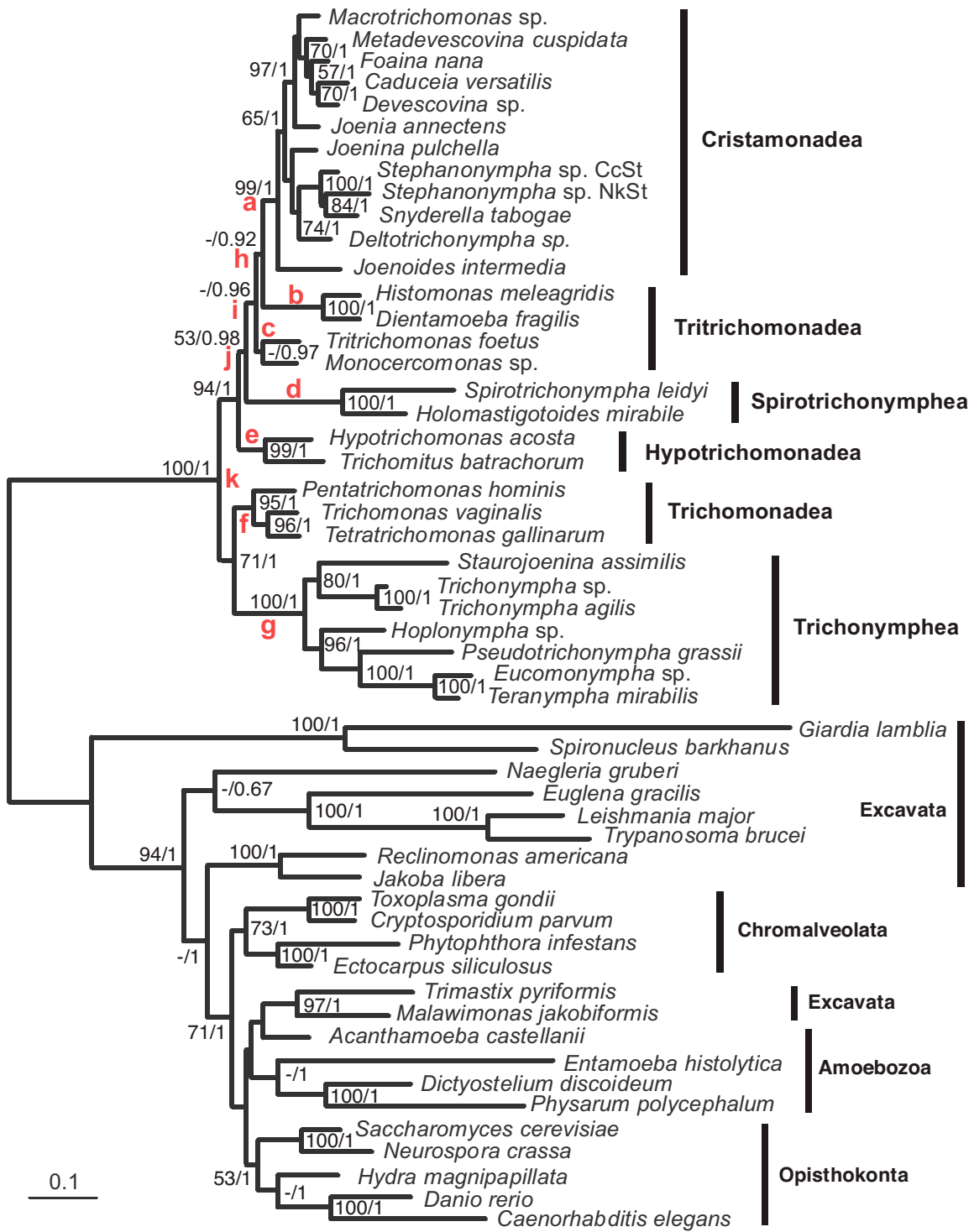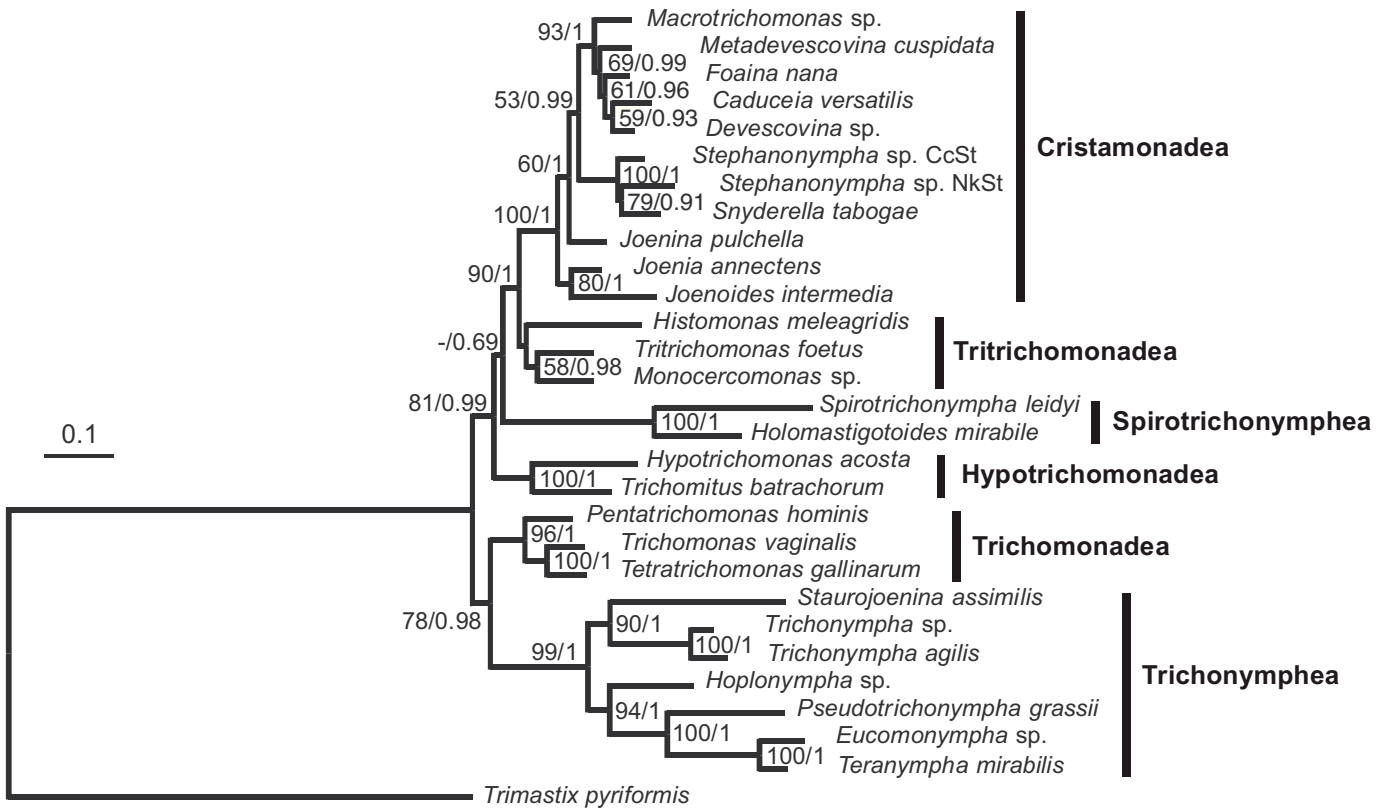ancestor of parabasalids —
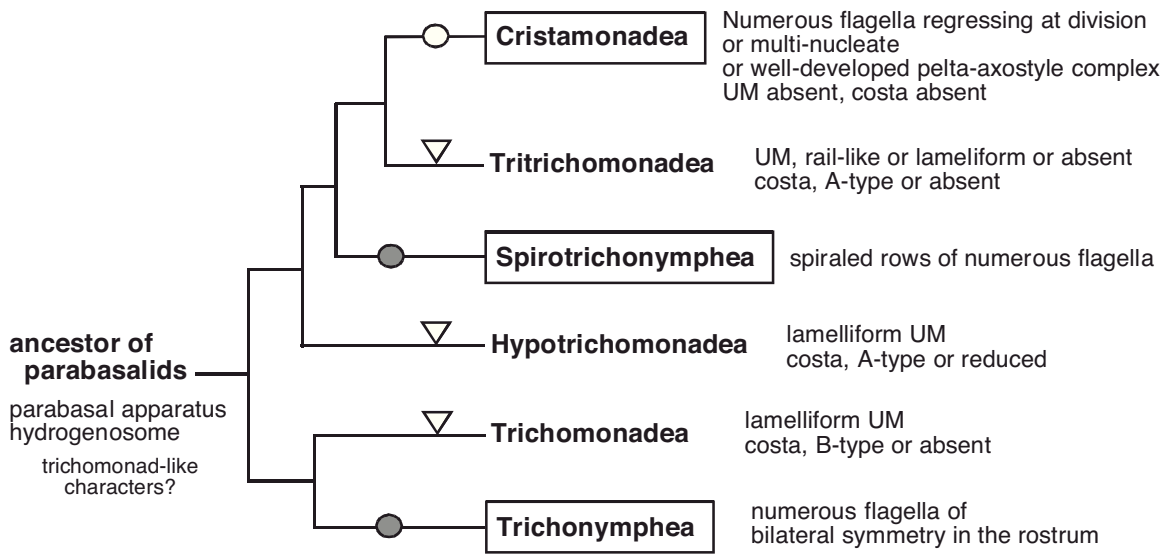parabasal apparatus
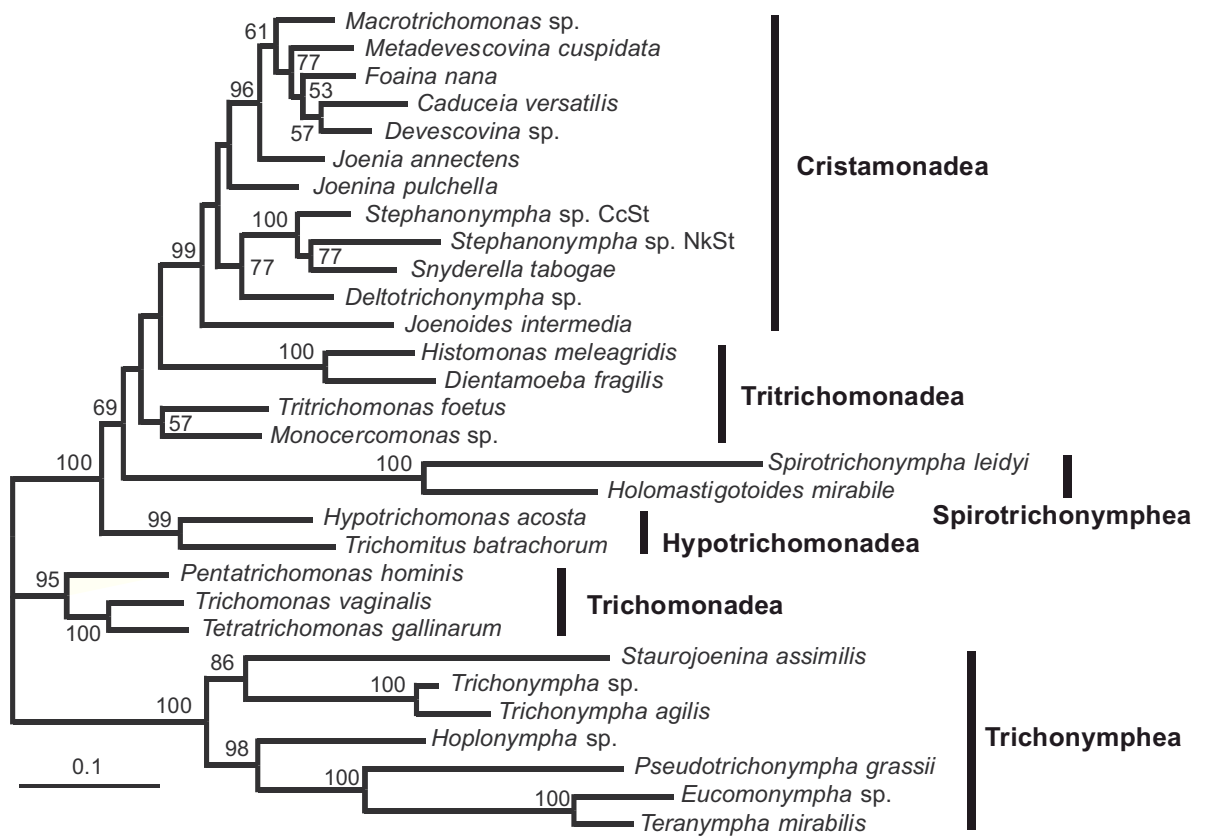hydrogenosome
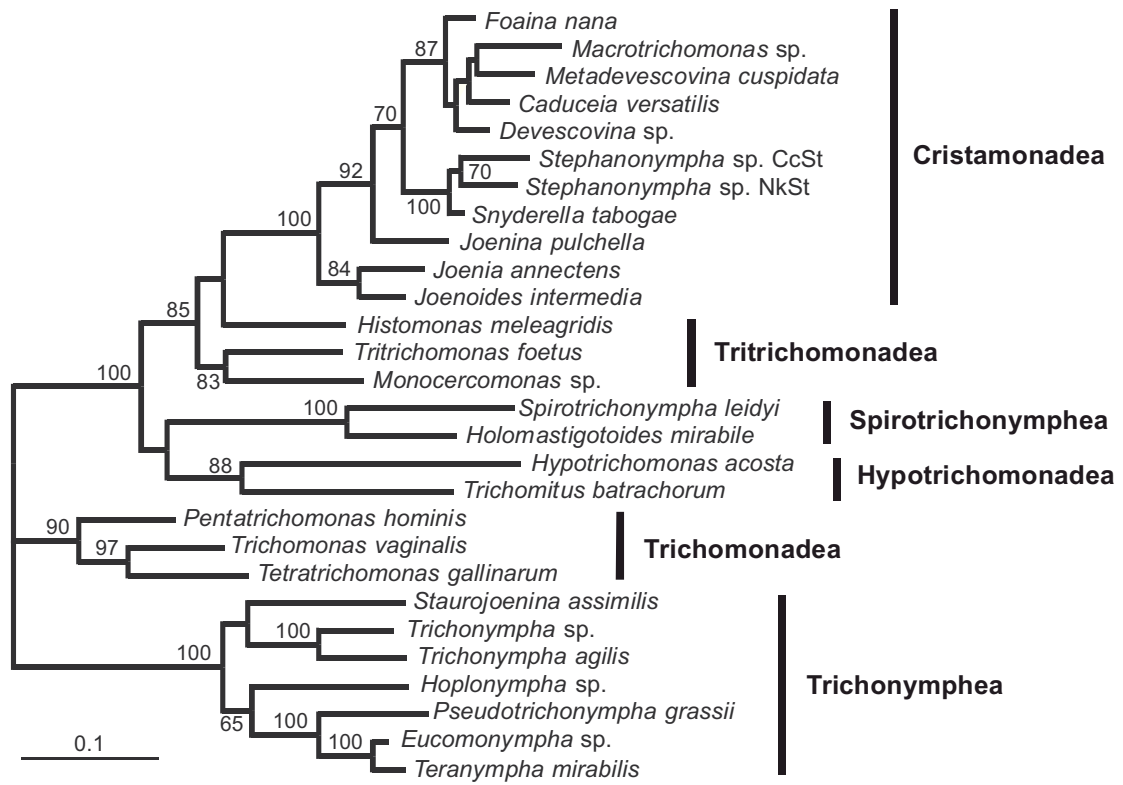trichomonad-like
characters?

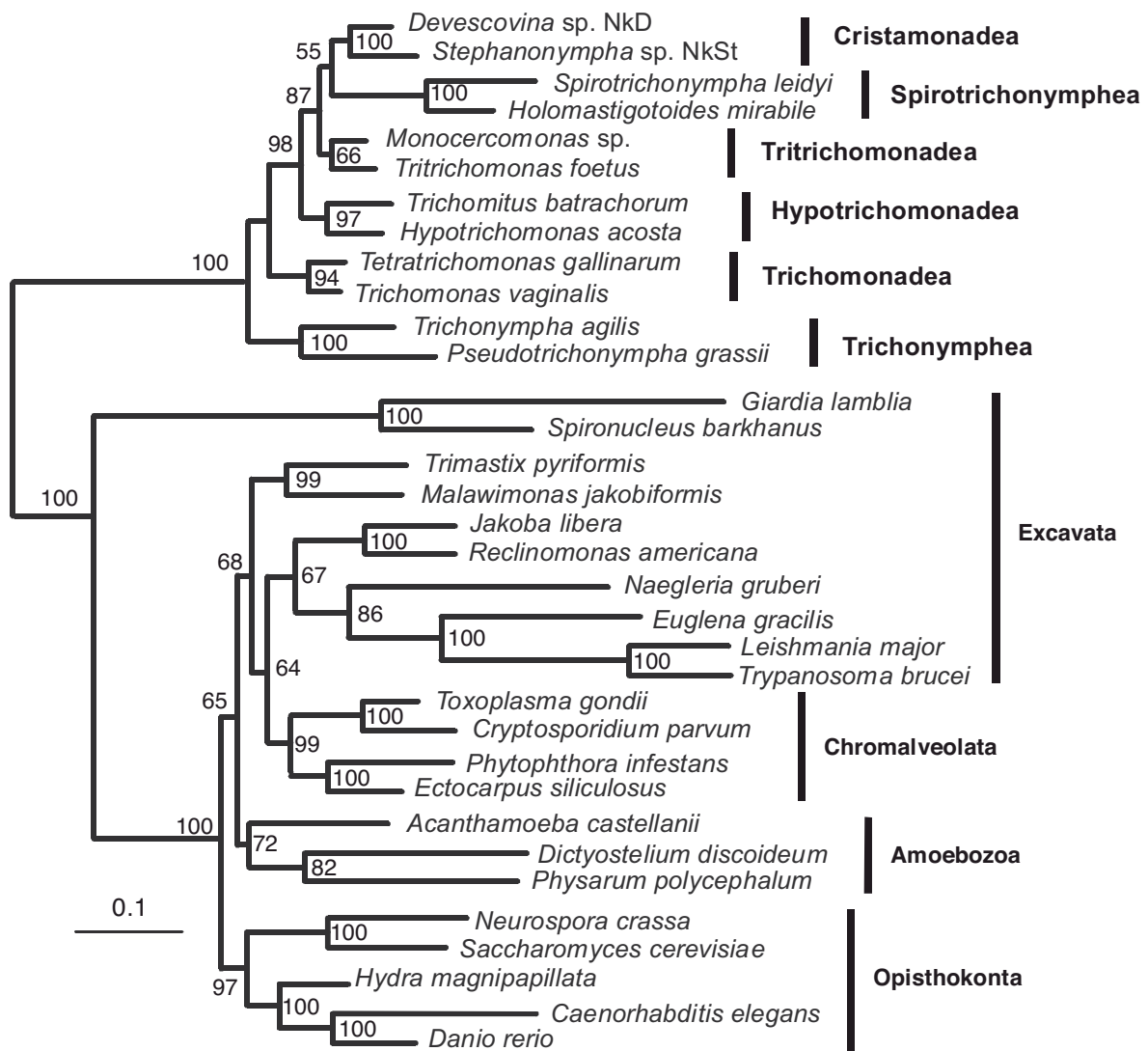Figure 5

Figure S1

Figure S2

Figure S3

Table S1. Removal of long-branch outgroup taxa and the effect on bootstrap values for major nodes within Parabasalia.

Node[a]	Number of outgroup taxa removed[b]										
	0	1	2	3	4	5	6	7	8	9	10
a	99	99	99	99	99	99	99	100	99	99	99
b	100	100	100	100	100	100	100	100	100	100	100
c	35	34	40	39	38	33	34	37	0	30	36
d	100	100	100	100	100	100	100	100	100	100	100
e	97	96	96	97	97	95	96	94	92	94	96
f	93	90	90	89	88	92	91	95	93	96	94
g	100	100	100	100	100	100	100	100	100	100	100
h	25	29	33	33	32	28	29	31	0	22	28
i	42	42	45	48	44	43	45	47	0	40	44
j	53	46	56	57	61	54	58	60	55	54	55
k	90	92	90	93	94	92	89	84	85	81	86
l	74	68	60	61	72	80	79	82	77	84	80
m	100	100	100	100	100	100	100	100	100	100	100

The distance from the inferred root position to each outgroup taxon was calculated in the ML tree shown in Figure 3 using TreeStat (http://tree.bio.ed.ac.uk/software/treestat/). The longest branched outgroup taxon was progressively excluded from the bootstrap analyses of 100 replicates using RAxML with the CAT model (GTRCAT for SSU rRNA gene and PROTCATWAG for the protein sequences). Note that the root position did not change through the removal analyses and that the bootstrap support values for this rooting were substantially stable (see nodes k and l).

[a] Nodes a to j are those indicated in Figure 3. Node k corresponds to the grouping of Cristamonadea, Tritrichomonadea, Spirotrichonymphea, and Hypotrichomonadea, whereas node l to the grouping of Trichomonadea and Trichonymphea. Node m corresponds to the monophyly of Parabasalia.

[b] Outgroup taxa removed progressively were 1, *Giardia*; 2, *Trypanosoma*; 3, *Leishmania*; 4, *Entamoeba*; 5, *Spironucleus*; 6, *Euglena*; 7, *Physarum*; 8, *Naegleria*; 9, *Caenorhabditis*; and 10, *Trimastix*.

Table S2. Species used for the outgroup and their sequence accession numbers used for the concatenation.

Species	SSU rRNA	Actin	EF-1α	α-tubulin	β-tubulin
Giardia lamblia	NW_001844081	EDO76979	XP_001704547	XP_001706843	XP_001707388
Spironucleus barkhanus	AY646679	EST data[a]	AAC47211	AAC47209	EST data[a]
Trimastix pyriformis	AF244903	TPL00000260	ABC54653	TPL00000212	TPL00001398
Malawimonas jakobiformis	EF455761	ABX25969	ABC54649	AF267181	AF267185
Naegleria gruberi	M18732	AAF37002	ABC54650	CAA56939	XP_002669519
Euglena gracilis	M12677	AAC99646	ACO50110	AAK37832	AAK37834
Leishmania major	NC_007268	AAA19789	XP_001682264	CAJ02503	CAJ06135
Trypanosoma brucei	M12676	XP_828467	XP_828111	XP_001218934	XP_001218933
Reclinomonas americana	AY117417	AAX09575	ABC54651	AF267182	AF267190
Jakoba libera	AY117418	JLL00000765	ACO50113	JLL00000756	JLL00000772
Cryptosporidium parvum	AF164102	AAM28417	XP_00138834	XP_625871	XP_627803
Toxoplasma gondii	U03070	CAJ20602	CAJ20335	XP_002364807	AAA30146
Phytophthora infestans	NW_003303738	XP_002898250	XP_002905383	EEY54372	XP_002908783
Ectocarpus siliculosus	L43062	CBJ30601	CBJ32894	CBJ28184	CBN79445
Entamoeba histolytica	X64142	AAA29085	XP_651869	–	–
Dictyostelium discoideum	AM168071	XP_636169	XP_645978	XP_637058	XP_646162
Physarum polycephalum	X13160	CAA30629	AAB69706	CAA28712	AAA29974
Acanthamoeba castellanii	AF260724	CAA23399	AAU94656	AAZ80770	AAZ80771
Saccharomyces cerevisiae	HQ174900	AAA34391	CAA55620	AAA35180	CAA24603
Neurospora crassa	X04971	AAC78496	XP_964868	EAA29668	EAA28433
Hydra magnipapillata	HQ392522	XP_002154696	XP_002160595	XP_002159229	XP_00216191
Danio rerio	BX537263	AAO38846	NP_571338	NP_919369	NP_942104
Caenorhabditis elegans	EU196001	CAA34720	P53013	BAA03909	NP_499367

[a] The EST sequences in the public database (GW585169-GW589878, and EC585128-EC586011) were searched and assembled to produce in silico-translated amino acid sequence.

Part III

_____

MOLECULAR EPIDEMIOLOGY OF *BLASTOCYSTIS* SP. IN

ITALY AND LEBANON

III-1 <u>Intoduction</u>

Blastocystis sp. is an emerging parasite with a worldwide distribution. Although *Blastocystis* sp. isolates from humans and animals have been reported to be morphologically indistinguishable, a considerable genetic diversity has been revealed among isolates with the identification of at least 13 STs. Each of the STs exhibited sufficient genetic divergence to be classified as separate species. This protozoa is often identified as the most common eukaryotic organism reported in human fecal samples and its prevalence has shown a dramatic increase in recent years. Indeed, *Blastocystis* sp. is one of the few enteric parasites with a prevalence that often exceeds 5% in the general population of developed countries, can reach 30% to 60% in developing countries, and largely exceeds 40% in individuals with chronic gastrointestinal illness such as IBS. This parasite is also frequently found in immunocompromised individuals (HIV/AIDS and cancer patients) and a higher risk of *Blastocystis* sp. infection has been found in humans with close animal contact. Such a prevalence of *Blastocystis* sp. in the human population and its potential for transmission from animals naturally raise questions about the impact of this parasite in public health and the significance of its genetic diversity on pathogenicity of variants.

Epidemiological studies have reported the frequency of STs from symptomatic and asymptomatic patients in several countries. These studies permit characterization of the transmission dynamics and pathogenicity of the different *Blastocystis* sp. STs and tracking infection and contamination sources. However, in Italy as in Lebanon, only preliminary epidemiological data are available concerning this parasite. Therefore, STs of a large number of *Blastocystis* sp. clinical isolates were identified from symptomatic patients in the Italian population **(Article 4)** and from symptomatic and asymptomatic patients in the Lebanese population **(Article 5)**.

III-2 <u>Results</u>

Article 4 (see above)

<u>Meloni D</u>, Sanciu G, Poirier P, El Alaoui H, Chabé M, Delhaes L, Dei-Cas E, Delbac F, Fiori PL, Di Cave D, Viscogliosi E. Molecular subtyping of *Blastocystis* sp. isolates from symptomatic patients in Italy. Parasitology Research, 2011, 109:613-619

Article 5

El Safadi D, <u>Meloni D</u>, Poirier P, Wawrzyniak I, El Alaoui H, Chabé M, Delhaes L, Dei-Cas E, Delbac F, Fiori PL, Hamze M, Viscogliosi E. Molecular epidemiology of human *Blastocystis* sp. isolates in Lebanon. European Journal of Clinical Microbiology and Infectious Diseases, 2011, in preparation

Brief description of Article 5: In this study performed in collaboration with the AZM Center for Research in Biotechnology of Tripoli, 220 stool specimens were collected in several hospitals and examined by light microscopy for the presence of *Blastocystis* sp. Among these samples, 42 were positive. For each positive patient, clinical and epidemiological data are notified using a standardized questionnaire. This questionnaire refers to the immune status of the patient, presence of gastrointestinal or dermatologic symptoms, association with IBS, presence of other parasitic infections, contact with pets or farm animals, and recent history of travelling. DNA is directly extracted from fecal samples and amplified by PCR using *Blastocystis* sp. specific primers targeting a 600 bp-fragment of the SSU rDNA coding region as described elsewhere (Souppart et al. 2009, 2010 and Article 4). After cloning of the PCR product, two clones are arbitrarily selected and sequenced. For subtyping, the sequences are compared with those of all known *Blastocystis* sp. STs available in GenBank using the Blast program. Among the 42 positive samples through direct light microscopy of smears, 19 were successfully amplified following the PCR protocol and genotyped. The remaining 23 samples were negative because of the presence of known PCR inhibitors in fecal samples. However all these 23 samples are positive by real-time quantitative PCR (data not shown), a more sensitive approach (Poirier et al. 2011), and we are waiting for the sequences of these PCR products. The 19 isolates belong to ST3 (53%), ST1 and ST2 (21%) and ST4 (5%). Among the 19 patients, 10 were asymptomatic and 9 symptomatic. At present time, the sample size remains too low and statistical analysis has not yet been performed to correlate patient symptomatic status and *Blastocystis* sp. ST. However, we note that ST1 isolates are only found in symptomatic patients. Interestingly, in a recent *in vivo* study, Hussein et al. (2008) have shown that human ST1 isolates obtained from symptomatic patients induced 25% mortality in rats.

III-3 <u>Conclusions</u>

In Italy, stool specimens were collected from a total of 30 Italian patients living in or in the vicinity of Rome and Sassari. All these patients presented various gastrointestinal troubles

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

and were thus classified as symptomatic. A total of 34 *Blastocystis* sp. isolates corresponding to 26 single and 4 mixed infections (presence of 2 STs in the same individual) were subtyped. From this molecular approach, the ST distribution in the present Italian population was as follows: ST3 47.1%; ST2 20.6%, ST4 17.7%, ST1 8.8% and ST7 and ST8 2.9%. Together with a previous preliminary report including the subtyping of seven *Blastocystis* sp. isolates, a total of seven STs (with the addition of ST5) have been found in Italian symptomatic patients.

In Lebanon, 220 stool samples were examined for the presence of *Blastocystis* sp. Among these samples, 42 were positive indicating a large prevalence of 19% in the Lebanese population. To our knowledge, this represents the first data concerning the prevalence of this parasite in Lebanon. To date, among these 42 samples, 19 of them have been genotyped. The ST distribution in the present Lebanese population was as follows: ST3 (53%) followed by ST1 and ST2 (21%) and ST4 (5%). Among the 19 patients whose *Blastocystis* sp. isolates were genotyped, 10 were asymptomatic and 9 symptomatic. Interestingly, ST1 isolates were only found in symptomatic patients.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Molecular subtyping of Blastocystis sp. isolates from symptomatic patients in Italy

Dionigia Meloni, Giovanna Sanciu,
Philippe Poirier, Hicham El Alaoui,
Magali Chabé, Laurence Delhaes,
Eduardo Dei-Cas, Frederic Delbac, et al.

Volume 108 · Number 5 · May 2011

Parasitology
Research

Springer

Springer

Springer

ORIGINAL PAPER

Molecular subtyping of *Blastocystis* sp. isolates from symptomatic patients in Italy

**Dionigia Meloni · Giovanna Sanciu · Philippe Poirier · Hicham El Alaoui ·
Magali Chabé · Laurence Delhaes · Eduardo Dei-Cas · Frederic Delbac ·
Pier Luigi Fiori · David Di Cave · Eric Viscogliosi**

Abstract *Blastocystis* sp. is the most common eukaryotic parasite in the intestinal tract of humans. Due to its potential impact in public health, we determined the *Blastocystis* sp. subtypes (STs) and their relative frequency in symptomatic patients living in or in the vicinity of two Italian cities (Rome and Sassari). A total of 34 *Blastocystis* sp. isolates corresponding to 26 single and 4 mixed infections were subtyped using partial small subunit ribosomal RNA gene sequencing. From this molecular approach, the ST distribution in the present Italian population was as follows: ST3 (47.1%), ST2 (20.6%), ST4 (17.7%), ST1 (8.8%), and ST7, and ST8 (2.9%). As in almost all countries worldwide, ST3 was the most common ST reinforcing the hypothesis of its human origin. Together with a previous preliminary report, a total of seven STs (with the addition of ST5) have been found in Italian symptomatic patients. The wide range of STs identified in the Italian population suggest that *Blastocystis* sp. infection is not associated with specific STs even if some STs (ST1–ST4) are predominant as reported in all other countries. Since most of the STs identified in Italian patients are zoonotic, our data raise crucial questions concerning the identification of animal reservoirs for *Blastocystis* sp. and the potential risks of transmission to humans.

D. Meloni · G. Sanciu · M. Chabé · L. Delhaes · E. Dei-Cas ·
E. Viscogliosi (✉)
Center for Infection and Immunity of Lille (CIIL), Institut Pasteur
of Lille, Inserm U1019, CNRS UMR 8204, University Lille Nord
de France, Biology and Diversity of Emerging Eukaryotic
Pathogens (BDEEP),
1 rue du Professeur Calmette, BP 245, 59019 Lille Cedex, France
e-mail: eric.viscogliosi@pasteur-lille.fr

D. Meloni · G. Sanciu · P. Luigi Fiori
Department of Biomedical Sciences, Division of Experimental
and Clinical Microbiology, University of Sassari,
Viale San Pietro 43/B,
07100 Sassari, Italy

D. Meloni · G. Sanciu · P. Luigi Fiori
Center for Biotechnology Development and Biodiversity
Research, University of Sassari,
07100 Sassari, Italy

P. Poirier · H. El Alaoui · F. Delbac
Laboratoire Microorganismes: Génome et Environnement
(LMGE), CNRS UMR 6023, University Blaise Pascal of
Clermont-Ferrand,
Avenue des Landais,
63177 Aubière Cedex, France

D. Di Cave
Department of Public Health and Cell Biology, University of
Rome Tor Vergata,
Via Montpellier 1,
00133 Rome, Italy

Introduction

Blastocystis sp. is an anaerobic parasitic protist, inhabiting the lower gastrointestinal tract of humans and a wide range of animals (for reviews, see Abe et al. 2002; Stenzel and Boreham 1996; Tan 2004, 2008) and responsible for frequent community infections. With the application of modern molecular phylogenetic approaches, *Blastocystis* sp. was unambiguously placed among the Stramenopiles, a complex and heterogeneous evolutionary assemblage of "botanical" protists comprising heterotrophic and photosynthetic representatives (Arisue et al. 2002; Silberman et al. 1996). Interestingly, *Blastocystis* sp. is the only Stramenopile known to commonly cause infection in humans. At the morphological level, four major forms of *Blastocystis* sp. have been described in stools and/or in vitro cultures: vacuolar, granular, amoeboid, and cystic

stages (for reviews, see Stenzel and Boreham 1996; Tan 2008). The two former forms are the most easily recognizable and frequently observed in laboratory culture and stool samples. Numerous data demonstrated that the water- and environmentally resistant infective cyst undoubtedly represents the transmissible stage of the parasite (Moe et al. 1997; Yoshikawa et al. 2004a). Taking into account these observations as those of in vitro encystation studies (Chen et al. 1999; Suresh et al. 1993), a life cycle for *Blastocystis* sp. was proposed with the cyst as the infectious stage (Tan 2008).

To date, *Blastocystis* sp. is the most common intestinal parasite found in human feces and considered an emerging parasite with a worldwide distribution. Its prevalence is by far higher than that of other unicellular intestinal parasites such as *Giardia*, *Entamoeba*, and *Cryptosporidium* (Amin 2002; Boorom et al. 2008; Östan et al. 2007; Su et al. 2009). It can reach 30% to 60% in developing countries and 1.5% to 20% in developed countries (for reviews, see Souppart et al. 2009; Tan 2008). This difference can be explained by poor hygiene practices, close animal contact, and consumption of contaminated food or water (Leelayoova et al. 2008; Li et al. 2007). Indeed, the fecal–oral route is considered to be the main mode of transmission of this parasite and several studies (Noël et al. 2005; Parkar et al. 2007, 2010; Yan et al. 2007; Yoshikawa et al. 2009) provided molecular-based evidence supporting the zoonotic potential of *Blastocystis* sp..

Blastocystis sp. isolates from humans and other animals have been reported to be morphologically indistinguishable. However, extensive genetic variation among numerous *Blastocystis* sp. isolates from both humans and animals has been mainly observed by PCR-restriction fragment length polymorphism and PCR using sequenced-tagged-site (STS) primers (for reviews, see Stensvold et al. 2007; Tan et al. 2008). This considerable genetic divergence among isolates was subsequently confirmed by molecular phylogenies mainly inferred from small subunit (SSU) rRNA gene sequences (Abe 2004; Arisue et al. 2003; Jones et al. 2009; Noël et al. 2005; Özyurt et al. 2008; Parkar et al. 2010; Scicluna et al. 2006; Souppart et al. 2009, 2010; Stensvold et al. 2009a; Whipps et al. 2010; Yoshikawa et al. 2009). From some of these molecular analyses, a consensus on *Blastocystis* sp. terminology was proposed (Stensvold et al. 2007), in which all human, mammalian, and avian isolates should be designated *Blastocystis* sp. and assigned to one of nine subtypes (STs: ST1 to ST9). In this new classification, each of the STs exhibited sufficient genetic diversity to be classified as separate species. Thereafter, a new ST was identified from primates and artiodactyls and designated as *Blastocystis* sp. ST10 (Stensvold et al. 2009a). More recently, three additional STs (ST11 to ST13) were identified from zoo animals (Parkar et al. 2010).

It remains widely debated in the literature whether *Blastocystis* sp. is a truly pathogenic organism (for reviews, see Boorom et al. 2008; Stensvold et al. 2009b; Tan 2008). This is mainly due to the fact that *Blastocystis* sp. can be found in both symptomatic and asymptomatic patients (Dogruman-Al et al. 2008; Eroglu et al. 2009; Souppart et al. 2009). However, accumulating recent in vitro and in vivo studies shed new light on the pathogenic power of this parasite (Elwakil and Hewedi 2010; Hussein et al. 2008; Puthia et al. 2005, 2006, 2008) and a model for pathogenesis of *Blastocystis* sp. was proposed (Tan 2008). Therefore *Blastocystis* sp. has been implicated in various gastrointestinal aspecific symptoms such as diarrhea, abdominal pain, nausea, fatigue, constipation, flatulence, and/or vomiting (for reviews, see Boorom et al. 2008; Stensvold et al. 2009c; Tan 2008), but this parasite may also play a significant role in several chronic gastrointestinal illnesses of unknown causes such as irritable bowel syndrome (IBS) and inflammatory bowel syndrome (IBD) (Dogruman-Al et al. 2009; Jones et al. 2009; Tai et al. 2010; Yakoob et al. 2004, 2010a, b; Yamamoto-Furusho and Torijano-Carrera 2010). In a recent study excluding patients with other intestinal pathogenic protists and bacteria, Dogruman-Al et al. (2010) reported that most IBS patients (76%) were infected with *Blastocystis* sp. Numerous cases were also reported regarding the association of *Blastocystis* sp. infection and urticaria (Hameed et al. 2010; Katsarou-Katsari et al. 2008). In addition, *Blastocystis* sp. has increasingly been implicated for diarrheal illness in immunocompromised individuals including HIV/AIDS and cancer patients (Tan et al. 2009). The studies that have addressed ST-dependent differences in pathogenicity in isolates from symptomatic and asymptomatic individuals (Dogruman-Al et al. 2008; Eroglu et al. 2009; Özyurt et al. 2008; Souppart et al. 2009; Stensvold et al. 2009c; Tan et al. 2008) have provided conflicting results mainly due to the limited number of patients examined in each epidemiological survey. However, in patients with IBS, IBD, and chronic diarrhea, either ST3 (Dogruman-Al et al. 2009; Jones et al. 2009) or ST1 (Yakoob et al. 2010a) were the most commonly found. Similarly, either ST3 (Hameed et al. 2010; Katsarou-Katsari et al. 2008) or ST2 (Vogelberg et al. 2010) were shown to be predominant in patients with urticaria.

Such a prevalence of *Blastocystis* sp. in the human population and its potential for transmission from animals naturally raise questions about the impact of this parasite in public health and the significance of its genetic diversity on pathogenicity of variants. Consequently, epidemiological studies have reported the frequency of STs from symptomatic and asymptomatic patients in numerous countries (for reviews, see Souppart et al. 2009; Tan 2008). These studies permit characterization of the transmission dynamics and

pathogenicity of the different *Blastocystis* sp. STs and tracking infection and contamination sources. However, in Italy, only preliminary epidemiological data are available with the molecular subtyping of no more than seven human *Blastocystis* sp. isolates (Mattiucci et al. 2010). Therefore, in the present study, genotypes of a large number of *Blastocystis* sp. clinical isolates were identified from symptomatic patients living in or in the vicinity of Rome and Sassari (Sardinia).

Materials and methods

Patient selection

Samples were collected at the Laboratory of Parasitology of the University of Rome Tor Vergata (24 samples analyzed in the present study: DM1 to DM24; Table 1) and Department of Biomedical Sciences of the University of Sassari (six

samples: DM25 to DM30) during parasitological examinations of stools from patients living in or in the vicinity of these two Italian cities. *Blastocystis* sp. was identified in stool samples through direct light microscopy of smears. All patients were classified as symptomatic due to the presence of digestive symptoms. Stool samples were also examined for the absence of other parasitic infections. No information was available on potential viral or bacterial infections.

DNA extraction and molecular subtyping of isolates

Genomic DNA of *Blastocystis* isolates was directly extracted from fecal samples using the QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. For each sample, 2 μl of genomic DNA was submitted to a PCR using the *Blastocystis* sp.-specific sense primer Blasto F (5′-TCTGGTTGATCCTGCCAGT-3′) and antisense primer Blasto R (5′-AGCTTTTTAACTG-CAACAACG-3′). These primers targeting the SSU rRNA

Table 1 *Blastocystis* sp. subtypes among symptomatic Italian patients

Patients	Age/sex	*Blastocystis* sp. subtype[a]	Nucleotide differences[b]	Accession no.
DM1	35/F	3	0	JF274658
DM2	77/F	2	1	JF274659, JF274660
DM3	20/M	3	0	JF274661
DM4	48/M	1	0	JF274662
DM5	59/M	2	1	JF274663, JF274664
DM6	51/M	4	0	JF274665
DM7	32/M	3	0	JF274666
DM8	11/M	1	0	JF274667
DM9	12/F	3	5	JF274668, JF274669
DM10	47/M	3	0	JF274670
DM11	24/M	2	2	JF274671, JF274672
DM12	48/M	4	0	JF274673
DM13	17/M	3	1	JF274674, JF274675
DM14	62/M	Mixed 2 and 8	–[c]	JF274676, JF274677
DM15	30/F	Mixed 2 and 3	–[c]	JF274678, JF274679
DM16	20/M	3	1	JF274680, JF274681
DM17	31/F	3	2	JF274682, JF274683
DM18	74/F	4	0	JF274684
DM19	24/F	3	3	JF274685, JF274686
DM20	35/F	Mixed 1 and 3	–[c]	JF274687, JF274688
DM21	65/F	4	1	JF274689, JF274690
DM22	59/M	4	0	JF274691
DM23	46/M	4	0	JF274692
DM24	11/M	2	0	JF274693
DM25	77/F	Mixed 3 and 7	–[c]	JF274694, JF274695
DM26	81/M	3	3	JF274696, JF274697
DM27	30/M	3	0	JF274698
DM28	34/M	3	1	JF274699, JF274700
DM29	48/M	2	2	JF274701, JF274702
DM30	41/F	3	0	JF274703

[a] According to the new standard terminology by Stensvold et al. (2007)

[b] Determined in the common region of two clones sequenced for each sample

[c] Not determined in mixed infections consisting of two clones representing two different subtypes

gene were derived from those described previously by Scicluna et al. (2006) and allowed the amplification of the same genomic region. PCR amplification was carried out in 50 µl according to standard conditions for Platinum *Taq* High-Fidelity DNA polymerase (Invitrogen). PCR conditions were similar to those described in previous studies (Souppart et al. 2009, 2010). Subsequently, PCR products were separated by agarose gel electrophoresis, and bands of the expected size (approximately 600 bp) were purified using the Wizard SV Gel and PCR clean-up system (Promega, Madison, WI, USA). Purified PCR products were cloned in the T-vector, pCR 2.1-TOPO (Invitrogen) and amplified in *Escherichia coli* One Shot TOP10 competent cells. Mini-preparations of plasmid DNA were done using the QIAprep Spin Miniprep kit (Qiagen). Two positive clones containing inserts of approximately the expected size were arbitrarily selected for each sample and sequenced on both strands. The SSU rRNA gene sequences obtained in this study have been deposited in GenBank under Accession numbers JF274658 to JF274703. These new sequences were aligned with the use of the BioEdit v7.0.1 package (http://www.mbio.ncsu.edu:BioEdit/bioedit.html) then compared with all the *Blastocystis* SSU rRNA gene sequences available from the National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov) using the BLAST program. STs were identified by determining the exact match or closest similarity against all known *Blastocystis* sp. STs according to the last classification by Stensvold et al. (2007).

Results and discussion

In this study, stool specimens were collected from a total of 30 Italian patients, 19 males and 11 females, ranging in age from 11 to 81 years (Table 1). These patients presented various gastrointestinal troubles, most commonly diarrhea, abdominal pain, flatulence, constipation, and/or vomiting, and were thus classified as symptomatic in the present study. For subtyping of *Blastocystis* sp. isolates, DNA was directly extracted from each stool sample then submitted to PCR amplification using a *Blastocystis* sp.-specific primers pair. This amplified domain of the SSU rDNA coding region has been shown to provide enough data to assign isolates to specific STs unambiguously (Scicluna et al. 2006) and was consequently used in recent molecular epidemiological surveys of French (Souppart et al. 2009) and Egyptian (Souppart et al. 2010) human populations. This fragment was eluted and cloned into the pCR 2.1-TOPO vector, and two clones were completely sequenced for each of the 30 samples. Each of the SSU rDNA gene sequences obtained in this study showed very high similarity (from 97% to 100%) to homologous sequences of the other *Blastocystis* sp. isolates reported so far. The

comparison with representative sequences of all known *Blastocystis* sp. STs using the BLAST software allowed the direct genotyping of the new isolates (Table 1), according to the consensus terminology by Stensvold et al. (2007). For 14 of the 30 samples, the two sequenced clones were identical (Table 1) while for 12 others, the number of nucleotide differences between the two clones ranged from 1 to 5. As previously discussed (Santin et al. 2011; Scicluna et al. 2006; Souppart et al. 2009, 2010), this low number of differences between the two clones of a same sample is likely due to sequence variations between SSU rDNA gene copies within the same isolate. Therefore, we have considered that two different clones of the same ST identified in the same patient derived from the same strain of *Blastocystis* sp. However, the possibility of mixed infection with two distinct isolates of the same ST cannot be completely excluded. For the last four samples (DM14, DM15, DM20, and DM25; Table 1), we identified mixed infections containing two *Blastocystis* sp. STs suggesting multiple sources of infection in these patients. The prevalence of 13.3% of mixed infections in our study was roughly similar to that described in different countries and comprised between 2.6% and 14.3% (Böhm-Gloning et al. 1997; Dogruman-Al et al. 2008, 2009; Li et al. 2007; Souppart et al. 2009, 2010; Stensvold et al. 2009c). In these latter studies, most of the mixed infections were with ST1 and ST3, ST1 and ST2, or ST2 and ST3 which was expected because of the prevalence of these three STs in the respective populations. In the present analysis, we also identified two mixed infections with ST1 and ST3 and ST2 and ST3. However, the two others were not common with ST2 and ST8 and ST3 and ST7, respectively. The true distribution of mixed infections remains difficult to ascertain in a particular individual and likely underestimated as this depends on the method employed for subtyping. Indeed, in studies based on direct sequencing of PCR product, mixed infections result in mixed peaks in sequencing chromatograms which may be difficult to interpret (Parkar et al. 2007). In genotyping analyses based on the cloning and sequencing of the PCR product as in the present study, only two clones are routinely sequenced for each sample (Souppart et al. 2009, 2010). Recently, by sequencing up to ten clones, Santin et al. (2011) have identified three different STs in a primate isolate. Moreover, genotyping of parasites after in vitro propagation risks underestimating mixed infections since certain STs may outgrow others (Parkar et al. 2007; Yan et al. 2007). In addition, PCR-based genotype classification method using known STS primers can also be a source of underestimation since only seven STs (ST1 to ST7) can be identified so far through this approach (Yoshikawa et al. 2004b).

The majority of the samples included in this study (26/30) as well as in all the epidemiological surveys available

in the literature from numerous countries, represented single infections. With the addition of four mixed infections consisting of two different isolates, we analyzed a total of 34 isolates. As summarized in Table 1, ST3 was the most common in our Italian population (47.1%) followed by ST2 (20.6%), ST4 (17.7%), ST1 (8.8%), and ST7 and ST8 (2.9%). The ST distribution in the present study was quite similar to that found in other countries. Indeed, in almost all the studies reported so far, a large majority of human infections with *Blastocystis* sp. were attributable to ST3 isolates (an average of 61.2% according to Souppart et al. 2009; for recapitulative tables describing the distribution of *Blastocystis* sp. STs infecting humans in different geographic regions see Souppart et al. 2009 and Tan et al. 2008). Interestingly, although the number of isolates obtained in Sardinia was low (seven isolates from the six samples DM25 to DM30), the ST3 was dominant in this region (5/7 isolates) as this was also the case in the Rome area (11/27 isolates). Only few exceptions showed the higher prevalence of other STs such as ST4 and ST1 in Spain (Dominguez-Marquez et al. 2009) and Thailand (Thathaisong et al. 2003), respectively. However, data obtained from these two latter studies have to be considered with caution since it has been recently revealed that one of the primers used in their genotyping approach (reverse primer R1 annealing to a region of the SSU rRNA gene; Böhm-Gloning et al. 1997) might preferentially amplify some STs over others (Wong et al. 2008). The proportions of STs other than ST3 clearly differ between locations. In Italy as for instance in Turkey and Denmark, ST2 is the second most common variant while this is ST1 in France, Germany, Greece, Egypt, China, Bangladesh, Pakistan, and Singapore (Souppart et al. 2009). As stated above, ST4 could be dominant in Spain and is also commonly found in Denmark, Germany, and Italy. In the present study, we identified only one isolate belonging to ST7 and one other to ST8. This confirms that ST7 is common in Asia (Japan, China) and North Africa (Egypt) but rarely observed in European countries. Regarding ST8, this ST was only found so far in the UK and United States (Stensvold et al. 2009a; Whipps et al. 2010). All these data reinforce the hypothesis that ST1–ST4 isolates are largely dominant in all human populations worldwide while the others STs (ST5 to ST13) are found episodically. In addition, all the epidemiological surveys including the present study indicate that there is no correlation between *Blastocystis* sp. geographic origin and ST.

In Italy, Mattiucci et al. (2010) have recently performed the molecular genotyping of seven human *Blastocystis* sp. isolates obtained in the Rome area. These isolates were classified as ST1 (three isolates), ST5 (two isolates), and ST7 (two isolates). However, although the number of isolates included in this latter study was very low, the ST distribution was uncorrelated with that observed in the present analysis. In particular, Mattiucci et al. (2010) did not identify ST2 and ST3 isolates, while these two STs are nevertheless dominant and represent approximately 70% of our isolates. As stated above, the use by Mattiucci et al. (2010) of the R1 primer in their genotyping method could likely explain the differences in the ST distribution observed between the two studies. In their report, Mattiucci et al. (2010) indicated that most of the infected patients (6/7) were symptomatic. Consequently, together with our data, a total of seven STs (ST1, ST2, ST3, ST4, ST5, ST7, and ST8) have been identified in Italian symptomatic patients. Similarly, a large diversity of STs (currently four to six STs) was also found in several European, American, and Asian symptomatic populations (Souppart et al. 2009; Whipps et al. 2010). These epidemiological data also reinforce the hypothesis of the existence of pathogenic and non-pathogenic variants in different STs as experimentally demonstrated in animal models (Hussein et al. 2008).

The differences reported in the number of STs identified in human populations as well as their relative proportions might indicate different reservoirs and routes of transmission. The dominant ST3 was suggested to be the only ST of human origin, while the remaining STs were likely zoonotic according to their moderate host specificity (Noël et al. 2005; Stensvold et al. 2009a; Tan 2008). Consequently, the predominance of ST3 in Italy as in other countries might be mainly explained by large-scale human-to-human transmission (Yoshikawa et al. 2000). Concerning ST1–ST2 and ST4, their presence in the Italian population might be linked to zoonotic transmission mainly from farm animals/pigs/dogs and domestic rodents, respectively (Noël et al. 2005; Stensvold et al. 2009a; Tan 2008). Moreover, ST7 is primarily associated with avian isolates and given its apparent host specificity, it is highly likely that human infections due to such avian ST are not common and of zoonotic origin (Noël et al. 2005; Stensvold et al. 2009a; Tan 2008). Regarding ST5 isolates identified by Mattiucci et al. (2010) in two Italian patients, a zoonotic transmission from pigs/cattle is also highly probable since this ST is the dominant genotype in these animals (Noël et al. 2005; Stensvold et al. 2009c; Yan et al. 2007). Finally, ST8 has been isolated in one Italian patient. This ST is normally very rare in humans but is common in primates (Stensvold et al. 2009a). Interestingly, ST8 is also common in primate handlers as described by Stensvold et al. (2009a), suggesting a zoonotic transmission from primates to humans.

To our knowledge, this is the first large investigation into the molecular genotyping of human *Blastocystis* sp. isolates in Italy. In this country, by taking into account the preliminary data of Mattiucci et al. (2010), we revealed the presence of seven different known STs of variable prevalence in symptomatic patients and found a higher

prevalence of ST3, as in other regions of the world. According to our observations, it may be suggested that *Blastocystis* sp. infection in Italy can be due to different STs, most of them being of animal origin. Therefore, on the basis of the present study, future attempts to identify specific risk factors for human infection with different *Blastocystis* sp. STs have to be performed in Italy particularly in the aim at exploring the potential association of STs with specific animal reservoirs.

References

Abe N (2004) Molecular and phylogenetic analysis of *Blastocystis* isolates from various hosts. Vet Parasitol 120:235–242

Abe N, Nagoshi M, Takami K, Sawano Y, Yoshikawa H (2002) A survey of *Blastocystis* sp. in livestock, pets, and zoo animals in Japan. Vet Parasitol 106:203–212

Amin OM (2002) Seasonal prevalence of intestinal parasites in the United States during 2000. Am J Trop Med Hyg 66:799–803

Arisue N, Hashimoto T, Yoshikawa H, Nakamura Y, Nakamura G, Nakamura F, Yano TA, Hasegawa M (2002) Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. J Eukaryot Microbiol 49:42–53

Arisue N, Hashimoto T, Yoshikawa H (2003) Sequence heterogeneity of the small subunit ribosomal RNA genes among *Blastocystis* isolates. Parasitology 126:1–9

Böhm-Gloning B, Knobloch J, Walderich B (1997) Five subgroups of *Blastocystis hominis* isolates from symptomatic and asymptomatic patients revealed by restriction site analysis of PCR-amplified 16S-like rDNA. Trop Med Int Health 2:117–118

Boorom KF, Smith H, Nimri L, Viscogliosi E, Spanakos G, Parkar U, Li LH, Zhou XN, Ok UZ, Leelayoova S, Jones MS (2008) Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection. BMC Parasit Vectors 1:40

Chen XQ, Singh M, Howe J, Ho LC, Tan SW, Yap EH (1999) In vitro encystation and excystation of *Blastocystis ratti*. Parasitology 118:151–160

Dogruman-Al F, Dragci H, Yoshikawa H, Kurt Ö, Demirel M (2008) A possible link between subtype 2 and asymptomatic infections of *Blastocystis hominis*. Parasitol Res 103:685–689

Dogruman-Al F, Kustimur S, Yoshikawa H, Tuncer C, Simsek Z, Tanyuksel M, Araz E, Boorom K (2009) *Blastocystis* subtypes in irritable bowel syndrome and inflammatory bowel disease in Ankara, Turkey. Mem Inst Oswaldo Cruz 104:724–727

Dogruman-Al F, Simsek Z, Boorom K, Ekici E, Sahin M, Tuncer C, Kustimur S, Altinbas A (2010) Comparison of methods for detection of *Blastocystis* infection in routinely submitted stool samples, and also in IBS/IBD Patients in Ankara, Turkey. PLoS ONE 5:e15484

Dominguez-Marquez MV, Guna R, Munoz C, Gomez-Munoz MT, Borras R (2009) High prevalence of subtype 4 among isolates of *Blastocystis hominis* from symptomatic patients of a health district of Valencia (Spain). Parasitol Res 105:949–955

Elwakil HS, Hewedi HI (2010) Pathogenic potential *of Blastocystis hominis* in laboratory mice. Parasitol Res 107:685–689

Eroglu F, Genc A, Elgun G, Koltas IS (2009) Identification of *Blastocystis hominis* isolates from asymptomatic and symptomatic patients by PCR. Parasitol Res 105:1589–1592

Hameed DM, Hassanin OM, Zuel-Fakkar NM (2010) Association of *Blastocystis hominis* genetic subtypes with urticaria. Parasitol Res. doi:10.1007/s00436-010-2097-2

Hussein EM, Hussein AM, Eida MM, Atwa MM (2008) Pathophysiological variability of different genotypes of human *Blastocystis hominis* egyptian isolates in experimentally infected rats. Parasitol Res 102:853–860

Jones MS, Whipps CM, Ganac RD, Hudson NR, Boroom K (2009) Association of *Blastocystis* subtype 3 and 1 with patients from an Oregon community presenting with chronic gastrointestinal illness. Parasitol Res 104:341–345

Katsarou-Katsari A, Vassalos CM, Tzanetou K, Spanakos G, Papadopoulou C, Vakalis N (2008) Acute urticaria associated with amoeboid forms of *Blastocystis* sp. subtype 3. Acta Derm Venereol 88:80–81

Leelayoova S, Siripattanapipong S, Thathaisong U, Naaglor T, Taamasri P, Piyaraj P, Mungthin M (2008) Drinking water: a possible source of *Blastocystis* spp. subtype 1 infection in schoolchildren of a rural community in central Thailand. Am J Trop Med Hyg 79:401–406

Li LH, Zhou XN, Du ZW, Wang XZ, Wang LB, Jiang JY, Yoshikawa H, Steinmann P, Utzinger J, Wu Z, Chen JX, Chen SH, Zhang L (2007) Molecular epidemiology of human *Blastocystis* in a village in Yunnan province, China. Parasitol Int 56:281–286

Mattiucci S, De Rosa A, Gabrielli S, Iori A, Cancrini G (2010) Molecular identification of *Blastocystis* sp. subtypes from humans in Italy: preliminary results. Parassitologia 52:207

Moe KT, Singh M, Howe J, Ho LC, Tan SW, Chen XQ, Ng GC, Yap EH (1997) Experimental *Blastocystis hominis* infection in laboratory mice. Parasitol Res 83:319–325

Noël C, Dufernez F, Gerbod D, Edgcomb VP, Delgado-Viscogliosi P, Ho LC, Singh M, Wintjens R, Sogin ML, Capron M, Pierce R, Zenner L, Viscogliosi E (2005) Molecular phylogenies of *Blastocystis* isolates from different hosts: implications for genetic diversity, identification of species, and zoonosis. J Clin Microbiol 43:348–355

Östan I, Kilimcioglu AA, Girginkardesler N, Özyurt BC, Limoncu ME, Ok UZ (2007) Health inequities: lower socio-economic conditions and higher incidences of intestinal parasites. BMC Public Health 7:342

Özyurt M, Kurt O, Molbak K, Nielsen HV, Haznedaroglu T, Stensvold CR (2008) Molecular epidemiology of *Blastocystis* infections in Turkey. Parasitol Int 57:300–306

Parkar U, Traub RJ, Kumar S, Mungthin M, Vitali S, Leelayoova S, Morris K, Thompson RCA (2007) Direct characterization of *Blastocystis* from faeces by PCR and evidence of zoonotic potential. Parasitology 134:359–367

Parkar U, Traub RJ, Vitali S, Elliot A, Levecke B, Robertson I, Geurden T, Steele J, Drake B, Thompson RC (2010) Molecular characterization of *Blastocystis* isolates from zoo animals and their animal-keepers. Vet Parasitol 169:8–17

Puthia MK, Vaithilingam A, Lu J, Tan KS (2005) Degradation of human secretory immunoglobulin A by *Blastocystis*. Parasitol Res 97:386–389

Puthia MK, Sio SW, Lu J, Tan KS (2006) *Blastocystis ratti* induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells. Infect Immun 74:4114–4223

Puthia MK, Lu J, Tan KS (2008) *Blastocystis ratti* contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner. Eukaryot Cell 7:435–443

Santin M, Gomez-Munoz MT, Solano-Aguilar G, Fayer R (2011) Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. Parasitol Res. doi:10.1007/s00436-010-2097-2

Scicluna SM, Tawari B, Clark CG (2006) DNA barcoding of *Blastocystis*. Protist 157:77–85

Silberman JD, Sogin ML, Leipe DD, Clark CG (1996) Human parasite finds taxonomic home. Nature 380:398

Souppart L, Sanciu G, Cian A, Wawrzyniak I, Delbac F, Capron M, Dei-Cas E, Boorom K, Delhaes L, Viscogliosi E (2009) Molecular epidemiology of human *Blastocystis* isolates in France. Parasitol Res 105:413–421

Souppart L, Moussa H, Cian A, Sanciu G, Poirier P, El Alaoui H, Delbac F, Boorom K, Delhaes L, Dei-Cas E, Viscogliosi E (2010) Subtype analysis of *Blastocystis* isolates from symptomatic patients in Egypt. Parasitol Res 106:505–511

Stensvold CR, Suresh GK, Tan KSW, Thompson RCA, Traub RJ, Viscogliosi E, Yoshikawa H, Clark CG (2007) Terminology for *Blastocystis* subtypes—a consensus. Trends Parasitol 23:93–96

Stensvold CR, Alfellani MA, Nørskov-Lauritsen S, Prip K, Victory EL, Maddox C, Nielsen HV, Clark CG (2009a) Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. Int J Parasitol 39:473–479

Stensvold CR, Nielsen HV, Mølbak K, Smith HV (2009b) Pursuing the clinical significance of *Blastocystis*—diagnostic limitations. Trends Parasitol 25:23–29

Stensvold CR, Lewis HC, Hammerum AM, Porsbo LJ, Nielsen SS, Olsen KE, Arendrup MC, Nielsen HV, Molbak K (2009c) *Blastocystis*: unravelling potential risk factors and clinical significance of a common but neglected parasite. Epidemiol Infect 137:1655–1663

Stenzel DJ, Boreham PFL (1996) *Blastocystis hominis* revisited. Clin Microbiol Rev 9:563–584

Su F-H, Chu F-Y, Li C-Y, Tang H-F, Lin Y-S, Peng Y-J, Su Y-M, Lee S-D (2009) *Blastocystis hominis* infection in long-term care facilities in Taiwan: prevalence and associated clinical factors. Parasitol Res 105:1007–1013

Suresh K, Ng GC, Ramachandran NP, Ho LC, Yap EH, Singh M (1993) In vitro encystment and experimental infections of *Blastocystis hominis*. Parasitol Res 79:456–460

Tai WP, Hu PJ, Wu J, Lin XC (2010) Six ulcerative colitis patients with refractory symptoms co-infective with *Blastocystis hominis* in China. Parasitol Res. doi:10.1007/s00436-010-2164-8

Tan KSW (2004) *Blastocystis* in humans and animals: new insights using modern methodologies. Vet Parasitol 126:121–144

Tan KSW (2008) New insights on classification, identification, and clinical relevance of *Blastocystis* spp. Clin Microbiol Rev 21:639–665

Tan TC, Suresh KG, Smith HV (2008) Phenotypic and genotypic characterization of *Blastocystis hominis* isolates implicates subtype 3 as a subtype with pathogenic potential. Parasitol Res 104:85–93

Tan TC, Ong SC, Suresh KG (2009) Genetic variability of *Blastocystis sp.* isolates obtained from cancer and HIV/AIDS patients. Parasitol Res 105:1283–1286

Thathaisong U, Worapong J, Mungthin M, Tan-Ariya P, Viputtigul K, Sudatis A, Noonai A, Leelayoova S (2003) *Blastocystis* isolates from a pig and a horse are closely related to *Blastocystis hominis*. J Clin Microbiol 41:967–975

Vogelberg C, Stensvold CR, Monecke S, Ditzen A, Stopsack K, Heinrich-Grafe U, Pohlmann C (2010) *Blastocystis* sp. subtype 2 detection during recurrence of gastrointestinal and urticarial symptoms. Parasitol Int 59:469–471

Whipps CM, Boorom K, Bermudez LE, Kent ML (2010) Molecular characterization of *Blastocystis* species in Oregon identifies multiple subtypes. Parasitol Res 106:827–832

Wong KHS, Ng GC, Lin RTP, Yoshikawa H, Taylor MB, Tan KSW (2008) Predominance of subtype 3 among *Blastocystis* isolates from a major hospital in Singapore. Parasitol Res 102:663–670

Yakoob J, Jafri W, Jafir N, Khan R, Islam M, AsimBeg M, Zaman V (2004) Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*. Am J Trop Med Hyg 70:383–385

Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, Islam M, Khan R (2010a) Irritable bowel syndrome: is it associated with genotypes of *Blastocystis hominis*. Parasitol Res 106:1033–1038

Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, Islam M, Khan R (2010b) *Blastocystis hominis* and *Dientamoeba fragilis* in patients fulfilling irritable bowel syndrome criteria. Parasitol Res 107:679–684

Yamamoto-Furusho JK, Torijano-Carrera E (2010) Intestinal protozoa infections among patients with ulcerative colitis: prevalence and impact on clinical disease course. Digestion 82:18–23

Yan Y, Su S, Ye J, Lai X, Lai R, Liao H, Chen G, Zhang R, Hou Z, Luo X (2007) *Blastocystis* sp. subtype 5: a possibly zoonotic genotype. Parasitol Res 101:1527–1532

Yoshikawa H, Abe N, Iwasawa M, Kitano S, Nagano I, Wu Z, Takahashi Y (2000) Genomic analysis of *Blastocystis hominis* strains isolated from two long-term health care facilities. J Clin Microbiol 38:1324–1330

Yoshikawa H, Yoshida K, Nakajima A, Yamanari K, Iwatani S, Kimata I (2004a) Fecal–oral transmission of the cyst form of *Blastocystis hominis* in rats. Parasitol Res 94:391–396

Yoshikawa H, Wu Z, Kimata I, Iseki M, Ali IKMD, Hossain MB, Zaman V, Haque R, Takahashi Y (2004b) Polymerase chain reaction-based genotype classification among human *Blastocystis hominis* populations isolated from different countries. Parasitol Res 92:22–29

Yoshikawa H, Wu Z, Pandey K, Pandey BD, Sherchand JB, Yanagi T, Kanbara H (2009) Molecular characterization of *Blastocystis* isolates from children and rhesus monkeys in Kathmandu, Nepal. Vet Parasitol 160:295–300

Part IV

_____

MULTIPLE COINFECTIONS WITH *BLASTOCYSTIS* SP. IN A

FRENCH PATIENT

IV-1 <u>Introduction</u>

Molecular PCR-based diagnostic approaches for *Blastocystis* sp. identification have been described. ST-specific diagnostic primers, also referred to as STS primers, were developed from random amplification of polymorphic DNA analysis of *Blastocystis* sp. isolates (Yoshikawa et al. 1998). However, this method allows only the amplification of 7 STs of the 13 identified so far. Other groups characterized isolates by PCR of the SSU rRNA gene followed by Restriction Fragment Length Polymorphism (RFLP) or cloning of the PCR product coupled with dideoxy sequencing of a limited number of clones. Because of their limitations, such genotyping methods likely underestimate the prevalence of mixed intra- and inter-ST infections leading to erroneous results and conclusions in the context of epidemiological studies. To test this hypothesis, we analyzed the genetic diversity of *Blastocystis* sp. isolates in a French patient considered at high risk of mixed infection through his lifestyle in rural area and long history of travelling.

IV-2 <u>Results</u>

Article 6 (see above)

<u>Meloni D</u>, Poirier P, Mantini C, Noël C, Gantois N, Wawrzyniak I, El Alaoui H, Chabé M, Delhaes L, Dei-Cas E, Fiori PL, Delbac F, Viscogliosi E. Mixed human intra- and inter-subtype infections with the parasitic Stramenopile *Blastocystis* sp. Journal of Clinical Microbiology, 2011, submitted

IV-3 <u>Conclusions</u>

To our knowledge, this is the first evidence of mixed intra- and inter-ST infections with *Blastocystis* sp. in a same individual. The sequencing of a large number of clones allows us identifying several isolates belonging to each of three detected STs (ST2, ST3, and ST4). According to her lifestyle and history of travelling, multiple potential contamination sources are possible that could naturally explain the identification of a large number of *Blastocystis* sp. isolates. In further molecular epidemiological studies, patients at risk of mixed infection should be identified in order to optimize the genotyping approach as in the present study by increasing the number of sequenced clones.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

SHORT-FORM PAPER

TITLE: Mixed human intra- and inter-subtype infections with the parasitic Stramenopile *Blastocystis* sp.

RUNNING TITLE: Mixed human infections with *Blastocystis* sp.

Dionigia Meloni,[1,2] Philippe Poirier,[3] Cléa Mantini,[1] Christophe Noël,[4] Nausicaa Gantois,[1] Ivan Wawrzyniak,[3] Hicham El Alaoui,[3] Magali Chabé,[1] Laurence Delhaes,[1] Eduardo Dei-Cas,[1] Pier Luigi Fiori,[2] Frederic Delbac,[3] and Eric Viscogliosi[1*]

Institut Pasteur of Lille, Center for Infection and Immunity of Lille (CIIL), Inserm U1019, CNRS UMR 8204, University Lille Nord de France, EA4547, Lille Cedex, France[1];

Department of Biomedical Sciences, Division of Experimental and Clinical Microbiology, University of Sassari, Sassari, Italy[2];

Laboratoire Microorganismes: Génome et Environnement, CNRS UMR 6023, Clermont Université, Université Blaise Pascal, Aubière, France[3];

Geneius Laboratories Ltd, INEX Business Centre, Newcastle upon Tyne, UK[4]

* Corresponding author. Mailing address: Institut Pasteur of Lille, Center for Infection and Immunity of Lille (CIIL), Inserm U1019, CNRS UMR 8204, University Lille-Nord de France, Biology and Diversity of Emerging Eukaryotic Pathogens, EA4547, 1 rue du Professeur Calmette, BP 245, 59019 Lille cedex, France. Phone: 33 3 20 87 79 61. Fax: 33 3 20 87 78 88. E-mail: eric.viscogliosi@pasteur-lille.fr

ABSTRACT

Because of their limitations, current subtyping methods likely underestimate mixed human intra- and inter-subtype infections with *Blastocystis* sp. leading to erroneous data in the context of epidemiological studies. We confirmed this hypothesis by the identification of several isolates belonging to three subtypes in a patient considered at high risk of mixed infection through her lifestyle in rural area and long history of travelling.

KEYWORDS: *Blastocystis* sp.; Intestinal parasite; Mixed infection; Molecular epidemiology; Subtype

Blastocystis sp. is an enteric protozoa commonly occurring in a wide range of animals (17). It has a worldwide distribution and is often the most common intestinal parasite reported in human faecal samples (2). The main mode of transmission of this parasite is the faecal-oral route through consumption of contaminated food and water (17). Its pathogenic potential in humans remained debated because the infection can be asymptomatic. However, recent *in vitro* and *in vivo* data together with those of the analysis of its genome, allowed proposing a model for pathogenesis of this parasite (3, 17). Accumulating reports suggest an association between *Blastocystis* sp. and a variety of gastrointestinal disorders, including irritable bowel syndrome (4) and acute urticarial lesions associated with minor digestive symptoms (6). This parasite has also increasingly been implicated for diarrheal illness in immunocompromised individuals including HIV and cancer patients (18). Although morphologically indistinguishable, phylogenies inferred from SSU rRNA gene sequences revealed a considerable genetic divergence among *Blastocystis* sp. isolates from humans and animals and a total of 13 subtypes (STs) have so far been identified (10, 15, 16). A large majority of human infections with *Blastocystis* sp. were attributable to ST3 isolates (14) suggesting that ST3 represents the only ST of human origin (9). All the others STs of supposed animal origin are zoonotic and almost all are able to infect humans in different frequencies. Therefore a higher risk of *Blastocystis* sp. infection has been shown in people living in rural area and/or with close animal contact (19, 22).

Numerous epidemiological studies have reported the frequency of STs from symptomatic and asymptomatic patients in several countries (14, 17) in the aim to identify a possible link between ST and pathogenicity and to track infection and contamination sources. For subtyping of *Blastocystis* sp. isolates, polymerase chain reaction (PCR), employing ST-specific primers, is a commonly used method but it allows only the amplification of 7/13 STs

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

(20). Isolates can also be characterized by amplification of informative domains of the SSU rRNA gene using *Blastocystis* sp.-specific primers followed by Restriction Fragment Length Polymorphism (RFLP) or cloning of the PCR product coupled with sequencing of a limited number of clones (1, 14). Because of their limitations, such subtyping methods likely underestimate the prevalence of mixed infections with more than one *Blastocystis* sp. ST as well as with isolates deriving from the same ST in a particular individual. Consequently these methodological drawbacks could seriously distort the results and understanding of epidemiological studies.

To test this hypothesis, we analyzed the genetic diversity of *Blastocystis* sp. isolates in a French patient, a 60-years old woman, considered at high potential risk of mixed infection with this parasite. Indeed, the patient showed several risk factors for contamination such as living in a rural area, eating fruits and vegetables from her garden or picked up in the country, and buying food in local markets. Moreover, she has a long history of travelling (Malta, Cuba, USA, Maghreb, Réunion Island) during the five last years and eats local food during her trips. As food-handler in Hospital, the patient without any gastrointestinal or skin symptoms, presented for a routine parasitic stool examination. *Blastocystis* sp. was identified through direct light microscopy (x400) of wet smears. DNA extraction from stool sample and PCR using *Blastocystis* sp.-specific primers and High-fidelity *Taq* DNA polymerase were performed as previously described (8, 14). The 600 bp-amplified domain of the SSU rDNA coding region has been shown to provide sufficient information for accurate subtyping (13). The PCR fragment was purified, cloned into the pCR 2.1-TOPO vector (Invitrogen, Groningen, The Netherlands), and amplified in TOP10 competent cells. Fifty positive clones (Cl1 to Cl50) and not only two as in previous studies (8, 14) were arbitrarily selected and sequenced. These SSU rRNA gene sequences were aligned with available sequences of *Blastocystis* sp. isolates representative of ST1 to ST10 using the BioEdit v7.0.1 package. ST11 to ST13 SSU rRNA sequences were not available for this particular domain (10). Maximum likelihood phylogenetic analysis of sequences was performed with PhyML 3.0 (5) using the GTR (general time reversible) + ☐ (gamma distribution of rates with four rate categories) + I (proportion of invariant sites) model of evolution determined with ModelTest 3.7 (12). Bootstrap values were obtained from 1,000 replicates.

The STs identified in our analysis as well as their relationships (Fig. 1) were identical to those described in previous phylogenetic studies including a large sampling of *Blastocystis* sp. isolates and based on complete SSU rRNA gene sequences (7, 9) or inferred from sequences of the present SSU rRNA amplified-gene domain (11, 16). In our tree, three

sequences corresponding to Chinese human isolates and named STX clearly belonged to ST7 and did not form an additional ST as previously suggested (7). We also noted that the newly identified ST10 (16) was closely related to ST8 casting doubt about its justification. Analyzing the sequence data obtained in the present study, we revealed that the patient was infected by three *Blastocystis* sp. STs. Indeed, 38 (76%), 6 (12%), and 6 (12%) out of the 50 clones derived from ST3, ST2, and ST4, respectively. Clustering of the 50 sequences was also confirmed by BLAST analysis. Since primers used in PCR amplification are not known to be selective for a specific ST, the present results strongly suggested that *Blastocystis* sp. ST3 parasites were predominant in the patient. According to our alignment, ST3 clones showed 86.9 to 100% identity (591 shared nucleotide positions), indicating a large intra-ST variation. Intra-ST2 and intra-ST4 variations were more limited, since ST2 and ST4 clones showed 96 to 99.6% and 97.1 to 99.8% identity, respectively. Obviously, these variations cannot be only due to known sequence differences between SSU rRNA gene copies within the same *Blastocystis* sp. isolates. Indeed, 38 SSU rDNA gene copies have been identified in the genome of a *Blastocystis* sp. ST7 isolate (3) and complete sequences are available for 22 of these copies. In the gene domain common to that used in our study (575 shared positions), SSU rRNA gene copies only differed by up to nine nucleotides to each other i.e. until 1.6% difference (Fig. 2). Consequently, even if information regarding the molecular heterogeneity in the SSU rDNA coding region of other STs is not available yet, our data strongly suggested that the patient was infected by several isolates of each detected ST (ST2, ST3, ST4). Moreover, intra-ST sequence polymorphism may naturally reflect the potential existence of subgroups within the same ST. According to the usually accepted source of parasites of these STs, ST3 multi-infections of the patient might be rather explained by human-to-human transmission (21) whereas the presence of ST2 and ST4 isolates might be linked to zoonotic transmission from primates, farm animals, or domestic rodents (9, 17).

To our knowledge, this is the first evidence of mixed intra- and inter-ST infections with *Blastocystis* sp. in a same individual. According to his lifestyle and travelling history, this patient had a high potential risk of mixed infections with this parasite through multiple potential contamination sources. This hypothesis is consistent with the identification of several isolates belonging to each of three detected STs. In the context of molecular epidemiological studies, patients at risk of mixed infection should be identified in order to optimize the genotyping approach by sequencing a large number of clones. In the future, with the rise of comparative genomics, development of Multi Locus Sequence Typing (MLST)

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

methods could optimally assess the diversity of *Blastocystis* sp. isolates from individual patients.

Nucleotide sequence accession numbers. The SSU rRNA gene sequences obtained in this study have been deposited in GenBank under accession numbers JN942527 to JN942576.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

REFERENCES

1. **Abe, N., Z. Wu, and H. Yoshikawa.** 2003. Molecular characterization of *Blastocystis* isolates from birds by PCR with diagnostic primers and restriction fragment length polymorphism analysis of the small subunit ribosomal RNA gene. Parasitol. Res. 89:393-396.

2. **Boorom, K. F., H. Smith, L. Nimri, E. Viscogliosi, G. Spanakos, U. Parkar, L.-H. Li, X.-N. Zhou, U. Z. Ok, S. Leelayoova, and M. S. Jones.** 2008. Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection. BMC Parasit. Vectors 1:40.

3. **Denoeud, F., M. Roussel, B. Noël, I. Wawrzyniak, C. Da Silva, M. Diogon, E. Viscogliosi, C. Brochier-Armamet, A. Couloux, J. Poulain, B. Segurans, V. Anthouard, C. Texier, N. Blot, P. Poirier, G. C. Ng, K. S. W. Tan, F. Antiguenave, O. Jaillon, J. Aury, F. Delbac, P. Wincker, C. P. Vivarès, and H. El Alaoui.** 2011. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol. 12:R29.

4. **Dogruman-Al, F., Z. Simsek, K. Boorom, E. Ekici, M. Sahin, C. Tuncer, S. Kustimur, and A. Altinbas.** 2010. Comparison of methods for detection of *Blastocystis* infection in routinely submitted stool samples, and also in IBS/IBD Patients in Ankara, Turkey PLoS ONE 5:e15484.

5. **Guindon, S., F. Lethiec, P. Duroux, and O. Gascuel.** 2005. PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 33:W557-559.

6. **Katsarou-Katsari, A., C. M. Vassalos, K. Tzanetou, G. Spanakos, C. Papadopoulou, and N. Vakalis.** 2008. Acute urticaria associated with amoeboid forms of *Blastocystis* sp. subtype 3. Acta Derm. Venereol. 88:80-81.

7. **Li, L.-H., X.-P. Zhang, S. Lv, L. Zhang, H. Yoshikawa, Z. Wu, P. Steinmann, J. Utzinger, X.-M. Tong, S.-H. Chen, and X.-N. Zhou.** (2007). Cross-sectional surveys and subtype classification of human *Blastocystis* isolates from four epidemiological settings in China. Parasitol. Res. 102:83-90.

8. **Meloni, D., G. Sanciu, P. Poirier, H. El Alaoui, M. Chabé, L. Delhaes, E. Dei-Cas, F. Delbac, P. L. Fiori, D. Di Cave, and E. Viscogliosi.** 2011 Molecular subtyping of *Blastocystis* isolates from symptomatic patients in Italy. Parasitol. Res. 109:613-619.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

9. **Noël, C., F. Dufernez, D. Gerbod, V. P. Edgcomb, P. Delgado-Viscogliosi, L. C. Ho, M. Singh, R. Wintjens, M. L. Sogin, M. Capron, R. Pierce, L. Zenner, and E. Viscogliosi.** 2005. Molecular phylogenies of *Blastocystis* isolates from different hosts: implications for genetic diversity, identification of species, and zoonosis. J. Clin. Microbiol. 43:348-355.

10. **Parkar, U., R. J. Traub, S. Vitali, A. Elliot, B. Levecke, I. Robertson, T. Geurden, J. Steele, B. Drake, and R. C. Thompson.** 2010. Molecular characterization of *Blastocystis* isolates from zoo animals and their animal-keepers. Vet. Parasitol. 169:8-17.

11. **Petrasova, J., M. Uzlikova, M. Kostka, K. J. Petrzelkova, M. A. Huffman, and D. Modry.** 2011. Diversity and host specificity of *Blastocystis* in syntopic primates on Rubondo Island, Tanzania. Int. J. Parasitol. 41:1113-1120.

12. **Posada, D., and K. A. Crandall KA.** 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817-818.

13. **Scicluna, S. M., B. Tawari, and C. G. Clark.** 2006. DNA barcoding of *Blastocystis*. Protist 157:77-85.

14. **Souppart, L., G. Sanciu, A. Cian, I. Wawrzyniak, F. Delbac, M. Capron, E. Dei-Cas, K. Boorom, L. Delhaes, and E. Viscogliosi.** 2009. Molecular epidemiology of human *Blastocystis* isolates in France. Parasitol. Res. 105:413-421.

15. **Stensvold, C. R., G. K. Suresh, K. S. Tan, K, R. C. Thompson, R. J. Traub, E. Viscogliosi, H. Yoshikawa, and C. G. Clark.** 2007. Terminology for *Blastocystis* subtypes - a consensus. Trends Parasitol. 23:93-96.

16. **Stensvold, C. R., M. A. Alfellani, S. Norskov-Lauritsen, K. Prip, E. L. Victory, C. Maddox, H. V. Nielsen, and C. G. Clark. 2009.** Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. Int. J. Parasitol. 39:473-479.

17. **Tan, K. S.** 2008. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. Clin. Microbiol. Rev. 21:639-665.

18. **Tan, T. C., S. C. Ong, and K. G. Suresh.** 2009. Genetic variability of *Blastocystis* sp. isolates obtained from cancer and HIV/AIDS patients. Parasitol. Res. 105:1283-1286.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

19. **Yan, Y., S. Su, J. Ye, X. Lai, R. Lai, H. Liao, G. Chen, R. Zhang, Z. Hou, and X. Luo.** 2007. *Blastocystis sp.* subtype 5: a possibly zoonotic genotype. Parasitol. Res. 101:1527-1532.

20. **Yoshikawa, H., I. Nagano, Z. Wu, E. H. Yap, M. Singh, and Y. Takahashi.** 1998. Genomic polymorphism among *Blastocystis hominis* strains and development of subtype-specific diagnostic primers. Mol. Cell. Probes 12:153-159.

21. **Yoshikawa, H., N. Abe, M. Iwasawa, S. Kitano, I. Nagano, Z. Wu, and Y. Takahashi.** 2000. Genomic analysis *of Blastocystis hominis* strains isolated from two long-term health care facilities. J. Clin. Microbiol. 38:1324-1330.

22. **Yoshikawa, H., Z. Wu, K. Pandey, B. D. Pandey, J. B. Sherchand, T. Yanagi, and H. Kanbara.** 2009. Molecular characterization of *Blastocystis* isolates from children and rhesus monkeys in Kathmandu, Nepal. Vet. Parasitol. 160:295-300.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

FIGURE LEGENDS

FIG. 1: Unrooted maximum likelihood tree of *Blastocystis* sp. isolates based on the 600 bp-domain sequences of the SSU rRNA gene. Numbers near the individual nodes indicate bootstrap values given as percentages. Values below 50% are not indicated. Clones sequenced in this study are called Cl1 to Cl50. Reference sequences from GenBank have the accession number preceded by the ST. The scale bar represents the average number of nucleotide replacements per site.

FIG. 2: Molecular heterogeneity of the SSU rDNA coding region of *Blastocystis* sp. ST7. Differences between 22 SSU rRNA gene copies in the 600 bp-domain common to that used for genotyping in this study are shaded. All the variable positions are located in a short region (positions 111 to 162 of the corresponding 600 bp-domain). As shown in the left column, SSU rRNA gene copies are located within numerous of the 54 scaffolds determined in the final assembled genome of *Blastocystis* sp. ST7 (3). Dots indicate sequence identity to 18S_scaffold0_1_1677, while asterisks indicate deletions.
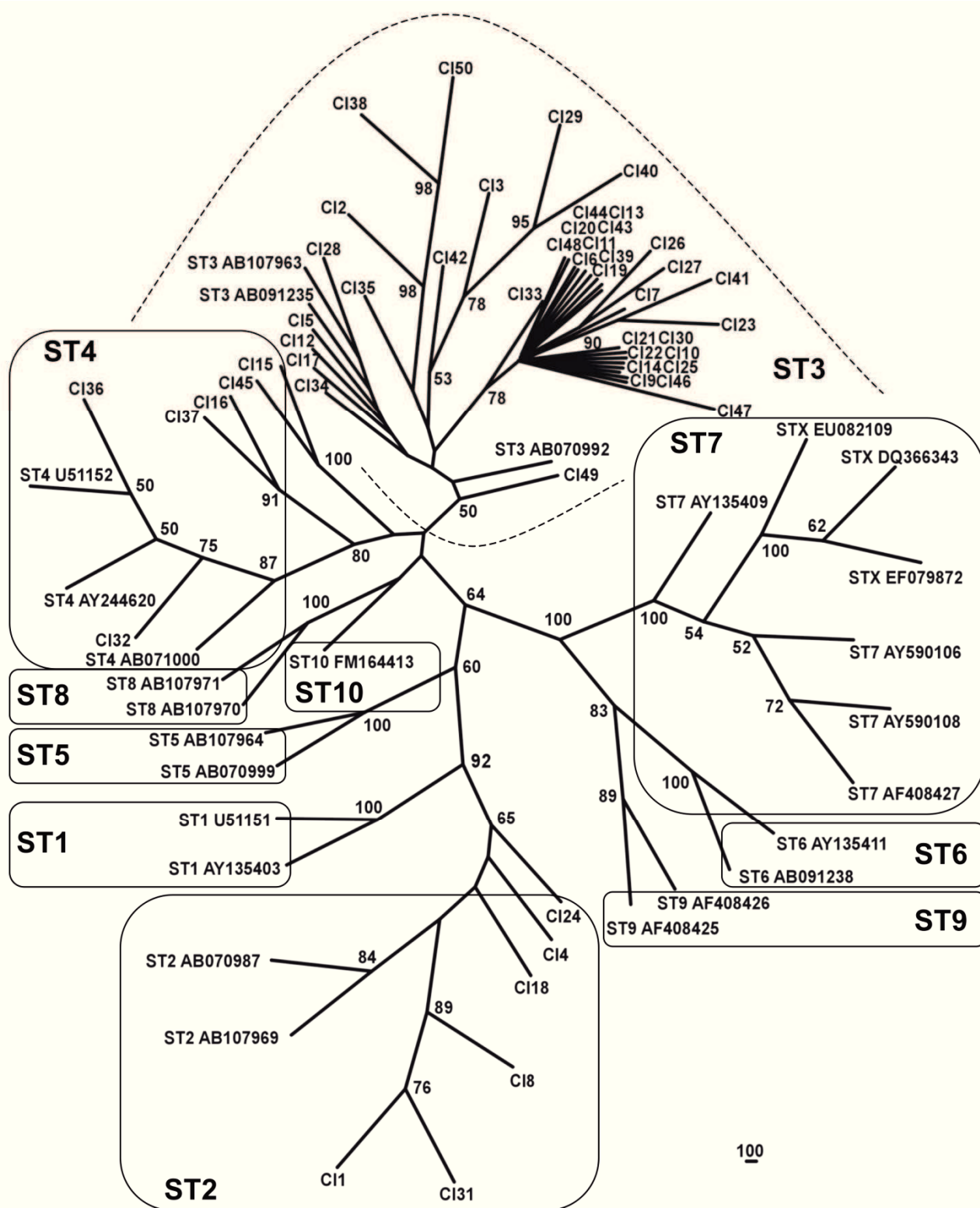
Figure 1

Nome e cognome: Dionigia Meloni

```
                          100                                                                         170
18S_scaffold0_1_1677      TGTTTTCTTTGATGGGGAACCATTAAAGAAGTACTTATTAGGCATAAAACCAA***AATTGTATTGTGAGT
18S_scaffold2_2_1804      ...............C.........................A...........TTT...GT.G.......
18S_scaf6_1800            ...............C.........................A...........TTT...GT.G.......
18S_scaffold6_2_1775      ...............C.........................A...........TTT...GT.G.......
18S_scaffold6_3_1800      ..........G....C.........................A...........TTT...GT.G.......
18S_scaffold9_2_1798      ...............C.........................***.....A...
18S_scaf11_1804           ...............C.........................A...........TTT...GT.G.......
18S_scaffold11_2_1804     ..........G....C.........................A...........TTT...GT.G.......
18S_scaf15_1772           ..........G....C.........................A...........TTT...GT.G.......
18S_scaf17_1808           ..........G....C.........................A...........TTT...GT.G.......
18S_scaf18_1_1808         ...............C.........................A...........TTT...GT.G.......
18S_scaf18_2_1801         ...............C.........................***.....A...
18S_scaf18_3_1808         ...............C.........................A...........TTT...GT.G.......
18S_scaf18_4_1809         ...............C.........................A...........TTT...GT.G.......
18S_scaf21_1797           .........................................***.....A...
18S_scaf24_1808           .........................................A...........TTT...GT.G.......
18S_scaf27_1801           ...............C.........................***.....A...
18S_scaf30_1800           ...............C.........................***.....A...
18S_scaf36_1780           ...............C.........................A...........TTT...GT.G.......
18S_scaf38_1819           ..........G....C.........................A...........TTT...GT.G.......
18S_scaf48_1806           ...............C.........................A...........TTT...GT.G.......
18S_scaf49_523            ...............C.........................A...........TTT...GT.G.......
```

Figure 2

General conclusions and perspectives

I-*Pentatrichomonas hominis*: pathogenicity and transmission

The trichomonad species *Pentatrichomonas hominis* colonizes the gastrointestinal tract of humans and many other mammals. There are very few studies on the prevalence of *Pentatrichomonas hominis* in the human population and it would probably not exceed 1.5% (Meloni et al. 1993; Crucitti et al. 2010). In parallel, its pathogenicity remains still much debated. It has often been suggested that *Pentatrichomonas hominis* was a commensal of the digestive tract and in some circumstances it could cause digestive problems (Honigberg 1978, 1990). In our study, we identified this trichomonad species in two patients.

In the first patient, a 33-year-old presenting different pathologies including IBS, trichomonads were not clearly visible in the stool. However these microorganisms were found in large numbers in the culture of the same stool after 10 days. In the second patient, an 11-month-old child, numerous trichomonads were found in the stool sample. In this case, the parasites were difficult to detect because all cytoskeletal structures typical of trichomonads were not visible, suggesting that the observed cells represented pseudocysts (Brugerolle 1976). Concerning the first patient, IBS was likely to be the main source of symptoms since we observed the clearance of trichomonads after IBS treatment without antiparasitic treatment with metronidazole. This implies that for still undetermined reasons, IBS has created favourable conditions for the development of *Pentatrichomonas hominis*. By contrast, in the second patient, *Pentatrichomonas hominis* is the probable causative agent of gastrointestinal troubles since the child did not present other intestinal parasites or pathogenic bacteria and metronidazole treatment led to clearance of trichomonads and clinical recovery. Our observations combined with those of previous studies (Yang et al. 1990; Chunge et al. 1991; Meloni et al. 1993; Ogunsanya et al. 1994) suggested that *Pentatrichomonas hominis* is a pathogen that can cause gastrointestinal disorders particularly in children aged 1 to 5 years.

In the first patient, numerous trichomonads were found after 10 days in stool culture suggesting a relatively prolonged survival of this parasite in water. Since the transmission of *Pentatrichomonas hominis* results in the consumption of contaminated food or water, its prolonged survival in the environment revealed the strong potential of contamination of this species through the fecal-oral route. In addition, we noted that the trichomonad sequences obtained from the two patients exhibited 98.9% to 99.6% identity to homologous sequences from *Pentatrichomonas hominis* strains obtained from different hosts including dogs and bovid, suggesting that the same species could colonize many mammalian species. This latter observation reinforced the hypothesis of the zoonotic potential of trichomonads. Indeed, the bovid species *Tritrichomonas foetus* and avian species of the genus *Tetratrichomonas* are

also able to colonize humans (Duboucher et al. 2007; Mantini et al. 2009). This also raised the question of the "human" origin of *Pentatrichomonas hominis* since this trichomonad species is frequently found in dogs and cats (Gookin et al. 2007; Kim et al. 2010).

Finally, according to the zoonotic character of *Pentatrichomonas hominis*, its strong potential of contamination by water as well as the difficulty to identify the pseudocysts in stool and environmental samples, the infection caused by this parasite is probably largely underestimated in humans and animals. Consequently, the prevalence of *Pentatrichomonas hominis* infections in large humans and animals populations has to be reassessed using molecular tools.

II-<u>Trichomonads colonize different unusual sites</u>

These flagellated protozoa are frequently found in the urogenital tracts (*Trichomonas vaginalis*), oral cavities (*Trichomonas tenax*), and intestinal tracts (*Pentatrichomonas hominis* and *Dientamoeba fragilis*) of humans (Honigberg 1990). It was thought that each human species had a specific tropism for its site of infection. However, it has been recently shown that these microorganisms could be frequently found outside their natural habitats, such as in the upper or lower respiratory tract of patients (pulmonary trichomonosis) with underlying pulmonary pathologies. In these clinical contexts, these trichomonads were generally identified as *Trichomonas tenax* on the basis of a limited number of questionable characters such as the shape and size of cells (Hersh 1985; Stratakis et al. 1999). This organism is usually regarded as a harmless commensal of the human oral cavity and is believed to enter the respiratory tract by aspiration from the contaminated oropharynx. As for the patient XR included in our study, the identification of *Trichomonas tenax* in the pleural fluid of several other patients has been confirmed using molecular tools (Mahmoud and Rahman 2004; Mallat et al. 2004; Bellanger et al. 2008). Therefore *Trichomonas tenax* is very likely the most frequently trichomonad species in this unusual location (Mahmoud and Rahman 2004). Interestingly, other trichomonad species have also been identified in human lungs. These include *Trichomonas vaginalis* (Duboucher et al. 2003), *Pentatrichomonas hominis* (Jongwutiwes et al. 2000), *Tritrichomonas foetus* (Duboucher et al. 2006), a genital trichomonad found in bovids, and avian species belonging to the genus *Tetratrichomonas* (Kutisova et al. 2005; Mantini et al. 2009). The identification of strains belonging to the genera *Tritrichomonas* and *Tetratrichomonas* in humans clearly raise again the question of the as yet poorly understood zoonotic potential of trichomonads, since these taxa were thought to be of animal origin. Alternatively, these isolates could represent human-host-

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

adapted species, since Kutisova et al. (2005) failed to transmit *Tetratrichomonas* strains of human origin to birds.

Regarding the recent literature, the presence of trichomonads in the human respiratory tract cannot be hereafter considered unusual. Indeed, these microorganisms are found frequently in the course of human PCP (Duboucher et al. 2005, 2007) and of ARDS (Duboucher et al. 2007). More precisely, trichomonads were detected as coinfecting agents with *Pneumocystis* in 60% of PCP patients and were found in 30% of ARDS patients. Because trichomonads are microaerophilic protozoa, they are likely not able to cause pulmonary disease by themselves and require favourable conditions for their development. With regard to PCP and ARDS, it was hypothesized that the development of these protozoa in the alveolar lumens was linked to local hypoxic conditions rather than immunodepression (Duboucher et al. 2007, 2008). Indeed, in these two clinical contexts, the alveolar lumens are obliterated by fungi in the case of PCP or by fibrin and cellular debris in the case of ARDS. In parallel, the proliferation of trichomonads in the pleural cavity has been reported in more than 15 cases of empyema (Lewis et al. 2003; Wang et al. 2006; Gilroy et al. 2007; Mantini et al. 2009), in addition to that reported in our study for patient XR, and appeared to depend on the presence of bacterial species in addition to anaerobic conditions. This is due to the fact that these parasites are known to feed on bacteria. Besides, antibiotics and metronidazole against coinfection with empyema, as well as drugs active against PCP, have consistently cured patients of pulmonary trichomonosis (Duboucher et al. 2003, 2006; Bellanger et al. 2008). To date, the frequency of pulmonary trichomonosis is probably broadly underestimated because of the difficulty to identify amoeboid trichomonads in lungs (Duboucher et al. 2007, 2008), the lack of research of these parasites in almost all high-risk patients with lung diseases (Jongwutiwes et al. 2000), and the too-long and inappropriate storage conditions of the samples (Bellanger et al. 2008). Moreover, six trichomonad species have already been identified as causative agents of pulmonary trichomonosis, and the number of trichomonad taxa identified as potentially involved as coinfecting parasites of lung diseases will likely increase in the future because of the use of molecular diagnostic tools.

In our study, in addition to the pleural fluid of the patient XR, *Trichomonas tenax* was not surprisingly identified in the oral wash specimen of two patients (VJ and MOC) but also in the stool of the patient VJ and colic biopsy of the patient JF. Interestingly, all the sequences obtained in the present work from different patients and samples and identified as *Trichomonas tenax* exhibited 100% identity to each other and with other sequences available in databases from this trichomonad species. Since the same strain was identified in the oral

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

cavity, digestive tube, and stool of patients, it clearly raises the question of the transit of *Trichomonas tenax* through the gastrointestinal tract. Indeed, to our knowledge, this is the first report of the presence of *Trichomonas tenax* in the feces as well as in the digestive tract in humans. During its transit, the parasite has to survive in conditions of highly acidic pH (the pH of the gastric acid is comprised between 1.5 and 3.5). In case of *Trichomonas vaginalis*, some strains are capable of growth and multiplication at pH 4.5 (Lehker and Alderete 1990) but such data are not yet available for *Trichomonas tenax*. In further studies, the growth and morphological changes in some axenic strains of *Trichomonas tenax* in pH conditions identical to those of the digestive tract of humans will be analyzed. In the immunocompetent patient JF presenting with diarrhea, we detected a mixed infection with *Trichomonas tenax* and *Trichomonas vaginalis*. Strikingly this patient did not present other intestinal parasites or pathogenic bacteria suggesting that *Trichomonas tenax* or *Trichomonas vaginalis* could be responsible for gastrointestinal disorders. Finally, we also identified *Trichomonas vaginalis* in the urine of the patient FC with nervous breakdown, cephalea, and feeling of faintness. The presence of this parasite in urines is not completely unusual since this location has been already described in some case reports as in a woman (Loo et al. 2000), in a male with an eight-year history of sterile pyuria (Niewiadomski et al. 1998), and in a 3 month-old girl (Bellanger et al. 2008).

In conclusion, we confirmed that several trichomonad species are able to colonize different unusual sites under favourable condition. As shown in our study, this colonization is not only observed in immunodepressed patients. Therefore, the significance of the location of these parasites has to be clarified and particularly in lungs since pulmonary infections represent a public health priority. Several observations reviewed by Duboucher et al. (2007, 2008), suggest the potentially pathogenic effect of trichomonads on the alveolar epithelia of patients. However, in order to confirm this hypothesis, physiopathological studies of these parasites have to be performed by developing *in situ* hybridization methods for lung tissues of patients (Liebhart et al. 2006) and/or animal models of pulmonary coinfection by trichomonads and other microorganisms. Regardless of their location, trichomonads could play an active role in the extension of lesions and more generally aggravate a patient's poor physical condition and delay recovery.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

III-<u>Phylogenies of parabasalids based on the concatenation of sequences of different markers</u>

For various reasons, the recent phylogenies based on SSU rRNA gene sequences (Gerbod et al. 2002; Hampl et al. 2004, 2006, 2007; Ohkuma et al. 2005; Carpenter and Keeling 2007; Noda et al. 2009) were in conflict with those inferred from morphological characters (Honigberg 1963; Brugerolle 1976). Consequently, Brugerolle and Lee (2001), Brugerolle and Patterson (2001), Adl et al. (2005), and finally Cepicka et al. (2010) proposed new classifications of parabasalids. In the latter classification, it was proposed to divide the parabasalids into six classes (Cepicka et al. 2010). However, a great uncertainty remains with respect to the phylogenetic relationships among and within the classes and the root position of Parabasalia. To overcome the lack of resolution of some parabasalid lineages based on SSU rRNA gene sequences, phylogenies of parabasalids have been inferred from the sequences of additional molecular markers. This allowed increasing the number of informative characters through the concatenation of sequences, thereby largely overcoming the problems encountered in previous molecular phylogenetic studies. Thus GAPDH, EF1a, actin, and SSU rRNA sequences are now available from a large sampling of parabasalids covering the full diversity of this group. Our phylogenetic inference based on concatenated data set mostly supported the new classification proposed by Cepicka et al. (2010). Indeed, Hypotrichomonadea, Trichomonadea, Trichonymphea, Spirotrichonymphea, and Cristamonadea represented well-supported lineages. In addition, as repeatedly stated (Hampl et al. 2004; Noël et al. 2007; Noda et al. 2009; Cepicka et al. 2010), Tritrichomonadea formed a not well-supported distinct lineage and a sister-group of Cristamonadea. This suggested a possible revision of the taxonomic status of Tritrichomonadea (and also Cristamonadea). For instance, these two classes could be united in the same group Cristamonadea.

In parallel, the root of Parabasalia was investigated through analyses with outgroup taxa. In the present study, GAPDH sequences were excluded from the analyses based on the concatenated dataset since parabasalid genes were likely acquired from a bacterium (Viscogliosi and Müller 1998). The concatenated dataset outlined a more plausible new root position, located at the position dividing some hypermastigids, the Trichonymphea and the Trichomonadea group from the others suggesting that the most primitive parabasalid exhibited a trichomonad-like character. Trichomonads belonging to Tritrichomonadea, Hypotrichomonadea, and Trichomonadea present a costa and undulating membrane with some variations among these three classes. The primitive parabasalid likely possessed the

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

lamelliform undulating membrane, and the rail-like undulating membrane evolved later. The homologous protein components of both types of costa suggested their common origin (Viscogliosi et Brugerolle 1994) and the differentiation of the two types of costa probably occurred very early in parabasalid evolution.

Historically, monocercomonads for its simple cytoskeleton was considered as the most ancient lineage of parabasalids and a more complex structure and morphology have developed later during parabasalian evolution (Honigberg 1963; Brugerolle 1976). Earlier molecular phylogenetic studies have challenged this scheme suggesting Trichonymphea, the most elaborate, in the basal placement. In this study, the new root position indicated that the evolution of parabasalids is principally simple-to-complex, but the complexity has emerged independently in multiple lineages in different modes of flagellar system multiplication. For instance, the flagella of hypermastigids have independently evolved three times.

IV-Molecular epidemiology of *Blastocystis* sp. in Italy and Lebanon

The aim of our epidemiological studies was to determine the *Blastocystis* sp. STs present in different human populations and their relative frequency. These studies are of special interest for understanding the biology and assessing the public health significance of this parasite. Indeed, they permit characterization of the transmission dynamics and pathogenicity of the different *Blastocystis* sp. STs.

In the present work we first subtyped 34 *Blastocystis* sp. isolates from 30 symptomatic patients (4 mixed infections containing two different STs) presenting with various gastrointestinal troubles (diarrhea, abdominal pain, constipation and/or vomiting) and living in or the vicinity of Rome and Sassari (Sardinia). In this Italian population, six STs were identified, and ST3 was the most common (47.1%) followed by ST2 (20.6%), ST4 (17.7%), ST1 (8.8%), and ST7 and ST8 (2.9%). Globally, the ST distribution showed similarities with that found in other countries (for review see Souppart et al. 2009). Indeed, all these studies reported that a large majority of human *Blastocystis* sp. infections (an average of more than 60%) were attributable to ST3 isolates whereas the proportions of other STs differed between locations. Only few exceptions showed the higher prevalence of other STs in Spain (Dominguez-Marquez et al. 2009) and Thailand (Thathaisong et al. 2003). In a previous preliminary study, Mattiuci et al. (2010) genotyped six *Blastocystis* sp. isolates obtained in the Rome area from symptomatic patients and classified as belonging to ST1, ST5, and ST7. Consequently, together with our data, a total of seven STs were identified in Italian

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

symptomatic patients reinforcing the hypothesis of the existence of pathogenic and non pathogenic variants in different STs.

The dominant ST3 was suggested to be the only ST of human origin, while the remaining STs are zoonotic (Böhm-Gloning et al. 1997; Kaneda et al. 2001; Yoshikawa et al. 2004; Noël et al. 2005; Tan 2008). This hypothesis was recently reinforced by the determination of the predominance of ST3 in an urbanized area such as Singapore where there is a limited opportunity for zoonotic transmission (Wong et al. 2008). The predominance of this ST might be explained by large-scale human-to-human transmission in Italy as in other countries (Yoshikawa et al. 2000). Concerning other STs, their presence in the Italian population might be linked to zoonotic transmission from farm animals, pigs/dogs/cattle, birds, primates, and/or domestic rodents depending of the STs (Noël et al. 2005; Yan et al. 2007; Tan 2008; Stensvold et al. 2009). Since most of the STs found in the Italian population are zoonotic, our data raise crucial questions concerning the identification of animal reservoirs for this parasite. In further studies, our aim will be to identify specific risk factor for human infections with different *Blastocystis* sp. STs in Italy.

Despite the obvious interest of this parasite in public health, any data on the prevalence and genetic diversity of *Blastocystis* sp. was available for Lebanon. Therefore, in the present work, stool samples were collected from 220 symptomatic and asymptomatic patients in or in the vicinity of Tripoli, Northern Lebanon. *Blastocystis* sp. was identified in a total of 42 samples (19%) through direct light microscopy of smears. This prevalence is significant and similar to that observed in several developing nations as well as in USA (Tan 2008). However, prevalence data are largely dependent on the methods used for detection and direct light microscopy of fecal smears greatly underestimate the prevalence of *Blastocystis* sp. isolates compared to culture and nonquantitative and quantitative PCR (Suresh and Smith 2004; Stensvold et al. 2007; Poirier et al. 2011; Santin et al. 2011). Although PCR is considered the most sensitive technique to detect the parasite (Stensvold et al. 2007), only 19 of 42 positives samples were amplified following the PCR protocol used in previous epidemiological studies (Souppart et al. 2009, 2010). This might be due to the presence of known PCR inhibitors in fecal samples or to a low quantity of *Blastocystis* sp. DNA in these samples. The remaining 25 DNAs have been subsequently analyzed by quantitative PCR and were shown to be positive (Poirier et al. 2011). Expected sequences should allow us to complete this study in the coming months.

Concerning the 19 samples already subtyped, they all represented single infections. Therefore we analyzed a total of 19 isolates which belonged to 4 different STs: ST3 was the

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

most common ST in the present Lebanese population (53%) followed by ST1 (21%), ST2 (21%), and ST4 (6%). Among the 19 patients, 10 were asymptomatic and 9 symptomatic. At present time, the sample size remains too low and statistical analysis has not yet been performed to correlate symptomatic status of patients and *Blastocystis* sp. ST. This possible correlation has not been clearly demonstrated yet and published studies are contradictory (Kaneda et al. 2001; Yan et al. 2006, Dogruman-Al et al. 2008, 2009, Hussein et al. 2008; Özyurt et al. 2008, Tan et al. 2008; Eroglu et al. 2009; Souppart et al. 2009; Stensvold et al. 2009, 2011). However, in our study, we noted that ST1 isolates were only found in symptomatic patients suggesting a correlation between ST1 and digestive symptoms (Yan et al. 2006). Interestingly, in a recent *in vivo* study, Hussein et al. (2008) have shown that human ST1 isolates obtained from symptomatic patients induced 25% mortality in rats. In addition, the ST2 is mainly symptomatic in Lebanon as in France (Souppart et al. 2009) while ST3 is largely asymptomatic.

V-<u>Mixed infections with *Blastocystis* sp. isolates</u>

In epidemiological studies, molecular subtyping of *Blastocystis* sp. isolates is performed using different methods: PCR employing ST-specific primers (Yoshikawa et al. 1998) and amplification of informative domains of the SSU rRNA gene (Scicluna et al. 2006) followed by RFLP (Abe et al. 2003) or cloning coupled with sequencing of a limited number of clones (Souppart et al. 2009, 2010). However, these methods show some limitations that could seriously distort the results of these studies. For instance, PCR employing ST-specific primers allows only the amplification of 7/13 STs. Moreover, sequencing of a limited number of clones following cloning of PCR products provides only incomplete survey of the isolates potentially present in the sample. Indeed, regarding this latter method, only two clones are usually arbitrarily selected and sequenced. Using this method, the prevalence of mixed infections with two different STs was 7.5% in France (Souppart et al. 2009) and 5% in Egypt (Souppart et al. 2010). Although these values were roughly similar to those described in other countries such as Germany (5%; Böhm-Gloning et al. 1997) and Turkey (4.3%; Dogruman-Al et al. 2008) using other methods, they were likely underestimated. In addition, for numerous samples, the sequences of two clones showed some nucleotide differences to each other. The differences observed between the two clones (between 1 to 11 differences according to Souppart et al. 2009, 2010) of a same sample are unlikely to be due to PCR artefacts, as the study used high fidelity *Taq* DNA polymerase. Therefore, as suggested by Scicluna et al. (2006) and Souppart et al. (2009), the cause may be coinfection with two

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

variants within the same ST or sequence variations between SSU rDNA gene copies within the same isolate. In the genome of a *Blastocystis* sp. ST7 isolate (Denoeud et al. 2011), 38 SSU rDNA gene copies have been identified that differed by up to nine nucleotides to each other in the gene domain commonly used in epidemiological studies. This suggested that beyond this number of differences, a mixed infection with two isolates of the same ST could be considered in the same individual. The issue of mixed intra-ST infection has never been studied. Therefore, we have analyzed the genetic diversity of *Blastocystis* sp. isolates in a French patient considered at high potential risk of mixed infection with this parasite according to his lifestyle, by the sequencing of 50 clones. Interestingly, 76% of sequences belong to ST3, 12% to ST2, and 12% to ST4. According to our alignment of sequences, ST3 clones showed a large intra-ST variation. This intra-ST variation was lower within ST2 and ST4 clones. All these data revealed that this patient was infected by several isolates belonging to three different STs suggesting multiple potential contamination sources. As suggested previously (Yoshikawa et al. 2009), intra-ST polymorphism might reflect the existence of several subgroups within the same ST.

In further studies, patients at risk of mixed infection should be identified in order to optimize the genotyping approach and to prevent misunderstanding in the context of epidemiological surveys. In parallel, comparative genomics approaches should lead to the development of Multi Locus Sequence Typing (MLST) methods that could optimally assess the diversity of *Blastocystis* sp. isolates from individual patients.

References

A

Abe, N., et al. (2002) "A survey of *Blastocystis sp*. in livestock, pets, and zoo animals in Japan." Vet Parasitol 106(3):203-12.

Abe, N., et al. (2003a) "Molecular characterization of *Blastocystis* isolates from birds by PCR with diagnostic primers and restriction fragment length polymorphism analysis of the small subunit ribosomal RNA gene." Parasitol Res 89(5):393-6.

Abe, N., et al. (2003b) "Zoonotic genotypes of *Blastocystis hominis* detected in cattle and pigs by PCR with diagnostic primers and restriction fragment length polymorphism analysis of the small subunit ribosomal RNA gene." Parasitol Res 90(2):124-8.

Abe, N. (2004) "Molecular and phylogenetic analysis of *Blastocystis* isolates from various hosts." Vet Parasitol 120(3):235-42.

Adl, S. M., et al. (2005) "The new higher level classification of eukaryotes with emphasis on the taxonomy of protists." J Eukaryot Microbiol 52(5): 399-451.

Aguiar, J. I., et al. (2007) "Intestinal protozoa and helminths among Terena Indians in the State of Mato Grosso do Sul: high prevalence of *Blastocystis hominis*." Rev Soc Bras Med Trop 40(6):631-4.

Amos, W. B., et al. (1979) "Structure, protein composition and birefringence of the costa: a motile flagellar root fibre in the flagellate *Trichomonas*." J. Cell. Sci. 35:139-64.

Archibald, J .M., et al. (2000) "Origin and evolution of eukaryotic chaperonins: phylogenetic evidence for ancient duplications in CCT genes." Mol. Biol. Evol., 17:1456-1466.

Archibald, J. M., et al. (2002) "The chaperonins genes of jakobid and jakobid-like flagellates: implications for eukaryotic evolution." Mol. Biol. Evol., 19:422-431.

Arisue, N., et al. (2002) "Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data." J Eukaryot Microbiol 49(1):42-53.

Arisue, N., et al. (2003) "Sequence heterogeneity of the small subunit ribosomal RNA genes among *blastocystis* isolates." Parasitology 126(Pt 1):1-9.

Armentia, A., et al. (1993) "Urticaria by *Blastocystis hominis*. Successful treatment with paromomycin." Allergol Immunopathol (Madr) 21(4):149-51.

B

Bapteste, E. and Philippe H. (2002) "The potential value of indels as phylogenetic markers: position of trichomonads as a case study." Mol Biol Evol 19(6): 972-7.

Bellanger, A. P., et al. (2008) "Two unusual occurrences of trichomoniasis: rapid species identification by PCR." J Clin Microbiol 46(9): 3159-61**.**

Benchimol, M. (2009) "Hydrogenosomes under microscopy." Tissue and Cell 41(3): 151-168.

Benchimol, M. and W. De Souza (1983) "Fine structure and cytochemistry of the hydrogenosome of *Tritrichomonas foetus*." J Protozool 30(2): 422-5.

Benchimol, M., et al. (1996) "Morphogenesis of the hydrogenosome: an ultrastructural study." Biol Cell 87(3): 197-205.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Berchtold, M., et al. (1995) "Culture and phylogenetic characterization of *Trichomitus trypanoides* Duboscq & Grasse 1924, n. comb.: a trichomonad flagellate isolated from the hindgut of the termite *Reticulitermes santonensis* Feytaud." J Eukaryot Microbiol 42(4): 388-91.

Berchtold, M. & König H. (1995) "Phylogenetic position of the two uncultivated trichomonads *Pentatrichomonoides scroa* Kirby and *Metadevescovina extranea* Kirby from the hindgut of the termite *Mastotermes darwiniensis* Froggatt." System. Appl. Microbiol., 18:567-573.

Bishop, A. (1939) "A note upon the systematic position of *Trichomonas Keilini* (Bishop, 1935)." Parasitology, 31:469-472.

Böhm-Gloning, B., et al., (1997) "Five subgroups of *Blastocystis hominis* from symptomatic and asymptomatic patients revealed by restriction site analysis of PCR-amplified 16S-like rDNA." Trop Med Int Health. 2(8):771-8.

Boorom, K. F., et al. (2008) "Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection." BMC Parasit Vectors 1(1):40.

Breznak, J. A. (1982) "Intestinal microbiota of termites and other xylophagous insects." Annu Rev Microbiol 36: 323-43.

Brugerolle, G. (1975) "Etude de la cryptopleuromitose et de la morphogenèse de division chez *Trichomonas vaginalis* et chez plusieurs genres de trichomonadines primitives." Protistologica, 11:457-468.

Brugerolle, G. (1976) "Cytologie ultrastructurale, systématique et évolution des Trichomonadida." Ann. Station Biologique de Besse-en- Chandesse, 10: 1-90

Brugerolle, G., (1986) "Séparation des genres *Trimitus* (Diplomonadida) et *Tricercomitus* (Trichomonadida) d'après leur ultrastructure." Protistologica, 22: 31-37.

Brugerolle, G., et al. (1994) "Ultrastructural study of *Pentatrichomonoides* sp., a trichomonad flagellate from *Mastotermes darwiniensis*." Europ. J. Protistol., 30:372-378.

Brugerolle, G. and Lee J. J. (2001) "Phylum Parabasalia." pp 1996 - 1250. *In*: The illustrated guide to the Protozoa II. Eds: lee J.J., Leedale G.F., Bradbury P.C.' Society of Protozoologists, Lawrence, Kansas.

Brugerolle, G. and Patterson D.J. (2001) "Ultrastructure of *Joenina pulchella* Grassi, 1917 (Protista, Parabasalia), a reassessment of evolutionary trends in the parabasalids, and a new order Cristamonadida for devescovinids, calonymphids and lophomonad flagellates." Org Divers Evol 1:147–160

C

Camp, R. R., et al. (1974) "Study of *Dientamoeba fragilis* Jepps & Dobell. I. Electronmicroscopic observations of the binucleate stages. II. Taxonomic position and revision of the genus." J Protozool 21(1): 69-82.

Carlton, J. M., et al. (2007) "Draft Genome Sequence of the Sexually Transmitted Pathogen *Trichomonas vaginalis*." Science 315(5809): 207-212.

Carpenter, K. J., et al. (2009) "Morphology, phylogeny, and diversity of *Trichonympha* (Parabasalia: Hypermastigida) of the wood-feeding cockroach *Cryptocercus punctulatus*." J Eukaryot Microbiol 56(4): 305-13.

Carpenter, K. J. and Keeling P. J. (2007) "Morphology and phylogenetic position of Eucomonympha imla (Parabasalia: Hypermastigida)." J Eukaryot Microbiol. 54(4):325-32.

Cassano, N., et al. (2005) "Remission of delayed pressure urticaria after eradication of *Blastocystis hominis*."Acta Derm Venereol 85(4):357-8.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Cavalier-Smith, T. (2002) "The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa." Int J Syst Evol Microbiol 52(Pt 2): 297-354.

Cepicka, I., et al. (2006) "New evolutionary lineages, unexpected diversity, and host specificity in the parabasalid genus *Tetratrichomonas*." Mol Phylogenet Evol 39(2): 542-51.

Cepicka, I., et al. (2010) "Critical taxonomic revision of Parabasalids with description of one new genus and three new species." Protist 161(3): 400-33.

Chunge, R. N., et al. (1991) "Longitudinal study of young children in Kenya: intestinal parasitic infection with special reference to *Giardia lambia*, its prevalence, incidence and duration, and its association with diarrhoea and with other parasites." Acta Trop. 50:39-40.

Cimerman, S., et al. (2003) [*Blastocystosis*: nitazoxanide as a new therapeutic option]. Rev Soc Bras Med Trop. 36(3):415-7.

Cirioni, O., et al. (1999) "Prevalence and clinical relevance of *Blastocystis hominis* in diverse patient cohorts." Eur J Epidemiol 15(4):389-93.

Clemens, D. L. and Johnson P. J. (2000) "Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy." Mol Biochem Parasitol 106(2): 307-13.

Cleveland, L. R. (1961) "The centrioles of *Trichomonas* and their functions in cell reproduction." Arch. Protistenkd., 105:149-162.

Cotch, M. F., et al. (1997) "*Trichomonas vaginalis* associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group." Sex Transm Dis 24(6): 353-60.

Corbeil, L. B. (1994) "Vaccination strategies against *Tritrichomonas foetus*." Parasitol Today 10(3): 103-6.

Crucitti, T., et al. (2010) "*Trichomonas vaginalis* is highly prevalent in adolescent girls, pregnant women, and commercial sex workers in Ndola, Zambia." Sex. Transm. Dis. 37:223-227.

Cudmore, S. L., et al. (2004) "Treatment of infections caused by metronidazole-resistant *Trichomonas vaginalis*." Clin Microbiol Rev 17(4): 783-93, table of contents.

D

Dacks, J. B. and Redfield R. J. (1998) "Phylogenetic placement of *Trichonympha*." J Eukaryot Microbiol 45(4): 445-7.

Delgado-Viscogliosi, P., et al. (1996) "Tubulin post-translational modifications in the primitive protist *Trichomonas vaginalis*." Cell Motil Cytoskeleton 33(4): 288-97.

Delgado-Viscogliosi, P., et al. (2000) "Molecular phylogeny of parabasalids based on small subunit rRNA sequences, with emphasis on the Trichomonadinae subfamily." J Eukaryot Microbiol 47(1): 70-5.

Denoeud, F., et al. (2011) "Genome sequence of the stramenopile *Blastocystis sp*., a human anaerobic parasite: from adaptive evolution to pathogenic potential." Genome Biol. 12(3):R29

Dessi', D., et al. (2005) "Long-Term Survival and Intracellular Replication of *Mycoplasma hominis* in *Trichomonas vaginalis* Cells: Potential Role of the Protozoon in Transmitting Bacterial Infection" Infect ImmunFebruary; 73(2): 1180–1186.

Dessì, D., et al. (2006) "*Mycoplasma hominis* and *Trichomonas vaginalis*: a unique case of symbiotic relationship between two obligate human parasites." Front Biosci.1;11:2028-34.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Diaz, N., et al. (2010) "Rapid detection of coinfections by *Trichomonas vaginalis*, *Mycoplasma hominis*, and *Ureaplasma urealyticum* by a new multiplex polymerase chain reaction." Diagn Microbiol Infect Dis. 67(1):30-6.

Diaz, E., et al. (2003) "Epidemiology and control of intestinal parasites with nitazoxanide in children in Mexico." Am J Trop Med Hyg 68(4):384-5.

Dogruman-Al, F., et al. (2008) "A possible link between subtype 2 and asymptomatic infections of *Blastocystis hominis*." Parasitol Res 103(3):685-9.

Dogruman-Al, F., et al. (2009) "*Blastocystis* subtypes in irritable bowel syndrome and inflammatory bowel disease in Ankara, Turkey." Mem Inst Oswaldo Cruz 104(5):724-7.

Dogruman-Al, F., et al. (2010) "Comparison of methods for detection of *Blastocystis* infection in routinely submitted stool samples, and also in IBS/IBD Patients in Ankara, Turkey." PLoS One 5(11):e15484.

Dominguez-Marquez, M. V., et al. (2009) "High prevalence of subtype 4 among isolates of *Blastocystis hominis* from symptomatic patients of a health district of Valencia (Spain)." Parasitol Res 105(4):949-55.

Duboucher, C., et al. (2003) "Pulmonary coinfection by *Trichomonas vaginalis* and *Pneumocystis sp.* as a novel manifestation of AIDS." Hum Pathol 34(5): 508-11.

Duboucher, C., et al. (2005) "Frequency of trichomonads as coinfecting agents in *Pneumocystis pneumonia*." Acta Cytol 49(3): 273-7.

Duboucher, C., et al. (2006) "Trichomonads as superinfecting agents in *Pneumocystis pneumonia* and acute respiratory distress syndrome." J Eukaryot Microbiol 53 Suppl 1: S95-7.

Duboucher, C., et al. (2007a) "Pulmonary superinfection by trichomonads in the course of acute respiratory distress syndrome." Lung 185(5): 295-301.

Duboucher, C., et al. (2007b) "*Pneumocystis pneumonia*: immunosuppression, *Pneumocystis jirovecii*...and the third man." Nat Rev Microbiol 5(12).

Duboucher, C., et al. (2008) "Recent advances in pulmonary trichomonosis." Trends Parasitol 24(5): 201-2.

Dufernez, F., et al. (2007) "Morphological and molecular identification of non-*Tritrichomonas foetus* trichomonad protozoa from the bovine preputial cavity." J. Eukaryot. Microbiol. 54(2):161-8.

E

Edgcomb V.P., et al. (1998) "New insights into the phylogeny of trichomonads inferred from small subunit rRNA sequences." Protist, 149:359-366.

Elwakil, H. S., and Hewedi, H. I (2010) "Pathogenic potential *of Blastocystis hominis* in laboratory mice." Parasitol Res 107(3):685-9.

El-Sherbini, G. T., et al. (2009) "Efficacy of two plant extracts against vaginal trichomoniasis." J Egypt Soc Parasitol 39(1): 47-58.

Eroglu, F., et al. (2009) "Identification of *Blastocystis hominis* isolates from asymptomatic and symptomatic patients by PCR." Parasitol Res 105(6):1589-92.

Eroglu, F., and Koltas, I. S. (2010) "Evaluation of the transmission mode of *B. hominis* by using PCR method." Parasitol Res 107(4):841-5.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

F

Farmer, M. A. (1993) "Ultrastructure of *Ditrichomonas honigbergii* n. g., n. sp. (Parabasalia) and its relationship to amitochondrial protists." J. Eukaryot. Microbiol. 40:619-626.

Fröhlich, J. and König, H. (1999) "Rapid isolation of single microbial cells from mixed natural and laboratory populations with the aid of a micromanipulator." Syst Appl Microbiol. 22(2):249-57.

Fukura, K., et al. (1996) "Nucleotide sequence of the SrRNA gene and phylogenetic analysis of *Trichomonas tenax*." Microbiol Immunol 40(3): 183-8.

G

Gardner, W. A. Jr., et al. (1986) "*Trichomonas vaginalis* in the prostate gland." Arch Pathol Lab Med 110(5): 430-2.

Gardner, W. A. Jr., et al. (1987) "Subepithelial organisms in trichomonal cervicitis." Diagn Cytopathol 3(3): 227-9.

Gerbase, A. C., et al. (1998) "Global prevalence and incidence estimates of selected curable STDs." Sex Transm Infect 74 Suppl 1: S12-6.

Gerbod, D., et al. (2000) "Phylogenetic position of parabasalid symbionts from the termite Calotermes flavicollis based on small subunit rRNA sequences." Int Microbiol. 3(3):165-72.

Gerbod, D., et al. (2001) "Phylogenetic position of the trichomonad parasite of turkeys, *Histomonas meleagridis* (Smith) Tyzzer, inferred from small subunit rRNA sequence." J Eukaryot Microbiol 48(4): 498-504.

Gerbod, D., et al. (2002) "Molecular phylogeny of parabasalids inferred from small subunit rRNA sequences, with emphasis on the Devescovinidae and Calonymphidae (Trichomonadea)." Mol Phylogenet Evol 25(3): 545-56.

Gerbod, D., et al. (2004) "Molecular phylogenies of Parabasalia inferred from four protein genes and comparison with rRNA trees." Mol Phylogenet Evol 31(2): 572-80.

Giacometti, A., et al. (1999) "Irritable bowel syndrome in patients with *Blastocystis hominis* infection." Eur J Clin Microbiol Infect Dis 18(6):436-9.

Gilroy, S. A., et al. (2007) *Trichomonas* species empyema coinfection in an alcoholic female. Clin. microbial. Newsl. 29:69-71.

Goldman, L. M., et al. (2009) "Treatment of metronidazole-resistant *Trichomonas vaginalis*." Sex Health 6(4): 345-7.

Gookin, J. L., et al. (2007) "Identification of *Pentatrichomonas hominis* in feline fecal samples by polymerase chain reaction assay." Vet Parasitol 145(1-2): 11-5.

Grain, J. (1986) "The cytoskeleton in protists: nature, structure, and functions." *Int Rev Cytol 104: 153-249.*

Grassé P-P. (1952) "Protozoaires: Généralités. Flagellés." *In* : Traité de Zoologie. Grassé P-P. (ed .), Masson & Cie, Paris, vol .1, fasc.1.

Grosovsky B. D. D. and Margulis M. (1982) "Termite microbial communities." *In*: Experimental microbial ecology. Burns R.G. & Slater J.H. (eds.), Blackwell Scientific, Oxford, pp 519-532.

Gunderson, J., et al. (1995) "Phylogeny of trichomonads inferred from small-subunit rRNA sequences." J Eukaryot Microbiol 42(4): 411-5.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Gupta, R. and Parsi, K. (2006) "Chronic urticaria due to *Blastocystis hominis*." Australas J Dermatol 47(2):117-9.

H

Hameed, D. M., et al. (2010) "Association of *Blastocystis hominis* genetic subtypes with urticaria." Parasitol Res. 108(3):553-60.

Hampl, V., et al. (2004) "Critical analysis of the topology and rooting of the parabasalian 16S rRNA tree." Molecular Phylogenetics and Evolution 32(3): 711-723.

Hampl, V., et al. (2005) "Inference of the phylogenetic position of oxymonads based on nine genes: support for metamonada and excavata." Mol Biol Evol 22(12): 2508-18.

Hampl, V., et al. (2006) "Affiliation of *Cochlosoma* to trichomonads confirmed by phylogenetic analysis of the small-subunit rRNA gene and a new family concept of the order Trichomonadida." Int J Syst Evol Microbiol 56(Pt 1): 305-12.

Hampl, V., et al. (2007) "Morphological and molecular diversity of the monocercomonadid genera *Monocercomonas*, *Hexamastix*, and *Honigbergiella gen. nov.*" Protist 158(3): 365-83.

Hampl, V., et al. (2009) "Phylogenenomic analyses support the monophyly of Excavata and resolve relashionships among eukaryotic "supergroups"." Proc. Natl. Acad. Sci USA 106:3859-3864

Haresh, K., et al. (1999) "Isolate resistance of *Blastocystis hominis* to metronidazole." Trop Med Int Health 4(4):274-7.

Hersh, S. M. (1985) "Pulmonary trichomoniasis and *Trichomonas tenax*." J Med Microbiol 20(1): 1-10.

Hirata, T., et al. (2007) "Prevalence of *Blastocystis hominis* and *Strongyloides stercoralis* infection in Okinawa, Japan." Parasitol Res 101(6):1717-9.

Hoevers, J., et al. (2000) "Restriction-fragment-length polymorphism analysis of small-subunit rRNA genes of *Blastocystis hominis* isolates from geographically diverse human hosts." Parasitol Res 86(1):57-61.

Hjort, K., et al. (2010) "Diversity and reductive evolution of mitochondria among microbial eukaryotes." Philos Trans R Soc Lond B Biol Sci 365(1541): 713-27.

Hollande A. and Valentin J. (1969) "Appareil de Golgi, pinocytose, lysosomes, mitochondries, bactéries symbiotiques, atractophores et pleuromitose chez les Hypermastigines du genre *Joenia*. Affinités entre Joeniidés et Trichomonadines." Protistologica, 5:39-86

Hollande A. and Carruette-Valentin J. (1971) "Les atractophores, l'induction du fuseau et la division cellulaire chez les Hypermastigines, étude infrastructurale et révision systématique des Trichonymphines et des Spirotrichonymphines." Protistologica, 7:5-100.

Hollande A. and Carruette-Valentin J. (1972) "Le problème du centrosome et la cryptopleuromitose atractophorienne chez *Lophomonas striata*." Protistologica, 8:267-278.

Hollande A. (1974) "Etude comparée de la mitose syndinienne et de celle des Péridiniens libres et des Hypermastigines, infrastucture et cycle évolutif des Syndinides parasites de Radiolaires. " Protistologica, 10:413-451

Hollande A. and Valentin J. (1968) "Morphologie infrastructurale de *Trichomonas* (*Trichomitopsis* Kofoid et Swezy, 1919) *termopsidis*, parasite intestinal de *Termopsis augusticollis*" Walk. Critique de la notion de centrosome chez les Polymastigines. Protistologica, 4:127-145.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Honigberg, B. M. (1963) "Evolutionary and systematic relationships in the flagellate order Trichomonadida Kirby." J. Protozool. 10:20-63.

Honigberg B.M. (1978) "Trichomonads of importance in human medicine." *In*: Parasitic Protozoa. Kreier J.P. (ed.), Academic Press, New York, 2:275-454.

Honigberg B. M., et al. (1968) "Structure of *Pentatrichomonas hominis* as revealed by electron microscopy. " J. Protozool., 15:419-430.

Honigberg B. M. and Brugerolle G. (1990) "Structure."*In*: Trichomonads Parasitic in Humans. Honigberg B.M. (ed.), Springer-Verlag, Berlin, pp 2-35.

Huber, K. et al. (2006) "*Histomonas meleagridis* in turkeys: dissemination kinetics in host tissues after cloacal infection." Poult Sci 85(6): 1008-14.

Hussein, E. M., et al. (2008) "Pathophysiological variability of different genotypes of human *Blastocystis hominis* Egyptian isolates in experimentally infected rats." Parasitol Res 102(5):853-60.

I

Idris, N. S., et al. (2010) "Intestinal parasitic infection of immunocompromised children with diarrhoea: clinical profile and therapeutic response." J Infect Dev Ctries 4(5):309-17.

J

Jakobsen, E. B., et al. (1987) "*Trichomonas* species in a subhepatic abscess." Eur J Clin Microbiol 6(3): 296-7.

Jones, M. S., et al. (2009) "Association of *Blastocystis* subtype 3 and 1 with patients from an Oregon community presenting with chronic gastrointestinal illness." Parasitol Res 104(2):341-5.

Jongwutiwes, S., et al. (2000) "*Pentatrichomonas hominis* in empyema thoracis." Trans R Soc Trop Med Hyg 94(2): 185-6.

Joyon L., et al. (1969) "Compléments à l'étude des Trichomonadida et plus particulièrement de leur cinétide." Protistologica, 5:309-326.

K

Kaneda, Y., et al. (2001) "Ribodemes of *Blastocystis hominis* isolated in Japan." Am J Trop Med Hyg 65(4):393-6.

Katsarou-Katsari, A., et al. (2008) "Acute urticaria associated with amoeboid forms of *Blastocystis sp*. subtype 3." Acta Derm Venereol 88(1):80-1.

Keeling, P. J. (2002) "Molecular phylogenetic position of *Trichomitopsis termopsidis* (Parabasalia) and evidence for the Trichomitopsiinae." Europ. J. Protistol., 38:279-286.

Keeling, P. J. (2007) "Genomics. Deep questions in the tree of life." Science 317(5846): 1875-6.

Keeling, P. J. and Palmer J. D. (2000) "Parabasalian flagellates are ancient eukaryotes." Nature 405(6787): 635-7.

Keeling, P. J., et al. (1998) "Phylogenetic diversity of parabasalian symbionts from termites, including thephylogenetic position of *Pseudotrypanosoma* and *Trichonympha*." J Eukaryot Microbiol 45(6): 643-50.

Keeling, P. J., et al. (2005) "The tree of eukaryotes." Trends Ecol Evol 20(12): 670-6.

Keeling, P. J., and Doolittle, W. F. (1995) "Concerted evolution in protists: a polyubiquitin gene from *Trichomonas vaginalis.*" J. Mol. Evol., 41:556-562.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Kim, Y. A., et al. (2010) "PCR detection and molecular characterization of *Pentatrichomonas hominis* from feces of dogs with diarrhea in the Republic of Korea." Korean J Parasitol 48(1): 9-13.

Kirby, H. (1941) "Devescovinid flagellates of termites. I. The genus *Devescovina*." Univ. Calif. Publ. Zool., 45:1-92.

Krashin, J. W., et al. (2010) "*Trichomonas vaginalis* prevalence, incidence, risk factors and antibiotic-resistance in an adolescent population." Sex Transm Dis 37(7): 440-4.

Krieger, J. N. (1995) "Trichomoniasis in men: old issues and new data." Sex Transm Dis 22(2): 83-96.

Kubai, D. F. (1973) "Unorthodox mitosis in *Trichonympha agilis*: kinetochore differentiation and chromosome movement." J Cell Sci 13(2): 511-52.

Kulda J., et al. (1986) "Basic structure and function of the trichomonad cell." Acta Universitatis Carolinae-Biologica, 30:181-210.

Kulda, J. (1999) "Trichomonads, hydrogenosomes and drug resistance." Int J Parasitol 29(2): 199-212.

Kurniawan, A., et al. (2009) "Intestinal parasitic infections in HIV/AIDS patients presenting with diarrhoea in Jakarta, Indonesia." Trans R Soc Trop Med Hyg 103(9):892-8.

Kutisova, K., et al. (2005) "Tetratrichomonads from the oral cavity and respiratory tract of humans." Parasitology 131(Pt 3): 309-19.

L

Leelayoova, S., et al. (2008) "Drinking water: a possible source of *Blastocystis spp.* subtype 1 infection in schoolchildren of a rural community in central Thailand." Am J Trop Med Hyg 79(3):401-6.

Lehker, M. W. and Alderete J. F. (1990) "Properties of Trichomonas vaginalis grown under chemostat controlled growth conditions." Genitourin Med. 66(3):193-9.

Lehker, M. W. and Alderete J. F. (2000) "Biology of trichomonosis." Curr Opin Infect Dis 13(1): 37-45.

Leipe, D. D., et al. (1993) "Small subunit ribosomal RNA+ of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree." Molecular and Biochemical Parasitology 59(1): 41-48.

Levine, N. D., et al. (1980) "A newly revised classification of the protozoa." J Protozool 27(1): 37-58.

Lewis, K. L., et al. (2003) "Empyema caused by *Trichomonas*." Chest 123:291–292

Li, L. H., et al. (2007a) "Molecular epidemiology of human *Blastocystis* in a village in Yunnan province, China." Parasitol Int 56(4):281-6.

Li, L. H., et al. (2007b) "Cross-sectional surveys and subtype classification of human *Blastocystis* isolates from four epidemiological settings in China." Parasitol Res 102(1):83-90.

Liebhart, D., et al. (2006) "In-situ hybridization for the detection and identification of *Histomonas meleagridis* in tissues." J Comp Pathol 135(4): 237-42.

Lindmark, D. G. and Muller M. (1973) "Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism." J Biol Chem 248(22): 7724-8.

Lithgow, T. and Schneider A. (2010) "Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes." Philos Trans R Soc Lond B Biol Sci 365(1541): 799-817.

Long, H. Y., et al. (2001) "*Blastocystis hominis* modulates immune responses and cytokine release in colonic epithelial cells." Parasitol Res 87(12):1029-30.

Loo, C. K. and Guné, S. (2000) "*Trichomonas vaginalis* in urine cytology" Acta Cytol. 44(3):484-5.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

M

Mahmoud M. S. and Rahman G. A. (2004) "Pulmonary trichomoniasis: improved diagnosis by using polymerase chain reaction targeting *Trichomonas tenax* 18S rRNA gene in sputum specimens." J. Egypt Soc Parasitol. 34(1):197-211.

Malik, S-B., et al. (2011) "Phylogeny of prasitic Parabasalia and free living relatives inferred from conventional markers *vs. Rpb1*, a single-copy gene." PLoS ONE 6(6):e20774.

Mallat, H., et al. (2004) "Molecular characterization of *Trichomonas tenax* causing pulmonary infection." J Clin Microbiol 42(8): 3886-7.

Mantini C., et al. (2009) "Molecular characterization of a new *Tetratrichomonas* species in a patient with empyema." J Clin Microbiol 47: 2336-2339

Mantini C., et al. (2009) "Molecular identification and phylogenetic relationships of trichomonad isolates of galliform birds inferred from nuclear small subunit rRNA gene sequences. Parasitol Res. 106(1):163-70.

Mason, P. R., et al. (2005) "Seroepidemiology of *Trichomonas vaginalis* in rural women in Zimbabwe and patterns of association with HIV infection." Epidemiol Infect.133(2):315-23.

Mattern C. F. T. & Honigberg B. M. (1971) "Observations on the ondulatory motion of the costa in the Zooflagellate *Trichomitopsis termopsidis*." Trans. Amer. Micros. Soc. USA, 90:309-313.

McClelland, R. S., et al. (2007) "Infection with *Trichomonas vaginalis* increases the risk of HIV-1 acquisition." J Infect Dis 195(5): 698-702.

Meloni, B. P., et al. (1993) "The prevalence of *Giardia* and other intestinal parasites in children, dogs and cats from aboriginal communities in the Kimberley." Med J Aust 158(3): 157-9.

Mentel, M., et al. (2008) "Protein Import into Hydrogenosomes of *Trichomonas vaginalis* Involves both N-Terminal and Internal Targeting Signals: a Case Study of Thioredoxin Reductases." Eukaryotic Cell 7(10): 1750-1757.

Miller, W. C., et al. (2005) "The prevalence of trichomoniasis in young adults in the United States." Sex Transm Dis 32(10): 593-8.

Mirza, H. and Tan, K. S. (2009) "*Blastocystis* exhibits inter- and intra-subtype variation in cysteine protease activity." Parasitol Res 104(2):355-61.

Mirza, H., et al. (2010) "A Rapid, High-Throughput Viability Assay for *Blastocystis spp.* Reveals Metronidazole Resistance and Extensive Subtype-Dependent Variations in Drug Susceptibilities." Antimicrob Agents Chemother. 55(2):637-48.

Moe, K. T., et al. (1996) "Observations on the ultrastructure and viability of the cystic stage of *Blastocystis hominis* from human feces." Parasitol Res 82(5):439-44.

Moe, K. T., et al. (1997) "Experimental *Blastocystis hominis* infection in laboratory mice." Parasitol Res 83(4):319-25.

Moghaddam, D. D., et al. (2005) "*Blastocystis hominis* and the evaluation of efficacy of metronidazole and trimethoprim/sulfamethoxazole." Parasitol Res 96(4):273-5.

Moreira, D., et al. (2002) "Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations." Mol. Biol. Evol., 19:189-200.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

N

Nasirudeen, A. M. and Tan, K. S. (2004) "Isolation and characterization of the mitochondrion-like organelle from *Blastocystis hominis.*" J Microbiol Methods 58(1):101-9.

Niewiadomski, S., et al. (1998) "Cytologic identification of trichomonas vaginalis in urine from a male with long-standing sterile pyuria." Acta Cytol. 42(4):1060-1.

Nigro, L., et al. (2003) "A placebo-controlled treatment trial of *Blastocystis hominis* infection with metronidazole." J Travel Med 10(2):128-30.

Noda, S., et al. (2009) "Molecular phylogeny of Parabasalids with emphasis on the order Cristamonidida and its complex morphological evolution." Mol Phylogenetec Evol. 52:217-224.

Noel, C., et al. (2003) "Morphogenesis during division and griseofulvin-induced changes of the microtubular cytoskeleton in the parasitic protist, *Trichomonas vaginalis*." Parasitol Res 89(6): 487-94

Noel, C., et al. (2003) "Phylogenetic analysis of *Blastocystis* isolates from different hosts based on the comparison of small-subunit rRNA gene sequences." Mol Biochem Parasitol 126(1):119-23.

Noel, C., et al. (2005) "Molecular phylogenies of *Blastocystis* isolates from different hosts: implications for genetic diversity, identification of species, and zoonosis." J Clin Microbiol 43(1):348-55.

Noel, C., et al. (2007) "Molecular phylogenetic position of the genera Stephanonympha and Caduceia (Parabasalia) inferred from nuclear small subunit rRNA gene sequences" J Eukaryot Microbiol. 54(1):93-9.

O

Ogunsanya, T. I., et al. (1994) "A study of the aetiological agents of childhood diarrhoea in Lagos, Nigeria." J Med Microbiol 40(1): 10-4.

O'Hara, C. M., et al. (1980) "Immunoperoxidase staining of *Trichomonas vaginalis* in cytologic material." Acta Cytol 24(5): 448-51.

Ohkuma, M., et al. (1998) "Phylogenetic identification of the symbiotic hypermastigote *Trichonympha agilis* in the hindgut of the termite Reticulitermes speratus based on small-subunit rRNA sequence." J Eukaryot Microbiol 45(4): 439-44.

Ohkuma, M., et al. (2000) "Phylogenetic identification of hypermastigotes, *Pseudotrichonympha*, *Spirotrichonympha*, *Holomastigotoides*, and parabasalian symbionts in the hindgut of termites." J Eukaryot Microbiol 47(3): 249-59.

Ohkuma, M. (2003) "Termite symbiotic systems: efficient bio-recycling of lignocelluloses". Appl Microbiol Biotechnol. 61(1):1-9.

Ohkuma, M., et al. (2005) "Molecular phylogeny of parabasalids inferred from small subunit rRNA sequences, with emphasis on the Hypermastigea." Mol Phylogenet Evol 35(3): 646-55.

Ohkuma, M., et al. (2007) "Comparison of four protein phylogeny of parabasalian symbionts in termite guts." Molecular Phylogenetics and Evolution 42(3): 847-853.

Ohkuma, M. (2008) "Symbioses of flagellates and prokaryotes in the gut of lower termites." Trends Microbiol 16(7):345-52.

Ok, U. Z., et al. (1999) "Effect of trimethoprim-sulfamethaxazole in *Blastocystis hominis* infection." Am J Gastroenterol 94(11):3245-7.

Özyurt, M., et al. (2008) "Molecular epidemiology of *Blastocystis* infections in Turkey." Parasitol Int 57(3):300-6.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

P

Parkar, U., et al. (2007) "Direct characterization of *Blastocystis* from faeces by PCR and evidence of zoonotic potential." Parasitology 134(Pt 3):359-67.

Parkar, U., et al. (2010) "Molecular characterization of *Blastocystis* isolates from zoo animals and their animal-keepers." Vet Parasitol 169(1-2):8-17.

Pegelow, K., et al. (1997) "Parasitological and nutritional situation of school children in the Sukaraja district, West Java, Indonesia." Southeast Asian J Trop Med Public Health 28(1):173-90.

Perez-Brocal, V., and Clark, C. G. (2008) "Analysis of two genomes from the mitochondrion-like organelle of the intestinal parasite *Blastocystis*: complete sequences, gene content, and genome organization." Mol Biol Evol 25(11):2475-82.

Petrin, D., et al. (1998) "Clinical and microbiological aspects of *Trichomonas vaginalis*." Clin Microbiol Rev 11(2): 300-17.

Philippe, H. and Adoutte, A. (1995) "How reliable is our current view of eukaryotic phylogeny?" *In*: Protistological Actualities. Brugerolle G. & Mignot J.P. (eds.), Proceedings of the Second European Congress of Protistology, pp 17-33.

Philippe, H. and Germot A. (2000) "Phylogeny of eukaryotes based on ribosomal RNA: long-branch attraction and models of sequence evolution." Mol Biol Evol 17(5): 830-4.

Philippe, H., et al. (2000) "The new phylogeny of eukaryotes." Curr Opin Genet Dev 10(6): 596-601.

Philippe, H. and Laurent J. (1998) "How good are deep phylogenetic trees?" Curr Opin Genet Dev 8(6): 616-23.

Pindak, F. F., et al. (1989) "Acquisition and retention of viruses by *Trichomonas vaginalis*." Genitourin Med 65(6): 366-71.

Poirier, P., et al. (2011) "Development and evaluation of a real-time PCR assay for detection and quantification of blastocystis parasites in human stool samples: prospective study of patients with hematological malignancies." J. Clin Microbiol. 49(3):975-83.

Prasad, K. N., et al. (2000) "Identification of enteric pathogens in HIV-positive patients with diarrhoea in northern India." J Health Popul Nutr 18(1):23-6.

Puthia, M. K, et al. (2005) "Degradation of human secretory immunoglobulin A by *Blastocystis.*" Parasitol Res 97(5):386-9.

Puthia, M. K., et al. (2006) "*Blastocystis ratti* induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells." Infect Immun 74(7):4114-23.

Puthia, M. K., et al. (2008) "*Blastocystis ratti* contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner." Eukaryot Cell 7(3):435-43.

R

Rappelli, P., et al. (1998) "*Mycoplasma hominis* parasitism of *Trichomonas vaginalis*." Lancet 352(9136): 1286.

Rayan, H. Z., et al. (2007) "Prevalence and clinical features of *Dientamoeba fragilis* infections in patients suspected to have intestinal parasitic infection." J Egypt Soc Parasitol 37(2):599-608.

Rene, B. A., et al. (2009) "Subtype analysis of *Blastocystis* isolates from *Blastocystis* cyst excreting patients." Am J Trop Med Hyg 80(4):588- 92.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Riisberg, I., et al. (2009) "Seven gene phylogeny of heterokonts." Protist 160(2):191-204.

Rivera, W. L. (2008) "Phylogenetic analysis of *Blastocystis* isolates from animal and human hosts in the Philippines." Vet Parasitol 156(3-4):178-82

Rivera, W. L. and Tan, M. A. (2005) "Molecular characterization of *Blastocysti*s isolates in the Philippines by riboprinting." Parasitol Res 96(4):253-7.

Rossignol, J. K., et al. (2005) "Effect of nitazoxanide in persistent diarrhea and enteritis associated with *Blastocystis hominis*." Clin Gastroenterol Hepatol 3(10):987-91.

Roger, A. J., et al. (1999a) "An evolution of elongation factor 1α as a phylogenetic marker for eukaryotes." Mol. Biol. Evol., 16:218-233.

Rybicka K., et al. (1972) "Fine structure of the mastigont system in culture forms of *Histomonas meleagridis* (Smith)." Protistologica, 8:107-120.

S

Sajid, M. and McKerrow, J. H. (2002) "Cysteine proteases of parasitic organisms." Mol Biochem Parasitol 120(1):1-21.

Saksirisampant, W., et al. (2003) "Intestinal parasitic infestations among children in an orphanage in Pathum Thani province." J Med Assoc Thai 86 Suppl 2:S263-70.

Saksirisampant, W., et al. (2006) "Prevalence of intestinal parasitic infections among school children in the central region of Thailand." J Med Assoc Thai 89(11):1928-33.

Santin, M., et al. (2011) "Development of a new PCR protocol to detect and subtype *Blastocystis spp*. from humans and animals." Parasitol Res. 109(1):205-12.

Sauvage, V., et al. (2009) "The role of ATP-binding cassette (ABC) proteins in protozoan parasites." Mol Biochem Parasitol 167(2):81-94.

Scicluna, S. M., et al. (2006) "DNA barcoding of *blastocystis*." Protist 157(1):77-85.

Schwebke, J. R. and Hook E. W., 3rd (2003) "High rates of *Trichomonas vaginalis* among men attending a sexually transmitted diseases clinic: implications for screening and urethritis management." J Infect Dis 188(3): 465-8.

Shiflett, A. M. and Johnson P. J. (2010) "Mitochondrion-Related Organelles in Eukaryotic Protists." Annu Rev Microbiol.

Silberman, J. D., et al. (1996) "Human parasite finds taxonomic home." Nature 380(6573):398.

Simpson, A. G., et al. (2006) "Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes." Mol Biol Evol 23(3): 615-25.

Simpson, A. G. and Roger A. J. (2004) "The real 'kingdoms' of eukaryotes." Curr Biol 14(17): R693-6.

Sio, S. W., et al. (2006) "Protease activity of *Blastocystis hominis*." Parasitol Res 99(2):126-30.

Smith, L. M., et al. (2002) "*Trichomonas vaginalis* infection in a premature newborn." J Perinatol 22(6): 502-3.

Sogin M.L. (1989) "Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs." Amer. Zool., 29:487-499.

Sogin, M. L. (1991) "Early evolution and the origin of eukaryotes." Curr Opin Genet Dev 1(4): 457-63.

Sogin, M. (1997) "History assignment: when was the mitochondrion founded?"Curr Opin Genet Dev 7(6): 792-9.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Sogin, M. L., et al. (1996) "Ancestral relationships of the major eukaryotic lineages." Microbiologia 12(1): 17-28.

Sogin, M. L. and. Silberman J. D (1998) "Evolution of the protists and protistan parasites from the perspective of molecular systematics." International Journal for Parasitology 28(1): 11-20.

Soper, D. (2004) "Trichomoniasis: under control or undercontrolled?" Am J Obstet Gynecol 190(1):281-90.

Souppart, L., et al. (2009) "Molecular epidemiology of human *Blastocystis* isolates in France." Parasitol Res 105(2):413-21.

Souppart, L., et al. (2010) "Subtype analysis of *Blastocystis* isolates from symptomatic patients in Egypt." Parasitol Res 106(2):505-11.

Stark, D. J., et al. (2006) "Dientamoebiasis: clinical importance and recent advances." Trends in Parasitology 22(2): 92-96.

Stechmann, A., et al. (2008) "Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes." Curr Biol 18(8):580-5.

Stensvold, C. R., et al. (2007a) "Terminology for *Blastocystis* subtypes--a consensus." Trends Parasitol 23(3):93-6.

Stensvold, C. R., et al. (2007b) "Detecting *Blastocystis* using parasitologic and DNA-based methods: a comparative study." Diagn Microbiol Infect Dis 59(3):303-7.

Stensvold, C. R., et al. (2008) "Symptomatic infection with *Blastocystis sp.* subtype 8 successfully treated with trimethoprim-sulfamethoxazole." Ann Trop Med Parasitol 102(3):271-4.

Stensvold, C. R., et al. (2009a) "Pursuing the clinical significance of *Blastocystis*-diagnostic limitations." Trends Parasitol 25(1):23-9.

Stensvold, C. R., et al. (2009b) "Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype." Int J Parasitol 39(4):473-9.

Stensvold, C. R., et al. (2009c) "*Blastocystis*: unravelling potential risk factors and clinical significance of a common but neglected parasite." Epidemiol Infect 137(11):1655-63.

Stensvold, C. R., et al. (2010) "Eradication of *Blastocystis* carriage with antimicrobials: reality or delusion?" J Clin Gastroenterol 44(2):85-90.

Stensvold, C. R. and Nielsen, H. V. (2011) "Comparison of microscopy and PCR for the detection of intestinal parasites in Danish patients supports incentive for molecular screening platforms." J Clin Microbiol [Epub ahead of print].

Stensvold, C. R., et al. (2011) "Blastocystis sp. subtype 4 is common in Danish Blastocystis-positive patients presenting with acute diarrhea." Am J Trop Med Hyg. 84(6):883-5.

Stensvold, R., et al. (2006) "Detection of Blastocystis hominis in unpreserved stool specimens by using polymerase chain reaction." J. Parasitol. 92(5):1081-7.

Stenzel, D. J. and Boreham, P. F. (1996) "*Blastocystis hominis* revisited." Clin Microbiol Rev 9(4):563-84.

Stratakis, D. F., et al. (1999). "[Pulmonary trichomoniasis: diagnosis based on identification of irritation in bronchoalveolar lavage]." Pneumologie 53(12): 617-9.

Straube, E., et al. (1991). "Peritonitis caused by *Trichomonas*-a preliminary case report." Zentralbl Bakteriol 275(3): 394-6.

Suresh, K., and Smith, H. (2004) "Comparison of methods for detecting *Blastocystis hominis*." Eur J Clin Microbiol Infect Dis 23(6):509-11.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Suresh, K., et al. (1993) "In vitro encystment and experimental infections of *Blastocystis hominis*." Parasitol Res 79(6):456-60.

Suresh, K., et al. (2005) "Viable *blastocystis* cysts in Scottish and Malaysian sewage samples." Appl Environ Microbiol 71(9):5619-20.

Suresh, K., et al. (2009) "*In vivo* encystation of *Blastocystis hominis*." Parasitol Res 104(6):1373-80.

Sutton, M. Y., et al. (1999) "Trichomoniasis in pregnant human immunodeficiency virus-infected and human immunodeficiency virus-uninfected Congolese women: prevalence, risk factors, and association with low birth weight." Am J Obstet Gynecol 181(3): 656-62.

Szarka, K., et al. (2002) "Neonatal pneumonia caused by *Trichomonas vaginalis*." Acta Microbiol Immunol Hung 49(1): 15-9.

T

Tachezy, J., et al. (2002) "Cattle pathogen *tritrichomonas foetus* (Riedmüller, 1928) and pig commensal *Tritrichomonas suis* (Gruby & Delafond, 1843) belong to the same species."J Eukaryot Microbiol. 49(2):154-63

Tai, W. P., et al. (2010) "Six ulcerative colitis patients with refractory symptoms co-infective with *Blastocystis hominis* in China." Parasitol Res. 108(5):1207-10.

Tan, K. S. (2004) "*Blastocystis* in humans and animals: new insights using modern methodologies." Vet Parasitol 126(1-2):121-44.

Tan, K. S. (2008) "New insights on classification, identification, and clinical relevance of *Blastocystis spp.*" Clin Microbiol Rev 21(4):639-65.

Tan, T. C. and Suresh, K. G. (2006) "Predominance of amoeboid forms of *Blastocystis hominis* in isolates from symptomatic patients." Parasitol Res 98(3):189-93.

Tan, T. C., et al. (2008) "Phenotypic and genotypic characterisation of *Blastocystis hominis* isolates implicates subtype 3 as a subtype with pathogenic potential." Parasitol Res 104(1):85-93.

Tan, T. C., et al. (2009) "Genetic variability of *Blastocystis sp.* isolates obtained from cancer and HIV/AIDS patients." Parasitol Res 105(5):1283-6.

Tasova, Y., et al. (2000) "Clinical significance and frequency of *Blastocystis hominis* in Turkish patients with hematological malignancy." Acta Med Okayama 54(3):133-6.

Thathaisong, U., et al. (2003) "*Blastocystis* isolates from a pig and a horse are closely related to *Blastocystis hominis*." J Clin Microbiol 41(3):967-75.

The ANOFEL Cryptosporidium National Network. Laboratory-based surveillance for Cryptosporidium in France, 2006–2009. EuroSurveill. 2010; 15(33):pii=19642. Available on line:http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19642.

U

Upcroft, J. A., et al. (2009) "Metronidazole resistance in *Trichomonas vaginalis* from highland women in Papua New Guinea." Sex Health 6(4): 334-8.

Upcroft, P. and Upcroft J. A. (2001) "Drug Targets and Mechanisms of Resistance in the Anaerobic Protozoa." Clin. Microbiol. Rev. 14(1): 150-164.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

V

Van Der Pol, B. (2007) "Editorial Commentary*: Trichomonas vaginalis* Infection: The Most Prevalent Nonviral Sexually Transmitted Infection Receives the Least Public Health Attention." Clinical Infectious Diseases 44(1): 23-25.

Viikki, M., et al. (2000) "Gynaecological infections as risk determinants of subsequent cervical neoplasia." Acta Oncol 39(1): 71-5.

Viscogliosi E. and Brugerolle G. (1994a) "Cytoskeleton in trichomonads." III. Study of the morphogenesis during division by using monoclonal antibodies against cytoskeletal structures. Europ. J. Protistol., 30:129-138.

Viscogliosi, E. and G. Brugerolle (1994b) "Striated fibers in trichomonads: costa proteins represent a new class of proteins forming striated roots." Cell Motil Cytoskeleton 29(1): 82-93.

Viscogliosi, E. and Muller M. (1998) "Phylogenetic relationships of the glycolytic enzyme, glyceraldehyde-3-phosphate dehydrogenase, from parabasalid flagellates." J Mol Evol 47(2): 190-9.

Viscogliosi, E., et al. (1996) "Phylogenetic implication of iron-containing superoxide dismutase genes from trichomonad species." Mol Biochem Parasitol 80(2): 209-214.

Vogel C and Morea V. (2006) "Duplication, divergence and formation of novel protein topologies." 28(10):973-8.

Vogelberg, C., et al. (2010) "*Blastocystis* sp. subtype 2 detection during recurrence of gastrointestinal and urticarial symptoms." Parasitol Int 59(3):469-71**.**

W

Wang, H.- K., et al. (2006) Case report: *Trichomonas* empyema with respiratory failure. Am. J. Trop. Med. Hyg. 75:1234-1236

Wawrzyniak, I., et al. (2008) "Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis*." Int J Parasitol 38(12):1377-82.

Whipps, C. M., et al. (2010) "Molecular characterization of *Blastocystis* species in Oregon identifies multiple subtypes." Parasitol Res 106(4):827-32.

Wong, K. H., et al. (2008) "Predominance of subtype 3 among *Blastocystis* isolates from a major hospital in Singapore." Parasitol Res 102(4):663-70.

Wu, B., et al. (2010) "*Blastocystis legumain* is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme." J Biol Chem 285(3):1790-8.

Y

Yakoob, J., et al. (2004a) "Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*." Am J Trop Med Hyg 70(4):383-5.

Yakoob, J., et al. (2004b) "*In vitro* susceptibility of *Blastocystis hominis* isolated from patients with irritable bowel syndrome." Br J Biomed Sci 61(2):75-7.

Yakoob, J., et al. (2010a) "Irritable bowel syndrome: is it associated with genotypes of *Blastocystis hominis*." Parasitol Res 106(5):1033-8.

Yamamoto-Furusho, J. K. and Torijano-Carrera, E. (2010) "Intestinal protozoa infections among patients with ulcerative colitis: prevalence and impact on clinical disease course." Digestion 82(1):18-23.

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

Yamin M.A. (1979) "Flagellates of the orders Trichomonadida Kirby, Oxymonadida Grassé, and Hypermastigida Grassi & Foà reported from lower termites (Isoptera families Mastotermitidae, Kalotermitidae, Hodotermitidae, Termopsidae, Rhinotermitidae, and Serritermitidae) and from the wood-feeding roach *Cryptocercus* (Dictyoptera: Cryptocercidae)." Sociobiology, 4:5-119.

Yan, Y., et al. (2006) "Genetic variability of *Blastocystis hominis* isolates in China." Parasitol Res 99(5):597-601.

Yan, Y., et al. (2007) "*Blastocystis sp.* subtype 5: a possibly zoonotic genotype." Parasitol Res 101(6):1527-32.

Yang, C. R., et al. (1990) "Diarrhoea surveillance in children aged under 5 years in a rural area of Hebei Province, China." J Diarrhoeal Dis Res 8(4): 155-9.

Yoshikawa, H., et al. (1998) "Genomic polymorphism among Blastocystis hominis strains and development of subtype-specific diagnostic primers." Mol Cell Probes. 12(3):153-9.

Yoshikawa, H., et al. (2000) "Genomic analysis *of Blastocystis hominis* strains isolated from two long-term health care facilities." J Clin Microbiol 38(4):1324-30.

Yoshikawa, H., et al. (2004a) "Fecal-oral transmission of the cyst form of *Blastocystis hominis* in rats." Parasitol Res 94(6):391-6.

Yoshikawa, H., et al. (2004b) "Polymerase chain reaction-based genotype classification among human *Blastocystis hominis* populations isolated from different countries." Parasitol Res 92(1):22-9.

Yoshikawa, H., et al. (2004c) "PCR-based identification of zoonotic isolates of *Blastocystis* from mammals and birds." Microbiology 150(Pt 5):1147-51.

Yoshikawa, H., et al. (2009) "Molecular characterization of *Blastocystis* isolates from children and rhesus monkeys in Kathmandu, Nepal." Vet Parasitol 160(3-4):295-300.

Yoshikawa, H., et al. (2009) "Molecular characterization of *Blastocystis* isolates from children and rhesus monkeys in Kathmandu, Nepal." Vet Parasitol 160(3-4):295-300.

Yule, A., et al. (1989) "Bovine trichomoniasis." Parasitol Today 5(12): 373-7.

Z

Zaman, V., and Zaki, M. (1996) "Resistance *of Blastocystis hominis* cysts to metronidazole." Trop Med Int Health 1(5):677-8.

Zhang, Z. F. and Begg C. B. (1994) "Is *Trichomonas vaginalis* a cause of cervical neoplasia? Results from a combined analysis of 24 studies." Int J Epidemiol 23(4): 682-90.

Appendix

Appendix

Article 7 (see above)

Monchy S, Sanciu G, Jobard M, Rasconi S, Gerphagnon M, Chabé M, Cian A, <u>Meloni D</u>, Niquil N, Christaki U, Viscogliosi E, Sime-Ngando T. Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. Environmental Microbiology, 2011, 13:1433-1453

Article 8

Monchy S, Grattepanche JD, Breton E, Sanciu G, <u>Meloni D</u>, Chabé M, Delhaes L, Dolan J, Viscogliosi E, Sime-Ngando T, Christaki U. Microplanktonic community structure in a coastal system characterised by *Phaeocystis* blooms, inferred tag pyrosequencing data and microscope observations. PLoS ONE, 2011, in preparation

The first year of my thesis was funded by an Erasmus grant and by a grant from my host laboratory in France. In this context, in parallel with my studies on the biology and diversity of parabasalids and *Blastocystis* sp., I spent several months in projects focused on the environmental biodiversity of some eukaryotic groups in pelagic **(Article 7)** and marine **(Article 8)** ecosystems using metagenomic approaches.

Briefly, in **Article 7**, the eukaryotic and especially fungal biodiversity of two French Lakes was analyzed using two approaches: the classical approach consisting of cloning/sequencing of the 18S, ITS1, 5.8S, ITS2 and partial 28S region using primers designed for fungus sequences, and the pyrosequencing of 18S rRNA hypervariable regions using two primer sets (one universal for eukaryotes and one for fungi). In both lakes and methods, fungi were dominated by sequences from Chytridiomycota. Besides fungi, both approaches revealed other major eukaryote groups, with the highest diversity in the central areas of lakes. This study represents the first unveiling of microbial eukaryote and fungus diversity assessed with two complementary molecular methods, and is considered a major milestone towards understanding the dynamics and ecology of fungi in freshwater lake ecosystems, which are directly link to the abundance and distribution of taxa.

In **Article 8**, our study site was the Eastern English Channel, a marine ecosystem very important for fisheries and characterised by spring blooms of the prymnesiophyte *Phaeocystis globosa* and a diverse community of colonial diatoms. Some recent studies have also shown the importance *P. globosa* bloom on the community structure shifts of

Nome e cognome: Dionigia Meloni
Titolo della tesi: Molecular identification and evolution of protozoa belonging to the Parabasalia group and the genus *Blastocystis*
Tesi di dottorato in scienze Biomolecolari e biotecnologiche. Indirizzo: Microbiologia molecolare e clinica
Università degli studi di Sassari

heterotrophic prokaryote and heterotrophic eukaryote communities. As an extension to these studies, the objective of the present study was to determine the diversity of protozoa relative to the spring bloom of nano- and micro-eukaryotes using tag pyrosequencing. This overall objective translates into the following specific tasks: (i) To determine if some of the forms or "morphospecies" recognised as key organisms in carbon fluxes and putatively taxonomically identified to genus or species, are genetically one or more genera or species. Particular attention was given in the analysis of infra-specific diversity of Haptophyceae and the determination and quantification of abundant and rare *P. globosa* species since this group represents more than 90 % of the auto- and heterotrophic microplanktonic biomass during the bloom every year and (ii) to determine the complimentarily between molecular and traditional monitoring combining tag pyrosequencing data and microscope observations.

Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing

Sébastien Monchy,[1,2,3,4,5] Giovanna Sanciu,[2,3,4,5]
Marlène Jobard,[1] Serena Rasconi,[1]
Mélanie Gerphagnon,[1] Magali Chabé,[2,3,4,5]
Amandine Cian,[2,3,4,5] Dionigia Meloni,[2,3,4,5]
Nathalie Niquil,[6] Urania Christaki,[7]
Eric Viscogliosi[2,3,4,5] and Télesphore Sime-Ngando[1]*

[1]LMGE, Laboratoire Microorganismes: Génome et Environnement, UMR CNRS 6023, Clermont Université, Blaise Pascal, BP 80026, 63171 Aubière Cedex, France.
[2]Center for Infection and Immunity of Lille, Institut Pasteur de Lille, Biology and Diversity of Emerging Eukaryotic Pathogens (BDEEP), 1 rue du Professeur Calmette, BP 245, 59019 Lille cedex, France.
[3]Inserm U1019, 59000 Lille cedex, France.
[4]CNRS UMR 8204, 59021 Lille cedex, France.
[5]University Lille Nord de France, 59000 Lille cedex, France.
[6]LIENSs, Littoral, Environnement et Sociétés, UMR CNRS 6250, Université de La Rochelle, 2 rue Olympe de Gouge, 17000 La Rochelle, France.
[7]ULCO, Laboratoire d'Océanologie et Géoscience, UMR CNRS 8187, Université de Lille-Nord, 62930 Wimereux, France.

Summary

Water samples were collected along transects from the shore to the centre of two French lakes: the deep, volcanic, oligomesotrophic and low allochthonic-impacted Lake Pavin, and the productive and higher allochthonic-impacted Lake Aydat. The biodiversity was analysed using two approaches: the classical approach consisting of cloning/sequencing of the 18S, ITS1, 5.8S, ITS2 and partial 28S region using primers designed for fungus sequences, and the pyrosequencing of 18S rRNA hypervariable V2, V3 and V5 regions using two primer sets (one universal for eukaryotes and one for fungi). The classical approach yielded 146 (Lake Pavin) and 143 (Lake Aydat) sequences, corresponding to 46 and 63 operational taxonomic units (OTUs) respectively. *Fungi* represented half of the OTUs identified in Lake Pavin and 30% in Lake Aydat, and were dominated by sequences from *Chytridiomycota* found throughout Lake Pavin but mostly in the central pelagic zone of Lake Aydat. The pyrosequencing approach yielded 42 064 (Pavin) and 61 371 (Aydat) reads, of which 12–15% and 9–19% reads were assigned to fungi in Lakes Pavin and Aydat respectively. *Chytridiomycota* members were also dominant among these reads, with OTUs displaying up to > 33-fold overrepresentation in the centre compared with the riparian areas of Lake Aydat. Besides fungi, both approaches revealed other major eukaryote groups, with the highest diversity in the central areas of lakes. One of the major findings of our study was that the two lakes displayed contrasting spatial distributions, homogenous for Lake Pavin and heterogeneous for Lake Aydat, which may be related to their peculiarities. This study represents the first unveiling of microbial eukaryote and fungus diversity assessed with two complementary molecular methods, and is considered a major milestone towards understanding the dynamics and ecology of fungi in freshwater lake ecosystems, which are directly link to the abundance and distribution of taxa.

Introduction

Fungi represent one of the last frontiers of the undiscovered biodiversity that challenge the microbial ecology today. They have a crucial role in ecosystem functioning – or are crucial for the maintenance of ecological balance since they influence many environmental processes such as nutrient cycling in the food web. From studies in soil systems, fungi are known to be vital in recycling nutrients through the metabolism of complex organic materials (Treseder, 2005; Watling, 2005). In pelagic environments, we also hypothesize an important role of fungi in many ecosystem processes. However, our ecological knowledge of fungi in pelagic ecosystems is scant. The kingdom *Fungi* was previously subdivided into four main phyla:

Basidiomycota, *Ascomycota*, *Zygomycota* and *Chytridiomycota* (Alexopoulos *et al.*, 1996). But the two latter phyla do not emerge as monophyletic groups in recent phylogenetic analyses (James *et al.*, 2000; 2006a). Today, seven phyla are recognized in the phylogenetic classification of *Fungi* (Hibbett *et al.*, 2007). The *Chytridiomycota* is retained in a restricted sense, with *Neocallimastigomycota* and *Blastocladiomycota* representing segregate phyla of flagellated *Fungi*. Taxa traditionally placed in *Zygomycota* are distributed among *Glomeromycota* and several subphyla *incertae sedis*, including *Mucoromycotina*, *Entomophtoromycotina*, *Kickxellomycotina* and *Zoopagomycotina* (Hibbett *et al.*, 2007). The number of fungi present on earth was estimated to 1.5 million species (Hawksworth, 2001), from which approximately 97 000 have so far been identified (Kirk and Ainsworth, 2008). This suggests a large underestimation of the diversity of this eukaryotic group, because of methodological constraints, but also partly because of the undersampling of pelagic environments, which represent the largest reservoir for life on earth.

In aquatic systems, several culturing assays allowed the isolation of ascomycetous or basidiomycetous yeasts. Freshwater *Ascomycetes* occur in both lentic (Hyde and Goh, 1998a; 1999) and lotic habitats (Hyde and Goh, 1998b,c; 1999) and more than 500 species have been reported so far (Shearer, 2001; Shearer *et al.*, 2007). Although ascomycetous fungi, together with bacteria, are efficient decomposers of organic matter, their functional importance in lakes remains unclear (Raja and Shearer, 2008). Basidiomycetous yeasts were retrieved in marine habitats, particularly in deep waters (Bass *et al.*, 2007), and from high-altitude cold lakes (Libkind *et al.*, 2009).

Fungi are often thought to be present in pelagic systems mainly as allochthonous terrestrial inputs, which is usually not the case for *Chytridiomycota*, i.e. chytrids, which are known to produce zoospores (Wong *et al.*, 1998). There are over 1000 species of chytrids (Powell, 1993; Shearer *et al.*, 2007); many of them have been characterized as parasites of freshwater algae (Gromov *et al.*, 1999a,b; Ibelings *et al.*, 2004). One chytrid species (*Batrachochytrium dendrobatidis*) was also described as the causative agent of die-offs and population declines of amphibian species (Berger *et al.*, 1998). Previous studies have unveiled a large reservoir of unsuspected fungal diversity in pelagic zone of lakes (Van Donk and Ringelberg, 1983; Holfeld, 1998; Hyde and Goh, 1998a; Richards *et al.*, 2005), primarily of chytrids (Lefèvre *et al.*, 2007; 2008; Sime-Ngando *et al.*, 2011) known as aquatic life forms, as their propagules (i.e. uniflagellated zoospores) have specialized aquatic dispersal abilities. Pelagic freshwater species of chytrids mostly fall within the *Rhizophydiales* order, which contains many host-

specific parasitic fungi of various phytoplankton species, primarily diatoms (Jobard *et al.*, 2010; Rasconi *et al.*, 2011). These organisms are known to often reduce phytoplankton blooms and to play a crucial role in phytoplankton succession (Kagami *et al.*, 2004; 2007; Rasconi *et al.*, 2011). Several studies have demonstrated epidemic occurrences of chytrids infecting phytoplankton in lakes (Canter and Lund, 1948; 1951; Canter, 1950; Van Donk and Ringelberg, 1983; Kudoh, 1990; Holfeld, 1998).

In the deep-volcanic oligomesotrophic Lake Pavin (France), which is a typical pelagic environment because of the absence of river inflow and low human influences, the occurrence of these parasitic fungi was confirmed from direct microscopic observation of their sporangia (i.e. the infective stage) fixed on various classes of phytoplankton, including diatoms and chlorophytes (Rasconi *et al.*, 2009). Chytrids can be characterized by their zoospores that typically have a single, posteriorly directed flagellum (James *et al.*, 2000), corresponding to one of the major forms observed directly via epifluorescence microscopy. However, because chytrid life forms are not distinguishable from some small flagellated protists, they may have been miscounted as bacterivorous flagellates in previous studies (Carrias *et al.*, 1996; Carrias and Amblard, 1998). Carrias and colleagues (1996) reported that the dominant small unidentified heterotrophic flagellates present in Lake Pavin during a seasonal study were not able to ingest bacteria. Large grazing-resistant diatoms, primarily *Aulacoseira italica* and *Asterionella formosa*, are important components of Lake Pavin, where they typically account for 50–98% of the total phytoplankton biomass production during the spring bloom in the whole water column (Rasconi *et al.*, 2009). These diatoms are well known as preferential hosts for chytrids in Lake Pavin. Similarly, filamentous cyanobacterial blooms, often dominated by the highly chytrid-sensible species *Anabaena flosaquae*, are characteristics of autumn in the eutrophic Lake Aydat, France (Rasconi *et al.*, 2009). The activity of fungi is suspected to be an important factor in Lake Pavin, Lake Aydat, and other pelagic inland and coastal ecosystems, which regulates phytoplankton community structure and species successions. It has also been shown that the small flagellated zoospores produced by parasitic chytrids are efficiently grazed by zooplankton such as *Daphnia* (Kagami *et al.*, 2004). Such considerations suggest that chytrids are important microbial players in pelagic ecosystems, where they are involved both in the decomposition of organic matter from parasitic epidemics or saprotrophy, and in the transfer of matter and energy to higher trophic levels (Lefèvre *et al.*, 2008).

However, still little is known about the diversity of fungi in pelagic ecosystems. Most of the related studies are relatively recent (Richards *et al.*, 2005), including those in

Lakes Aydat and Pavin (Lefranc *et al.*, 2005; Lefèvre *et al.*, 2007; 2008). The latter studies were performed initially with the main aim to examine the phylogenetic composition of picoeukaryotes (< 5 µm), and ended up with the unveiling of an important diversity of fungi. For example, in Lakes Aydat and Pavin, studies on the diversity of small heterotrophic eukaryotes based on small subunit (SSU) rRNA gene sequences have shown that *Alveolata*, *Cercozoa*, *Stramenopile* and *Fungi* represented the major part of the total diversity (Lefèvre *et al.*, 2007; 2008). However, the molecular approach used in these studies, i.e. polymerase chain reaction (PCR) amplification and cloning of the entire SSU rDNA gene followed by analysis of the positive clones by restriction fragment length polymorphism (RFLP), and sequencing of the selected clones or operational taxonomic units (OTUs), represents a both costly and tedious method, providing only a limited analysis of the diversity associated with the number of clones analysed. Therefore, we have taken a complementary approach, by pyrosequencing of hypervariable regions of the SSU rRNA gene from selected samples. This method, already successfully used for the exploration of deep-sea community composition (Sogin *et al.*, 2006; Huse *et al.*, 2008), is cost-effective and offers around three orders of magnitude larger SSU rRNA sequencing, compared with classical approaches. In addition, it eliminates cloning bias and allows maximizing the number of organisms sampled in a run (Huse *et al.*, 2008).

The aim of this work was: (i) to determine the diversity of microbial eukaryotes, focusing on fungi, in the whole plankton fraction (size class 0.6–150 µm) using two complementary methodological approaches, (ii) to compare this diversity with earlier studies performed from the same lakes, (iii) to estimate the spatial variability in the diversity of pelagic fungi from samples collected along transects from the shore to the centre of the lakes, and (iv) to determine the differences in eukaryote and fungus diversity between two lakes differing by their trophic status. To our knowledge, this study is the first exploration of the spatial dynamics in the diversity of pelagic fungi using a combination of two complementary approaches. The complete sequencing of the SSU rDNA gene, considered as the classical approach, attributed to each sequence a precise taxonomical position and identified potential new clades, based on the use of fungus-designed primers. The second approach, consisting of pyrosequencing a large number of SSU rRNA hypervariable tags, provided a general overview of the eukaryote and fungus diversity, with the use of a universal eukaryote and a fungus-designed set of primers. This second method, referred to as the pyrosequencing approach, unveiled rare species, and gave quantitative information about the OTUs identified by the classical approach. In this study, we have unveiled all the major and minor fungus taxa and their spatial variability (i.e. between and within lakes), which can be considered as the first step towards the understanding of population dynamics and functional role of fungi in pelagic food webs.

Results

Overview of the retrieved diversity

The 'classical approach' yielded a total of 146 (Lake Pavin) and 143 (Lake Aydat) complete SSU rDNA gene sequences, considered an overview of the biodiversity in the pilot lakes at the sampling time. After clustering sequences sharing more than 99% identity, the total number of OTUs increased from 46 OTUs in Lake Pavin (Fig. 1) to 63 OTUs in Lake Aydat (Fig. 2). In Lake Pavin, half of the OTUs identified belonged to *Fungi* with 15, 7 and 1 representatives (i.e. OTUs) from the *Chrytridiomycota*, *Ascomycota* and *Basidiomycota* phyla respectively (Fig. 1, Table S1a). The other represented taxonomical groups comprised SSU rDNA gene OTUs from 11 *Viridiplantae*, 5 *Cryptophyta*, 4 *Alveolata*, 1 *Telonemida*, 1 *Ichtyosporea* and 1 *Stramenopile* (Fig. 1). In Lake Aydat, one-third of the OTUs belonged to *Fungi*, including 10 *Chytridiomycota*, 8 *Ascomycota* and 1 *Basidiomycota* (Fig. 2 and Table S1b). The diversity of chytrids was 50% lower in Lake Aydat compared with Lake Pavin. Similar to the findings in Lake Pavin, other eukaryotic groups were found in Aydat, with SSU rDNA gene OTUs corresponding to 21 *Alveolata*, 9 *Viridiplantae*, 6 *Cryptophyta*, 2 *Ichtyospora* and 2 *Stramenopile* (Fig. 2 and Table S1b). In addition, three *Choanoflagellida* and one *Cercozoan* OTUs were only retrieved from Lake Aydat (Fig. 2 and Table S1b). Within all the identified OTUs, only five were common to both lakes: two *Cryptophyta*, one *Alveolata* and two *Fungi* (one *Ascomycota* and one *Chytridiomycota*).

For both lakes, the rarefaction curves with > 97 and > 99% levels of sequence similarities did not reach saturation (Fig. 3A), indicating biodiversity underestimation using the classical approach. This result justifies the implementation of a complementary pyrosequencing approach, to improve the exploration of the hidden biodiversity, including rare species. With this approach, a total of 23 519 and 18 545 sequences from Lake Pavin, and of 30 337 and 31 034 sequences from Lake Aydat, using eukaryote and fungus-designed primers, were, respectively, obtained. For both lakes and for all sampling stations, the rarefaction curves continued to rise almost linearly with > 99% level of sequence similarity (data not shown), but approaching a plateau if applying > 97%, the apparent diversity being higher when using the fungus set of primers (Fig. 3B).
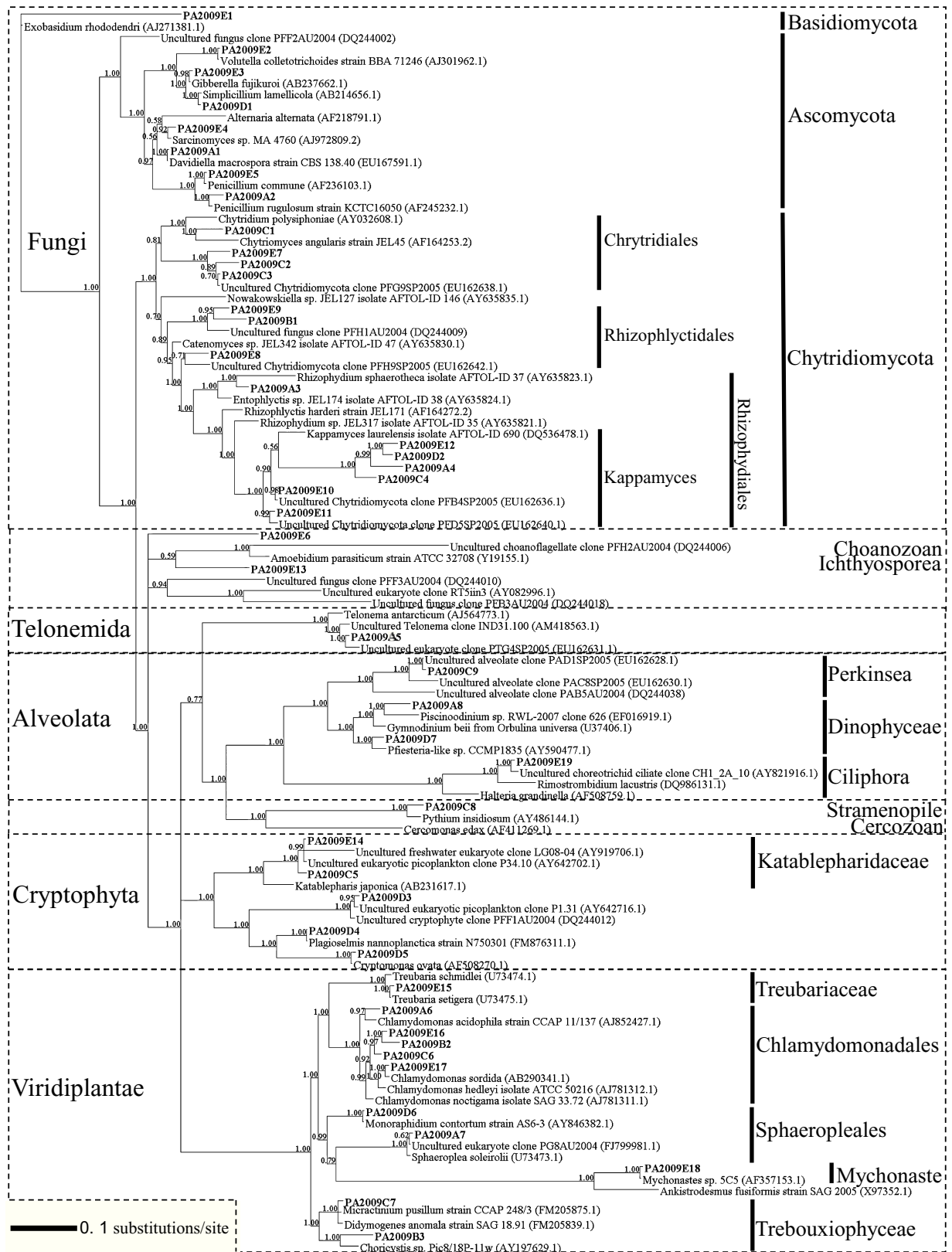
PA2009E1
Exobasidium rhododendri (AJ271381.1)
Uncultured fungus clone PFF2AU2004 (DQ244002)
PA2009E2
Volutella colletotrichoides strain BBA 71246 (AJ301962.1)
PA2009E3
Gibberella fujikuroi (AB237662.1)
Simplicillium lamellicola (AB214656.1)
PA2009D1
Alternaria alternata (AF218791.1)
PA2009E4
Sarcinomyces sp. MA 4760 (AJ972809.2)
PA2009A1
Davidiella macrospora strain CBS 138.40 (EU167591.1)
PA2009E5
Penicillium commune (AF236103.1)
PA2009A2
Penicillium rugulosum strain KCTC16050 (AF245232.1)
Chytridium polysiphoniae (AY032608.1)
PA2009C1
Chytriomyces angularis strain JEL45 (AF164253.2)
PA2009E7
PA2009C2
PA2009C3
Uncultured Chytridiomycota clone PFG9SP2005 (EU162638.1)
Nowakowskiella sp. JEL127 isolate AFTOL-ID 146 (AY635835.1)
PA2009E9
PA2009B1
Uncultured fungus clone PFH1AU2004 (DQ244009)
Catenomyces sp. JEL342 isolate AFTOL-ID 47 (AY635830.1)
PA2009E8
Uncultured Chytridiomycota clone PFH9SP2005 (EU162642.1)
Rhizophydium sphaerotheca isolate AFTOL-ID 37 (AY635823.1)
PA2009A3
Entophlyctis sp. JEL174 isolate AFTOL-ID 38 (AY635824.1)
Rhizophlyctis harderi strain JEL171 (AF164272.2)
Rhizophydium sp. JEL317 isolate AFTOL-ID 35 (AY635821.1)
Kappamyces laurelensis isolate AFTOL-ID 690 (DQ536478.1)
PA2009E12
PA2009D2
PA2009A4
PA2009C4
PA2009E10
Uncultured Chytridiomycota clone PFB4SP2005 (EU162636.1)
PA2009E11
Uncultured Chytridiomycota clone PFD5SP2005 (EU162640.1)

Fungi

Basidiomycota
Ascomycota
Chrytridiales
Rhizophlyctidales
Kappamyces
Chytridiomycota
Rhizophydiales

PA2009E6
Uncultured choanoflagellate clone PFH2AU2004 (DQ244006)
Amoebidium parasiticum strain ATCC 32708 (Y19155.1)
PA2009E13
Uncultured fungus clone PFF3AU2004 (DQ244010)
Uncultured eukaryote clone RT5iin3 (AY082996.1)
Uncultured fungus clone PFB3AU2004 (DQ244018)

Choanozoan
Ichthyosporea

Telonema antarcticum (AJ564773.1)
Uncultured Telonema clone IND31.100 (AM418563.1)
PA2009A5
Uncultured eukaryote clone PTG4SP2005 (EU162631.1)

Telonemida

Uncultured alveolate clone PAD1SP2005 (EU162628.1)
PA2009C9
Uncultured alveolate clone PAC8SP2005 (EU162630.1)
Uncultured alveolate clone PAB5AU2004 (DQ244038)
PA2009A8
Piscinoodinium sp. RWL-2007 clone 626 (EF016919.1)
Gymnodinium beii from Orbulina universa (U37406.1)
PA2009D7
Pfiesteria-like sp. CCMP1835 (AY590477.1)
PA2009E19
Uncultured choreotrichid ciliate clone CH1_2A_10 (AY821916.1)
Rimostrombidium lacustris (DQ986131.1)
Halteria grandinella (AF508759.1)

Alveolata

Perkinsea
Dinophyceae
Ciliphora

PA2009C8
Pythium insidiosum (AY486144.1)
Cercomonas edax (AF411269.1)

Stramenopile
Cercozoan

PA2009E14
Uncultured freshwater eukaryote clone LG08-04 (AY919706.1)
Uncultured eukaryotic picoplankton clone P34.10 (AY642702.1)
PA2009C5
Katablepharis japonica (AB231617.1)
PA2009D3
Uncultured eukaryotic picoplankton clone P1.31 (AY642716.1)
Uncultured cryptophyte clone PFF1AU2004 (DQ244012)
PA2009D4
Plagioselmis nannoplanctica strain N750301 (FM876311.1)
PA2009D5
Cryptomonas ovata (AF508270.1)

Cryptophyta

Katablepharidaceae

Treubaria schmidlei (U73474.1)
PA2009E15
Treubaria setigera (U73475.1)
PA2009A6
Chlamydomonas acidophila strain CCAP 11/137 (AJ852427.1)
PA2009E16
PA2009B2
PA2009C6
PA2009E17
Chlamydomonas sordida (AB290341.1)
Chlamydomonas hedleyi isolate ATCC 50216 (AJ781312.1)
Chlamydomonas noctigama isolate SAG 33.72 (AJ781311.1)
PA2009D6
Monoraphidium contortum strain AS6-3 (AY846382.1)
PA2009A7
Uncultured eukaryote clone PG8AU2004 (FJ799981.1)
Sphaeroplea soleirolii (U73473.1)
PA2009E18
Mychonastes sp. 5C5 (AF357153.1)
Ankistrodesmus fusiformis strain SAG 2005 (X97352.1)
PA2009C7
Micractinium pusillum strain CCAP 248/3 (FM205875.1)
Didymogenes anomala strain SAG 18.91 (FM205839.1)
PA2009B3
Choricystis sp. Pic8/18P-11w (AY197629.1)

Viridiplantae

Treubariaceae
Chlamydomonadales
Sphaeropleales
Mychonaste
Trebouxiophyceae

0. 1 substitutions/site

Fig. 1. Phylogenetic tree based on 18S rRNA gene sequences obtained from the Lake Pavin. The unrooted phylogenetic tree displayed 46 operational taxonomic units (OTUs) inferred from 146 18S rRNA gene sequences sharing less than 99% sequence identity. Representative from each OTU were submitted to a BLAST search (Altschul *et al.*, 1990) to determine the closest homologous 18S rRNA gene sequences from the non-redundant nucleotide database (NCBI). The OTUs and reference sequences were aligned using CLUSTALW (Thompson *et al.*, 1994), and the construction of the phylogenetic tree was performed using the Bayesian method from MrBayes3 software. The numbers at the nodes correspond to Bayesian posterior probabilities. The boxes (dash lines) showed the different phyla, where each taxonomical subdivisions are delimitated by vertical lines.

With the pyrosequencing approach, an important diversity of fungi was observed in the two investigated environments, but also of many other eukaryotes with the two sets of primer used (eukaryote and fungi). When considering the data obtained with the eukaryote primers, the most represented groups were *Stramenopile* (representing 27% and 3% of the pyrosequence reads in Pavin and Aydat respectively), *Chlorophyta* (19% and 11%), *Metazoa* (15% and 53%), *Fungi* (12% and 9%) and *Alveolata* (10% and 20%) (Fig. 4). The *Cercozoa* and *Telonemida* were also among the major groups in Lake Pavin, accounting for 7% and 8% of the sequences respectively (Figs 4 and 5). When considering the data obtained with the fungus-designed primers, enrichment in fungal sequences (15% and 19% of the pyrosequence reads in Pavin and Aydat respectively) was observed. With these fungus-designed primers, only a few metazoan and *Viridiplantae* sequences were retrieved, while the *Katablepharidophyta* (60% and 10% of the pyrosequence reads in Pavin and Aydat respectively) and *Cryptophyta* (10% and 36%) were among the dominant groups (Fig. 4). Finally, whatever the primer considered, it was not possible to assign a taxonomic group for 1–5% of the sequences from the different samples because of suspicious or ambiguous assignment of the SSU rDNA gene fragments in the database, while less than 1% of the sequences had no hit in the Silva SSU database.

The pyrosequencing method was used to identify dominant taxa and estimate their spatial distribution using data obtained from eukaryote primers for the overall biodiversity, and from fungal primers for the diversity of fungi. In the following sections, we describe the taxonomical groups and species unveiled by the complementary 'classical' and pyrosequencing approaches and we compare the diversity observed in the two pilot lakes, with focus on fungi.

The diversity of fungi

Chytridiomycota. In both lakes, the identified chytrids belonged to the *Chytridiales*, *Rhizophydiales* and *Rhizophlyctidales* orders. The chytrid sequences found here confirmed the high diversity of *Chytridiomycota*-affiliated sequences reported in previous studies from Lake Pavin (Lefèvre *et al.*, 2007; 2008) (Figs 1 and 2).

Members of *Rhizophydiales*, which occur in aquatic ecosystems primary as phytoplankton parasites (Letcher

et al., 2006; James *et al.*, 2006b), were the most represented fungi in Lake Pavin with seven OTUs (Fig. 1). Only one *Rhizophydiales* taxon was common to both lakes, as suggested by the 99.58% identity between the OTUs PA2009E11 (Pavin) and AY2009D2 (Aydat) sequences (Figs 1 and 2). These SSU rDNA gene sequences were closely related to PFD5SP2005 sequence previously identified in Lake Pavin and considered as being part of a new clade (Lefèvre *et al.*, 2008). These sequences were grouped with *Kappamyces laurelensis* in the *Kappamyces* subclade (Letcher and Powell, 2005). In Lake Pavin, six OTUs were clustered within the *Kappamyces* group (Fig. 1). Interestingly, four of them (PA2009C4, PA2009A4, PA2009E12 and PA2009D2) formed a new sister clade of the main *Kappamyces* group (Fig. 1) which was not found in the pool of sequences obtained from the Lake Aydat.

In both lakes, one OTU (PA2009C1 in Lake Pavin and AY2009B4 in Lake Aydat) belonged to the *Chytridium angularis* clade (James *et al.*, 2006b). This clade includes *Chytridium polysiphoniae*, a parasite of the marine brown alga *Pylaiella littoralis*, and *C. angularis* that grow on pollen and heat treated *Oedogonium* (a green algae) baits (Longcore, 1992). Interestingly, in Lake Pavin, three new sequences clustered together to form a sister clade of the *C. angularis* clade (Fig. 1). In Lake Aydat, AY2009C3 and AY2009B3 sequences form an early divergent novel clade within the *Chytridiomycota* (Fig. 2).

Concerning the order *Rhizophlyctidales*, one OTU (PA2009E8 in Lake Pavin and AY2009D3 in Lake Aydat) belonged to the *Rhizophlyctis* clade described by James and colleagues (2006b). Indeed, in this latter study, *Catenomyces persinius*, a cellulose decomposer found in aquatic habitats, currently included in the order *Blastocladiales* (phylum *Blastocladiomycota*), was assigned to the *Rhizophlyctis* clade (including *Rhizophlyctis rosea*), a sister clade to the *Spizellomycetales*. Based on molecular monophyly and zoospore ultrastructure, this *Rhizophlyctis* clade was recently designated as a new order, the *Rhizophlyctidales* (Letcher *et al.*, 2006). To note, in each lake, three OTUs constituted a novel clade, which appear to be the sister group of the *Rhizophlyctis* clade (Figs 1 and 2).

The proportion of each fungal OTU (classical approach) was quantified according to the number of matching pyrosequence reads (Table S1a and b). The two approaches consistently displayed the same dominant
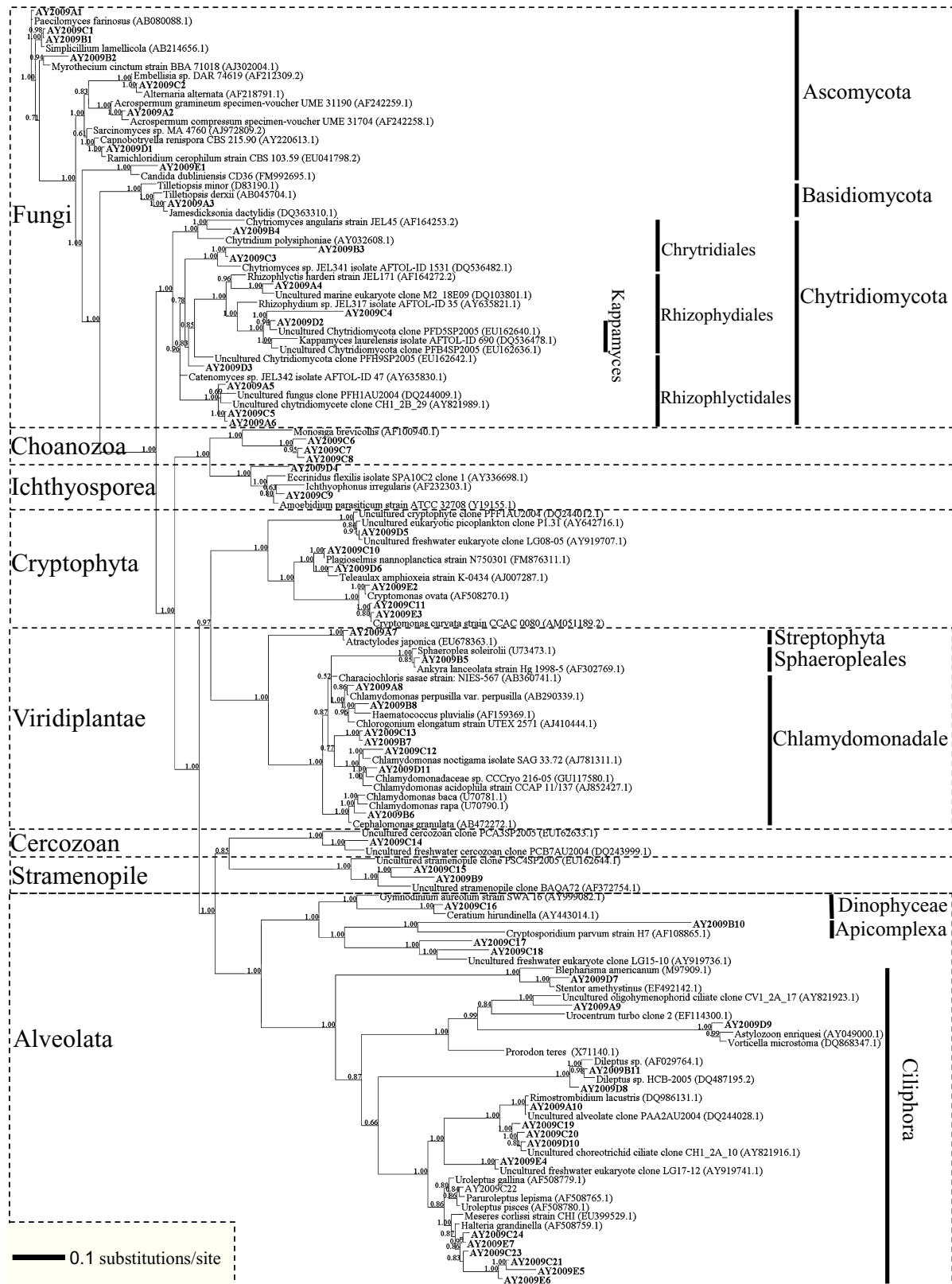
Fig. 2. Phylogenetic tree based on 18S rRNA gene sequences obtained from the Lake Aydat. The unrooted phylogenetic tree, similarly to Fig. 1, displayed 63 OTUs inferred from 143 rDNA gene sequences.
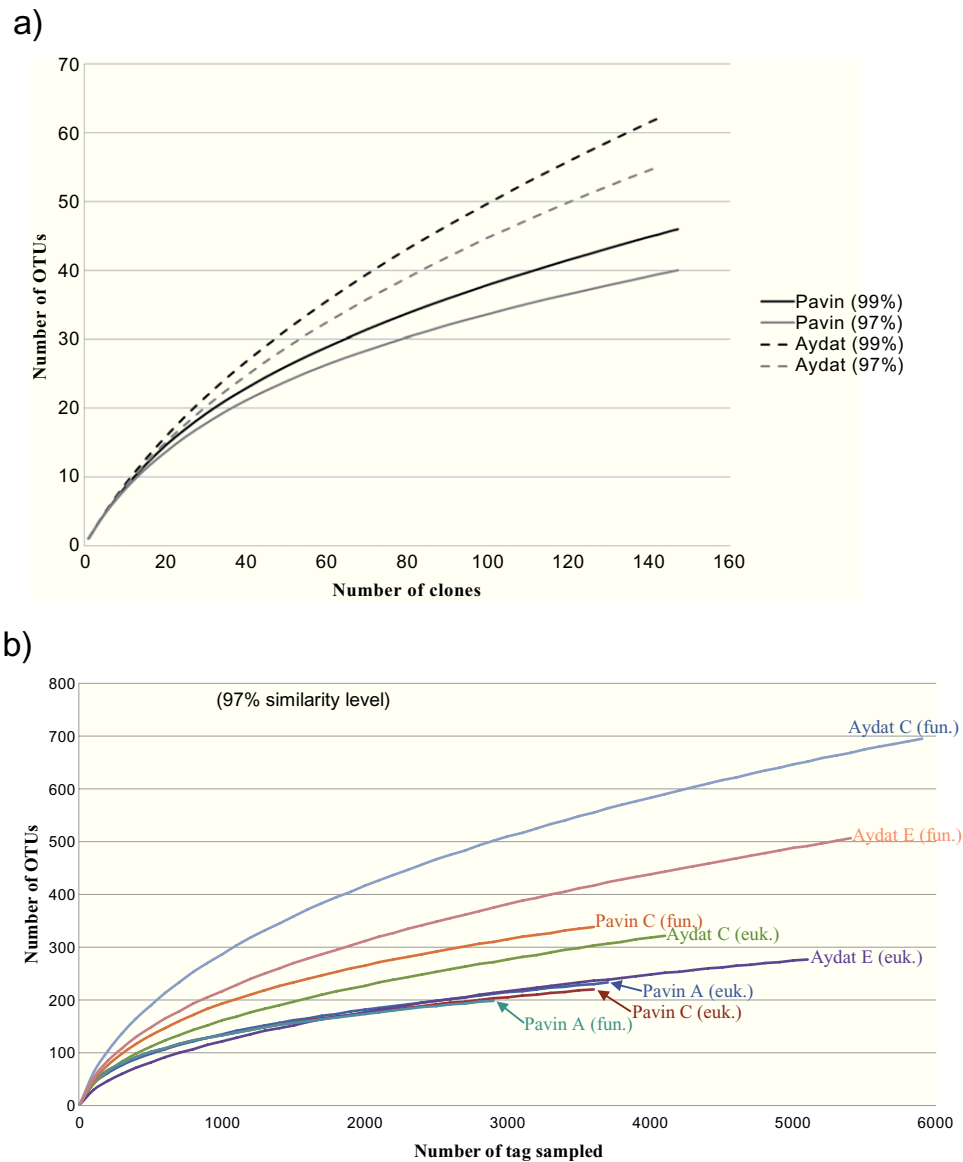
a)

b)

Fig. 3. Rarefaction curves. Rarefaction curves representing the numbers of OTUs versus the number of clones were plotted from sequences obtained with the classical (A) and the pyrosequencing (B) approaches. The curves were calculated using the program Mothur (Schloss *et al.*, 2009), with a cut-off value set to 0.01 and 0.03 (OTUs with differences that do not exceed 1% and 3%) for the analysis.
A. For the classical approach, all the 18S rRNA gene sequences obtained from Lake Pavin (full lines) and Lake Aydat (dash lines) were grouped into OTUs at a level of sequence similarity ≥ 99% (black lines) and ≥ 97% (grey lines).
B. The pyrosequence reads obtained from lakes, sampling stations, and using the two sets of primers [eukaryote (euk.) and fungus (fun.)], were curated (see *Experimental procedures*) and grouped into OTUs at a level of sequence similarity ≥ 97%.

species. In Lake Pavin, pyrosequence data obtained from both primer sets (eukaryotes and fungi) consistently displayed PA2009E8, PA2009E9, PA2009E10 and PA2009E11 as the dominant chytrids' OTUs (Table S1a). In Lake Aydat, the OTU AY2009C4 was consistently found highly dominant among the pyrosequence reads obtained from both sets of primers. This OTU represents 80% of the pyrosequence reads matching a fungal OTU identified from the classical approach (Table S1b), and corresponds

to a novel species having only 90.3% similarity with *Rhizophydium* sp. JEL317, its closest relative.

When combining pyrosequence data obtained with the two set of primers, additional sequences were retrieved compared with the sole use of the classical approach. Some *Chytriomyces* sp. JEL341 and representatives of the *Monoblepharidaceae* family were found in Lake Pavin, while several *Rhizophydiales* sp. AF033 sequences were retrieved in Lake Aydat (Figs 5 and 6, and Table S2).
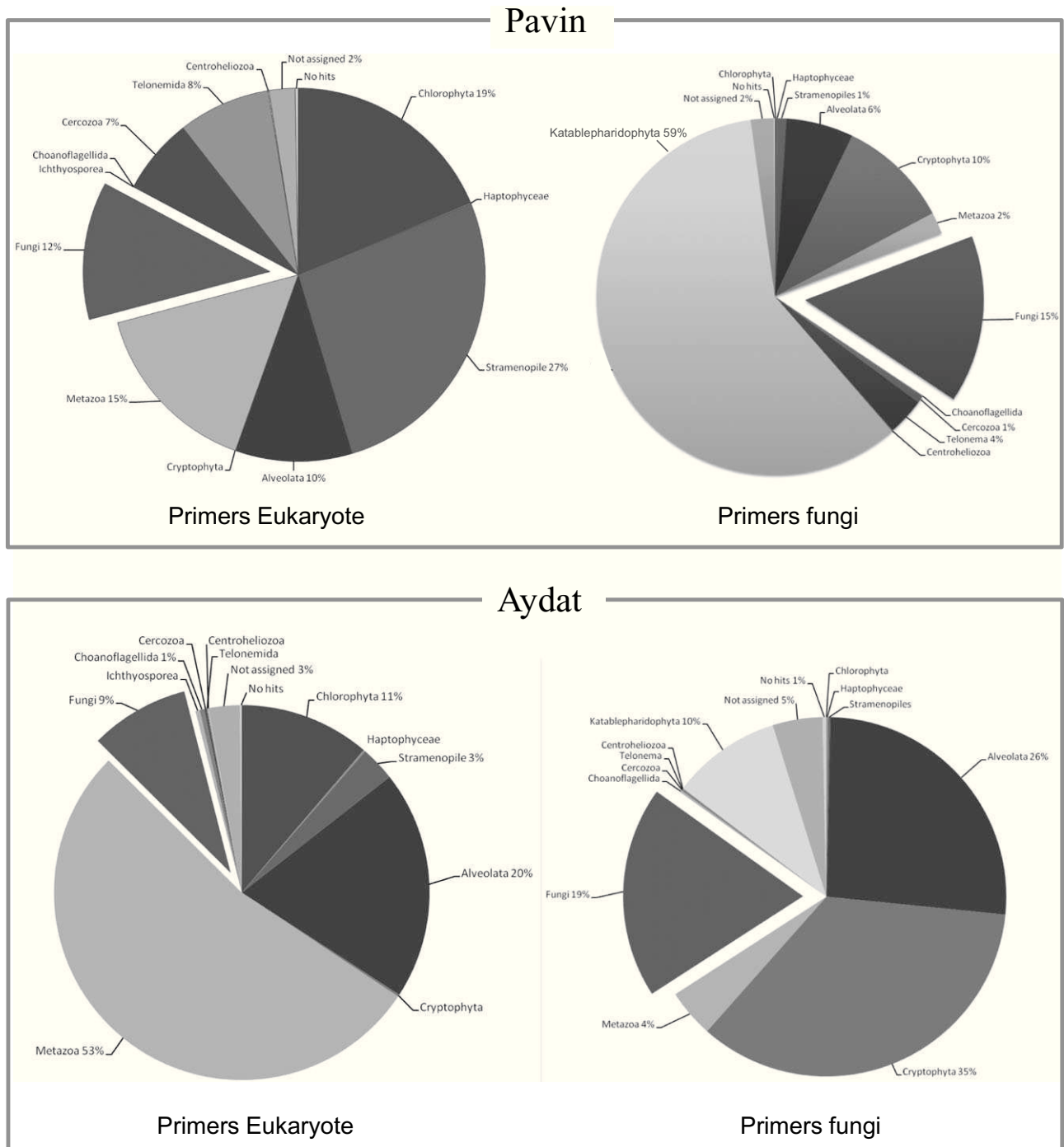
Fig. 4. Proportion of taxonomic groups identified in Lake Pavin and Lake Aydat using pyrosequencing of 18S rRNA gene hypervariable regions. The reads obtained from pyrosequencing of 18S rRNA hypervariable region were subject to BLASTN (Altschul *et al.*, 1990) search against the Silva SSU rRNA database (http://www.arb-silva.de/) to assign a taxonomic group. The pie diagram displayed the proportion of reads, obtained from both lakes sampling stations, and using the two primers sets (eukaryote and fungus), belonging to a particular phylum. 'No hit' corresponds to reads having no homologous sequence in the Silva database (threshold $E = 10^{-5}$). 'Not assigned' corresponds to reads having a match in the Silva database but without a precise taxonomic phylum assignment.
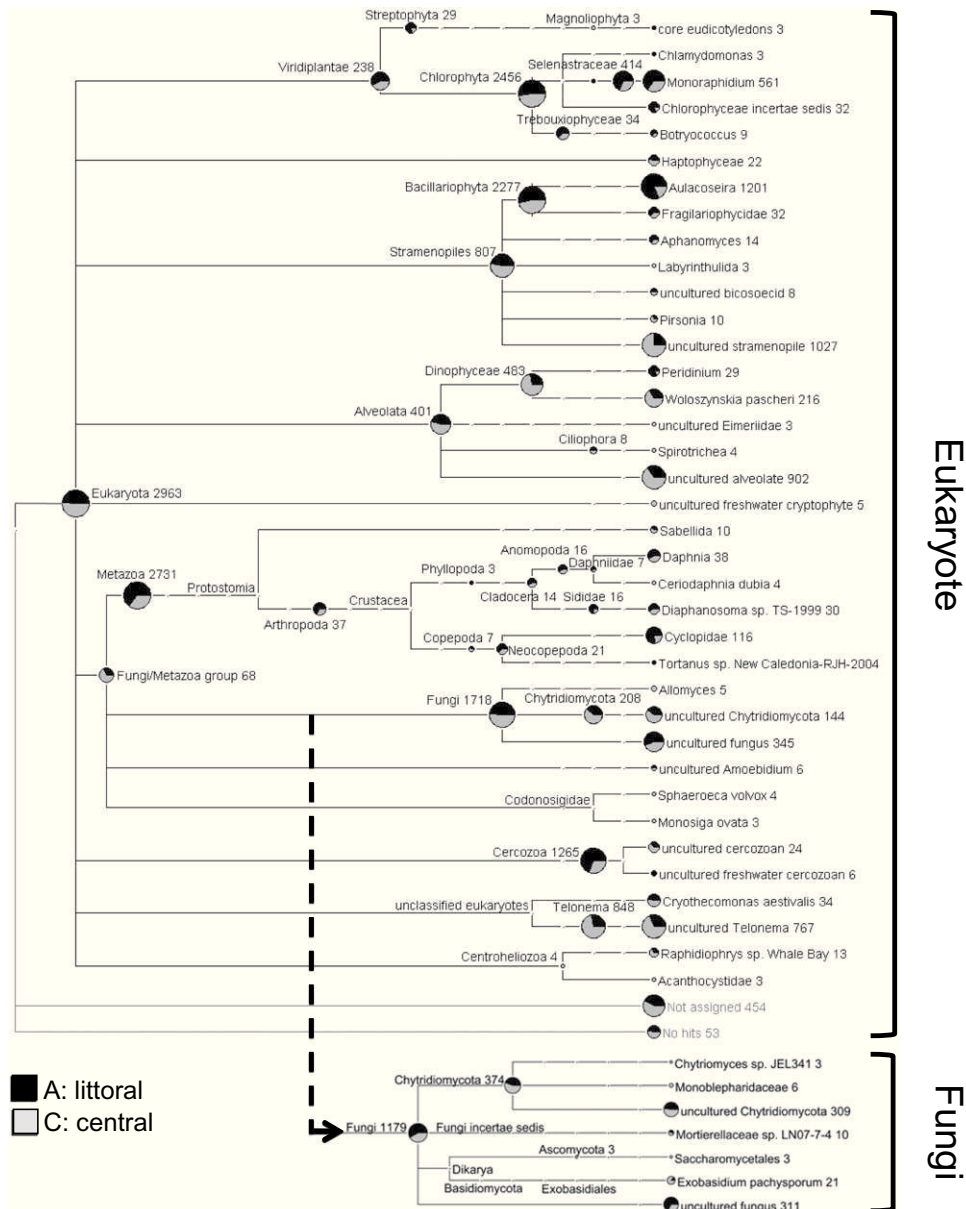
Fig. 5. Taxonomic assignment of the pyrosequences 18S rRNA reads obtained from Lake Pavin. The 18S rRNA hypervariable tag pyrosequences obtained from Lake Pavin were analysed using the software MEGAN, after BLASTN (Altschul *et al.*, 1990) search against the Silva SSU rRNA database. The MEGAN software (Huson *et al.*, 2007) plots on a schematic phylogenetic tree the number of pyrosequence reads matching a particular taxonomical group. Each taxonomic node is represented by a pie diagram, with littoral sample in black and central samples in grey colour, whose size is proportional to the number of assigned reads (given by the numbers). The tree displays all taxonomic groups identified from the assignment of pyrosequence reads obtained with eukaryote primers, while only the fungal Kingdom is given (on the bottom) for reads obtained using fungus-designed primers.

Other fungi. One *Ascomycota* taxon was common to both lakes and identified as OTUs PA2009D1 (Pavin) and AY2009B1 (Aydat). They indeed shared 99.4% identity, and were similar to *Simplicillium lamellicola*, a pathogen of another fungus, *Agaricus bisporus* (*Basidiomycota*) (Spatafora *et al.*, 2007). This pathogen was the dominant *Ascomycota* in both lakes (Table S1a and b). The other *Ascomycota* identified in the two lakes represented minor

species according to the low number of corresponding pyrosequence reads. In addition, pyrosequencing yielded *Saccharomycetales* as the sole group of *Ascomycota* in both lakes (Figs 5 and 6). Only one sequence per lake, retrieved from the SSU rDNA clone libraries, clustered within the *Basidiomycota* group, i.e. *Exobasidium rhododendri* in Lake Pavin and *Jamesdicksonia dactylidis* in Lake Aydat. Both of them are putative plant pathogens.

Fig. 6. Taxonomic assignment of the pyrosequences 18S rRNA reads obtained from Lake Aydat. Similarly to Fig. 5, the pyrosequence 18S rRNA assigned reads were displayed on a schematic phylogenetic tree. Pie diagrams, at each taxonomic nod, indicate the proportion of reads obtained from the central (grey colour) and littoral (black colour) sampling stations of Lake Aydat.

Compared with the classical SSU rDNA approach, pyrosequencing revealed the presence of *Exobasidium pachysporum* in both lakes and of three sequences affiliated to *Tilletiopsis minor* in Lake Aydat. In both lakes, the *Basidiomycota* sequences retrieved by the two approaches belonged to the class of *Exobasidiomycetidae*. The other fungus sequences, only identified through pyrosequencing, included some *Glomeromycota* (previously known as *Zygomycota*), such as *Mortierellaceae* sp. LN07-7-4 in the two lakes; and *Glomus mosseae* and representatives of the *Entomophthoraceae* family in Lake Aydat (Figs 5 and 6).

The diversity of other eukaryotes

Alveolata. In clone libraries from Lake Pavin, two *Alveolata* OTUs belonged to the *Dinophyceae*, one to the *Perkinsea* and one to the *Ciliophora* group (Fig. 1). In Lake Aydat, one-third of the identified OTUs from the clone libraries corresponded to *Alveolata*, from which 17 belong to the *Ciliophora*, one to the dinoflagellates and one to the *Apicomplexa* group (Fig. 2). One taxon of *Ciliophora* was common to both lakes, i.e. the OTUs PA2009E19 (Pavin) and AY2009D10 (Aydat) sharing 99.7% sequence identity. On their respective phylogenetic trees, these two OTUs

form new clades, supported by Bayesian posterior probabilities of 1.00 and 0.81 respectively. These clades clustered two additional OTUs in Lake Aydat (AY2009C19, AY2009C20), with *Rimostrombidium lacustris* as a sister group. In Lake Aydat, another new clade of *Ciliophora* clustering three OTUs (AY2009E5, AY2009C21 and AY2009E6) was supported by Bayesian posterior probabilities of 0.83. The *Apicomplexa* group was only found in Lake Aydat with one OTU (AY2009B10) having *Cryptosporidium parvum* as closest relative, i.e. a human protozoan parasite that can cause an acute short-term infection and become severe for immunocompromised individuals. Considering the phylogenetic distance between AY2009B10 and *C. parvum*, it was unlikely that Lake Aydat contained this parasite. Therefore, this sequence could represent a new clade or a *Cryptosporidium*-like organism. Consistent with the observations made from the classical approach, pyrosequence data displayed a larger diversity of *Alveolata* in Lake Aydat, compared with Lake Pavin (Figs 5 and 6), which contributed to the apparent overall higher eukaryote diversity in Lake Aydat.

Cryptophyta. Two *Cryptophyta* taxon, identified by the classical approach, were found in both lakes. The first one, corresponding to OTUs PA2009D3 (Pavin) and AY2009D5 (Aydat) sharing 99.6% sequence identity, clustered together with a sequence isolated from Lake Georges (LG08-05) (Richards *et al.*, 2005), and a sequence (P1.31) previously obtained from Lake Pavin (Lefranc *et al.*, 2005). The second, corresponding to OTUs PA2009D5 (Pavin) and AY2009E2 (Aydat) sharing 99.4% sequence identity, was closely related to *Cryptomonas ovata*. In both lakes, two distinct OTUs (having less than 99% identity) were related to *Plagioselmis nannoplanctica* strain N750301. In Lake Pavin, two OTUs (PA2009C5 and PA2009E14) clustered together in a clade with sequence P34-10 previously isolated from Lake Pavin (Lefranc *et al.*, 2005) and sequence LG08-04 from Lake Georges (Richards *et al.*, 2005). This clade, named LG-E in the phylogenetic tree based on sequences obtained from the Lake Georges (Richards *et al.*, 2005), showed branches that were consistently attracted either to the cryptomonads or to the glaucocystophytes. In fact, this branch derived from the *Katablepharis japonica* group for which SSU rDNA gene sequences were not available at that time, and belongs to the katablepharids, a distant sister group from the *Cryptophyta* (Okamoto and Inouye, 2005). These two OTUs (PA2009C5 and PA2009E14) and their corresponding katablepharids group were matching the highest number of pyrosequence reads obtained when using fungus-designed primers, but were poorly detected when using eukaryote primers (Fig. 4 and Table S1a). Similarly in

Lake Aydat, the OTUs AY2009E3 and AY2009C11, related to *Cryptomonas curvata* strain CCAC 0080, and their corresponding *Cryptophyta* group, were displaying the highest number of match with pyrosequence reads obtained from fungus-designed primers, but were also poorly detected using eukaryote primers (Table S1b).

Viridiplantae. This group of organisms was well represented in the clone libraries, accounting for 37% and 21% of the total numbers of sequences in Lakes Pavin and Aydat respectively. Representatives of the *Sphaeropleales* and *Chlamydomonadales* were found in both lakes, with the addition of the *Treubariaceae, Trebouxiophyceae* and *Mychonaste* groups in Lake Pavin (Figs 1 and 2). Two distinct new clades within the *Chlamydomonadale* group were identified in the two lakes, including the OTUs PA2009E16 and PA2009B2 in Lake Pavin (Fig. 3) and AY2009C13 and AY2009B7 in Lake Aydat (Fig. 4). Reads corresponding to *Viridiplantae* 18S rRNA genes were only retrieved from pyrosequencing when using eukaryote primers.

Minor groups. Stramenopiles were poorly identified in both lakes' clone libraries (Figs 1 and 2), but abundantly in Lake Pavin according to the pyrosequence data (Fig. 4). Through the classical approach, the *Telonemida* group was only identified in Lake Pavin, with the OTU PA2009A5. The pyrosequencing approach confirmed the abundance of *Telonemida* in Lake Pavin, and their rarity in Lake Aydat. On the opposite, *Choanozoa* sequences were only retrieved in Lake Aydat, both with the classical approach (i.e. OTUs, AY2009C6, AY2009C7 and AY2009C8, clustering together with *Monosiga ovata* and more distantly with *Monosiga brevicollis*) (Fig. 2) and with the pyrosequencing approach (Fig. 6). Finally, in addition to taxa identified with the classical approach, pyrosequencing using eukaryote primers allowed the retrieval of SSU rDNA sequences from numerous *Arthropoda* ($n = 312$ in Lake Pavin and 2502 in Lake Aydat), *Rhizaria* ($n = 1329$ and 59), a few *Haptophyceae* ($n = 22$ and 45) and *Centroheliozoa* ($n = 20$ and 9) (Figs 5 and 6). Overall, the pyrosequencing of 18S rRNA hypervariable regions confirmed the higher eukaryote diversity in Lake Aydat, compared with Lake Pavin (Figs 5 and 6).

Within-lake spatial variability

Lake Pavin. Both the classical and the pyrosequencing approaches showed a relatively homogenous spatial distribution of eukaryotes in Lake Pavin (Fig. 5 and Table S1a). The classical approach yielded 23 OTUs in the centre of the lake and 15, 18, 18 and 19 OTUs for stations A, B, D and E respectively (Fig. 1, Table S1a). The pyrosequence data, independently of the set of

primer used (eukaryote or fungi), displayed a similar number of reads corresponding to the main groups of taxa retrieved at the two sampling stations (Fig. 5). The number of pyrosequence reads corresponding to fungi decreased from the shore station A to the central station C, with a consistent number of reads related to uncultured fungi (i.e. having not clear taxonomical assignment) on the littoral zone (Fig. 5 and Table S2).

Chytridiomycota was the most identified group of fungi among the pyrosequence reads obtained using fungus-designed primers. Chytrid sequences increased from 25.8% ($n = 313$) of all the fungal reads in the riparian station A to 37.7% ($n = 379$) in the centre station C (Table S2). A similar result was observed when using eukaryote primers, but with about two times less chytrid reads retrieved from pyrosequencing. The dominant chytrids OTUs were equally found located on the littoral and central areas of the lake (Table S1a). The *Ascomycota* were consistently more frequently retrieved from the shore station by both approaches (classical and pyrosequencing) (Fig. 5 and Table S1a). The dominant *Ascomycota* (OTU PA2009D1) was almost exclusively located in the shore station (Table S1a). Finally, the *Basidiomycota* OTU (PA2009E1) retrieved with the classical approach was mostly found at the littoral station of Lake Pavin (Fig. 1, Table S1a). This contrasts with the general within-lake distribution of the other *Basidiomycota*, which were mostly retrieved in the central ($n = 18$) compared with the littoral ($n = 3$) stations of the lake according to pyrosequence data (Fig. 5 and Table S2).

The spatial patterns for other eukaryote species could also be estimated using data from pyrosequencing. In this respect, most of the eukaryotic groups displayed a homogeneous localization between the riparian and central sampling stations. Some differences were observed for *Cercozoa* which were mostly sampled in the littoral waters while, in contrast, *Telonemida* were preferentially sampled in the pelagic central waters of Lake Pavin. Finally, the *Cryptophyta* group, only unveiled when using fungus-designed primers, was threefold more abundant in the centre ($n = 1135$) compared with the littoral ($n = 347$) areas of the lake, according to the pyrosequencing (Table S2).

Lake Aydat. Both methods (classical and pyrosequencing) showed a more heterogeneous spatial distribution of species in Lake Aydat (Fig. 6 and Table S1b), compared with Lake Pavin. The classical approach yielded 24 OTUs in the central area, and 19, 19, 22 and 20 OTUs at stations A, B, D and E respectively (Fig. 2 and Table S1b). Among the eukaryotic groups, fungi displayed a decreasing proportion from station A (45%), B (36%), C (22%), D (25%) to station E (12%) relative to the total number of sequences retrieved by the classical approach. Fungi thus displayed a clear spatial pattern from station E to A, following the up–downstream flow of the river. With the pyrosequencing approach and independently of the set of primers used, fungi represented around 15% of all the pyrosequence reads at the central station C, but only 1–3% at the littoral station E (Fig. 6 and Table S2).

Chytrids spatial distribution, estimated from pyrosequencing and confirmed by the use of both primer sets (eukaryote and fungi), showed a preferential localization in the centre pelagial of Lake Aydat (Table S2). The classical approach showed that the OTU AY2009C4, representative of a novel clade within the *Rhizophydiales* order, was particularly abundant in the centre of the lake. When searching corresponding reads for AY2009C4 in the pyrosequence database obtained with fungus-designed primers, it appeared that matching reads were 33 times more abundant in samples taken from the central station C ($n = 701$) compared with the shore station E ($n = 21$) (Table S1b). This was confirmed by the pyrosequence data obtained from eukaryote primers for the OTU AY2009C4 which was 40 times more abundant in the central than in littoral stations. Two others OTUs, AY2009B3 and AY2009C3, clustering together with *Chytriomyces* sp. JEL341, were two- to threefold more represented in the central than in the littoral areas of the lake, according to pyrosequence data and independently of the set of the primers used (Table S1b).

The classical approach showed that the number of OTUs and of sequences composing OTUs belonging to *Ascomycota* decreased from the littoral sampling stations A and B (located close to the within-lake islands) to the rest of the lake (Table S1b). This gradient was confirmed by the pyrosequence data, with an overall higher number of *Saccharomycetales* at the station C ($n = 115$) compared with station E ($n = 70$) (Fig. 6 and Table S2). A low number of *Ascomycota* OTUs, identified by the classical method (Fig. 4), were matching pyrosequence reads at station E (Table S1b), thus reinforcing our hypothesis of a decreasing gradient from stations A to E for *Ascomycota* in Lake Aydat. *Basidiomycota* sequences were mainly retrieved in the riparian areas of this lake, by the two approaches. The AY2009A3 OTU, closely related to *Jamesdicksonia dactylidis*, was only retrieved by the classical method at station A (Table S1b). The other *Basidiomycota*, *E. pachysporum* ($n = 11$) and *T. minor* ($n = 3$), were exclusively retrieved by pyrosequencing at the sampling station E. These results are in contrast with those obtained in Lake Pavin, where sequences corresponding to *Basidiomycota* were mainly found in the centre of the lake. Other groups of fungi were identified solely by pyrosequencing, among which *Glomeromycota* were exclusively located in the littoral area ($n = 6$), while *Mortierellaceae* sp. LN07-7-4 ($n = 5$) was exclusively found in the centre of the lake (Fig. 6, Table S2).

Besides fungi, the spatial pattern for other eukaryotes was mainly estimated from the pyrosequence data obtained with eukaryote primers. The major groups of species displaying spatial distribution included *Viridiplantae* and *Alveolata* which were preferentially localized in the centre of the lake, while *Stramenopile* and *Arthropoda* were mainly located in the riparian area. The minor taxonomic groups, *Choanoflagellida*, *Centroheliozoa* and *Telonemida*, were mostly or exclusively found in the centre of Lake Aydat (Fig. 6, Tables S1b and S2). Finally, the kathablepharids group, only detected with the fungus-designed primers, was almost exclusively found in the centre ($n = 1559$) compared with the littoral ($n = 47$) areas of the lake (Table S2). The heterogeneous spatial distribution of the kathablepharid group in Lake Aydat contrasts with its homogenous distribution in Lake Pavin. In contrast to the observation made in Lake Pavin, *Cryptophyta* were homogeneously located in the central ($n = 3476$) and in the peripheric ($n = 2219$) areas of Lake Aydat. However, two *Cryptophyta* OTUs (AY2009C11 and AY2009E3) clustering with *C. curvata* strain CCAC 0080 were twofold more represented in the littoral than in the central areas of Lake Aydat (Table S1b).

Discussion

General and methodological considerations

This study's aim was to identify the diversity of microbial eukaryotes within two freshwater environments, Lakes Pavin and Aydat, Massif Central, France, which contrasted in terms of the influences of allochthonous terrestrial inputs. These two lakes indeed differ by their surface, altitude, depth, trophic status, geological origin, water flux (river, drainage basin) and anthropogenic activities (agricultural and industrial impact). Considering all these differences, a homogenous distribution of species was expected in Lake Pavin where typical pelagic organisms are characteristics of the whole water body. Conversely, we anticipated Lake Aydat would reveal a heterogeneous spatial variation of species, due to its exposition to enhanced allochthonous materials from terrestrial inputs. Few studies have investigated the diversity of picoeukaryotes in these two lakes (Lefranc *et al.*, 2005; Lefèvre *et al.*, 2007; 2008), and highlighted the importance and ecological significance of chytrids (Lefèvre *et al.*, 2008). Our work intended to complete these studies by identifying new fungi and describe their quantitative importance. This goal was achieved through two complementary methods to beneficiate from their respective advantages: a qualitative taxonomic identification of each species, and of new clades, by the 'classical cloning/sequencing approach', and the corresponding dominant and rare species by the pyrosequencing of the

18S rRNA gene hypervariable regions. In addition, the large number of reads obtained with the pyrosequencing was a powerful tool to evaluate the distribution of species between the littoral areas versus the central pelagial ones. While one set of primers for fungi was used for the classical approach (White *et al.*, 1990), two sets were utilized for pyrosequencing, i.e. a universal eukaryote set of primers (Casamayor *et al.*, 2002; Lopez-Garcia *et al.*, 2003; Lepere *et al.*, 2006), and a set of fungus-designed primers (Borneman and Hartin, 2000). Although the latter primers were not specific for the 18S rDNA genes of the sole fungi, they allowed a significant fungal enrichment of our database (compared with eukaryote primers), up to 2.6-fold for total fungi and 3.5-fold for chytrids. The fungal primers did not amplify 18S rDNA genes from the abundant groups of *Arthropoda* and *Viridiplantae*, leading to the enrichment in 18S rDNA gene sequences from other eukaryotes, primarily from fungi, i.e. compared with the eukaryote primers. The concurrent use of fungal primers was therefore justified. In addition, the fungus-designed primers allowed the identification of an increased number of fungi taxa compared with universal eukaryote primers (Figs 5 and 6). Nevertheless, designing primers specific for pelagic fungus rRNA genes is one of the major bottlenecks that need to be resolved for future studies. In this regard, the ITS1, 5.8S and ITS2 sequences that were generated with the classical approach used in this study will be very helpful to design new tools to follow the dynamics of particular fungal species in natural waters.

The diversity in Lake Aydat was greater than that observed in Lake Pavin, which is likely to be linked to the differences in the trophic status of theses lakes. It is interesting to observe that the two lakes shared only fives taxa in common (two *Cryptophyta*, one *Alveolata*, one *Ascomycota* and one *Chytridiomycota*), highlighting the difference between the two lakes and providing evidence that aspects of eukaryote microbial diversity are specific to certain aquatic environments. The rarefaction curves constructed from the 146 sequences in Lake Pavin and 142 sequences in Lake Aydat displayed a linear pattern, indicating a still undersampling of the total diversity. Consistently with this result, the same groups of organisms were identified, and confirmed the existence of new clades known from previous studies (Lefèvre *et al.*, 2007; 2008), in addition to the unveiling of many new rDNA gene sequences that were not previously retrieved. Finally, this low-throughput approach allows the retrieval of SSU rDNA sequences from species that were probably the dominant ones in the lakes. In contrast, the high-throughput pyrosequencing of 18S rRNA gene hypervariable regions was initiated in order to increase the sampling effort of rDNA gene sequences in the target lakes and to unveil a much extended eukary-

ote diversity, with a focus on fungi. Rarefaction curves computed from pyrosequence data displayed linear pattern if applying 99% sequence identity threshold, while it is approaching a plateau if applying 97% identity threshold (Fig. 3B). A 97% identity threshold is better adapted to computer rarefaction curves from deep pyrosequencing of rRNA reads (Kunin *et al.*, 2010). These patterns of rarefaction curves suggest a good sampling of the water samples taken from the lakes and showed higher apparent species richness when using fungus-designed primers compared with eukaryote primers. The variability within different regions of the SSU rRNA molecule and the length of the amplicon have a great effect on the apparent species richness, as it was shown when tracking the microbial diversity of the termite hindgut (Engelbrektson *et al.*, 2010). Species evenness and richness should not be directly compared between different regions of the rRNA molecules (Engelbrektson *et al.*, 2010). In addition, richness should also be temperate by some technological limits (alignments of large sets of sequences, pyrosequencing errors, statistical tools etc.) (Kunin *et al.*, 2010) and by the unknown variability of the multiple copy of rDNA operon from the same organism, as it was demonstrated for bacteria belonging to the genus *Vibrio* (Moreno *et al.*, 2002).

Pyrosequencing of rRNA gene hypervariable region reveal much greater eukaryote diversity in the studied lakes than in the previous studies (Lefranc *et al.*, 2005; Lefèvre *et al.*, 2007; 2008). The reason is that only the classical approach was used in these studies, unveiling dominant but not minor species. Rare species, corresponding to species that are maintained at a low number [arbitrarily chosen below 0.05–1% of the total number of pyrosequence reads (Sogin *et al.*, 2006; Huse *et al.*, 2008)], may play an important role in lakes by becoming dominant in response to environmental changes, in addition to represent a nearly inexhaustible source of genomic innovation (Sogin *et al.*, 2006). They indeed represent a hypothetically important reservoir of ecological redundancies that can buffer the effects of dramatic environmental shifts and, perhaps, insure the maintenance of basic biogeochemical processes in natural ecosystems (Sime-Ngando and Niquil, 2011). However, rare microbial eukaryotes are largely underexplored in aquatic systems, primarily in freshwaters. All studies on the 'rare biosphere' are indeed restricted to marine prokaryotes, with only one theoretical attempt considering microbial eukaryotes (Caron and Countway, 2009). The improved unveiling of eukaryote diversity, such as in the present study, thus opens great perspectives for future identification of novel 'species', genes and metabolic pathways in lakes by complete deep sequencing of complex community, similar to what was done in marine ecosystems (Venter *et al.*, 2004; Delong *et al.*, 2006).

Diversity patterns

A greater diversity of fungi and chytrids was observed in Lake Pavin compared with Lake Aydat. The classical approach showed that fungi represent half of the OTUs identified in Lake Pavin, and one-third in Lake Aydat. Both approaches used confirmed that Lake Pavin contains around 20% more fungal diversity than Lake Aydat, suggesting a preference or the capacity of most fungi unveiled to adapt to a pelagic lifestyle. In both lakes, the four phyla of the main fungal divisions were represented in our sequences. The presence of *Glomeromycota* 18S rDNA gene sequences in the littoral area of Lake Aydat is consistent with their distribution in terrestrial rather than aquatic habitats and further implies a contribution of specimens in our samples from allochthonous inputs. The low allochthonous water input thus probably explained their absence in Lake Pavin, which have a water retention time about 10-fold longer than Lake Aydat (Camus *et al.*, 1993; Sarazin *et al.*, 1995). It is often difficult to compare results obtained from pyrosequencing and classical approaches, because of the use of different primers. Despite the lack of replicates, a good correlation was observed between the two methods for the identification of dominant fungi: the highest number of pyrosequence reads corresponds to the fungus OTUs with the highest number of sequences in the classical approach (Table S1a and b). Still, the quantitative information given by pyrosequencing data should be interpreted with care, because of the inherent bias of the PCR technique and also because no information is available on the number of copies of rRNA operons per fungal cells. Yet, such comparison revealed that the OTUs identified by the classical approach corresponded to around 13.8% (Station A, Pavin), 9.3% (C, Pavin), 19.2% (C, Aydat) and 10.4% (E, Aydat) of the pyrosequence reads obtained with fungus-designed primers. This result suggests a realistic representation of the fungal complete 18S rDNA gene sequences obtained with the classical approach. Several dominant chytrids were specific to each lake. The chytrids OTUs PA2009E8 and PA200911were dominant in Lake Pavin, while the OTU AY2009C4 corresponds to the main chytrid sequence isolated from Lake Aydat. These dominant chytrids may play an important role in the lakes, but further characterization of these species is needed to better understand their dynamics and functional importance.

The classical approach reveals the presence of species belonging to the group of *Viridiplantae* in both lakes, and a large diversity of *Alveolata* in Lake Aydat. Among the *Alveolata*, one rDNA gene sequence was corresponding to a *C. parvum*-like organism. Additional studies and analysis are required to characterize this parasite-like organism and to determine, if it is of concern to human

safety. Within the *Alveolata* group, some *Perkinsus* rDNA sequences were identified, which have not been yet characterized earlier by traditional microscope observations because these organisms do not display distinctive morphological characteristics (Burreson *et al.*, 2005; Gestal *et al.*, 2006). In recent molecular studies, a growing number of *Perkinsus* have been identified from freshwater environments (Brate *et al.*, 2010), indicating that more efforts are needed to culture and characterize members of this lineage. Only one cercozoan sequence (AY2009C14) was found in Lake Aydat. This was surprising since many cercozoans were isolated from both lakes in previous studies (Lefranc *et al.*, 2005; Lefèvre *et al.*, 2007; 2008). It may be suggested that cercozoans were poorly represented in the two lakes at the time of our sampling.

The pyrosequencing of the 18S rDNA gene hypervariable regions led to the identification of many other eukaryotes. Among them, the *Cryptophyta/Katablepharidophyta* was the most represented group, accounting for more than 50% of the total pyrosequencing reads when using fungus-designed primers, but was not detected with eukaryote primers. On the contrary, the groups of *Viridiplantae*, *Stramenopile* and *Arthropoda*, were only retrieved when using eukaryote primers. These two examples highlight the importance of using a combination of primers sets, to accurately determine environmental biodiversity. Few metazoan rDNA sequences were retrieved by the pyrosequencing approach. Since most metazoan could not pass through the filtration process, these sequences may come from larva, decomposed dead cells or free DNA molecules. In this study, we consider ourselves to have unveiled the most exhaustive catalogue of eukaryotes, and particularly that of fungi, identified from freshwater environments, by taking advantage of the complementarities between the classical and pyrosequencing approach, and the specificity of different primers.

Spatial distribution

The spatial distribution of taxa within the two lakes, estimated from our data, should be interpreted with care since no biological repeats were performed. Nevertheless, some general patterns can be observed such as a more heterogeneous distribution of taxa in Lake Aydat compared with Lake Pavin (Figs 5 and 6, Table S1a and b). This observation may be linked to the specific characteristics of both lakes, including the oligomesotrophic homogenous Lake Pavin compared with the heterogeneous eutrophic status of Lake Aydat. In addition, the deep-crater Lake Pavin characterized by a low drainage basin, the absence of an influent river and a 10-fold longer retention time, compared with Lake Aydat, establishes a typical pelagic environment in the whole water body of the

lake. These characteristics may explain why chytrids, which are typical pelagic organisms, were located throughout Lake Pavin, and preferentially in the central water column of Lake Aydat, which is subject to a higher level of water flux and of terrestrial input. As an example, the Lake Aydat chytrid OTU AY2009C4, belonging to the *Rhizophydiales* order, was at least 33 times more represented in the centre compared with the littoral area of the lake, a spatial pattern that is apparently consistent with the typical pelagic lifestyle of chytrids. The *Arthropoda* were mostly located in the littoral areas of both lakes; it is particularly true for Lake Aydat. This localization can be explained by the food source of these metazoans that might be more located in littoral areas. Finally, pyrosequence data obtained using fungus-designed primers revealed a preferential central localization of *Cryptophyta* in Lake Pavin, and for *Katablepharidophyta*, its sister group, in Lake Aydat. In general, in both lakes, the greatest diversity of eukaryotes was observed at the central point.

Conclusions

Very few species were shared between Lakes Pavin and Aydat, which indicates a different community composition in the same regional area. Geographical, physical and chemical factors of the biotope influence the species community structure and spatial variability. The characterization of the biodiversity is the first, but perhaps not the most important step to understand the functioning of an ecosystem. Matching sequences to organisms represents one of the last frontiers of the undiscovered biodiversity that challenge pelagic microbial ecology today (Sime-Ngando and Niquil, 2011). This study has to be completed by further characterization of the species, at least for the dominant ones, from their mode of nutrition (autotrophy, heterotrophy, mixotrophy), to their interactions between each other (food web, saprotrophism, parasitism, commensalism etc.). In order to do that, we first need to learn how to grow them in the laboratory. Chytrids play a major role in lake ecology, primarily as parasites of inedible phytoplankton, and by transferring these matter and energy (in the form of zoospores) to higher trophic levels (Lefèvre *et al.*, 2008). Sequencing the genome of chytrids, developing molecular tools (mutagenesis) and performing phenotypic analysis will lead to a major breakthrough in the understanding of the functioning of lake ecosystems. This knowledge could be integrated into mathematical models to predict ecosystem functioning. In this respect, Lake Pavin, because of its species composition homogeneity, is an ideal model for this type of study. Although the data presented in this article are only a snapshot of the whole lake's biodiversity, the massive parallel pyrosequencing of the 18S rRNA gene hypervari-
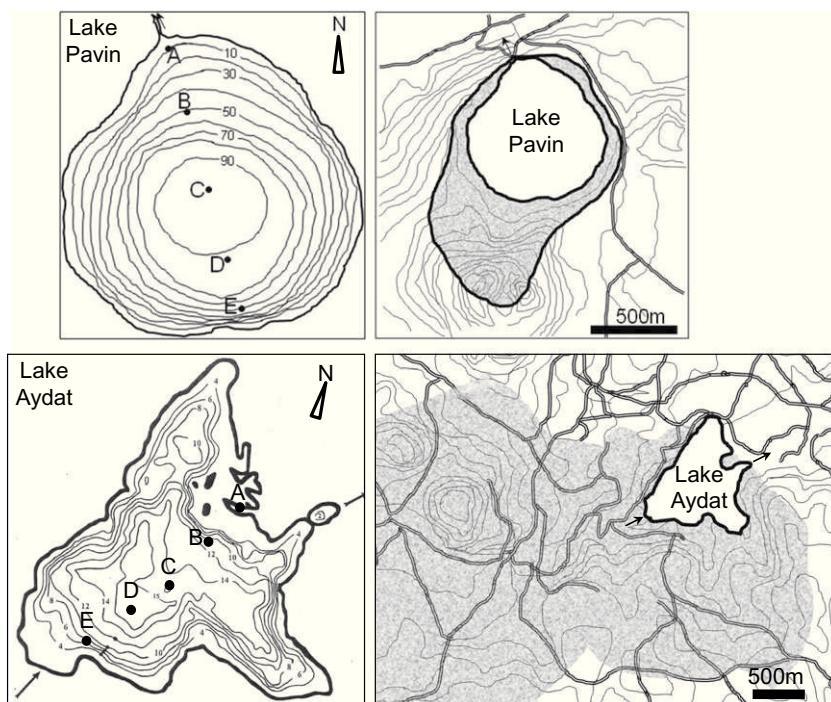
Fig. 7. Bathymetric maps (left) and water catchment area (right) from Lake Pavin and Lake Aydat. The bathymetric maps of Lake Pavin (from Delbecque, 1898) and Lake Aydat (from Rabette and Lair, 1999) display the location of the five sampling stations: two littoral (A and E), one central (C) and two intermediaries (B and D). On the bathymetric maps, the arrows show the flux of water and the numbers indicate the depth in metre. The two water catchment areas of Lake Pavin (from Lair, 1978) and Lake Aydat (this study) show the larger size of Lake Aydat drainage basin (grey shade) compared with Lake Pavin. Arrows indicate the flux of water, lines show the ground topography and roads are indicated by double lines.

able regions allows the assessment of the impact of space, time and complex environments on microbial communities. A timeline survey of the lake's biodiversity, with a regular sampling over several years, is needed. This will give crucial information to understand the succession of stable and unstable distribution of eukaryote community, the dynamics, interactions and successive blooms of species, and the changes of the eukaryote biodiversity due to climate change or anthropogenic activities.

Experimental procedures

Sampling procedure and DNA extraction

Lake Pavin (45°30′N, 2°53″W) is a meromictic, dimictic, oligomesotrophic lake situated in the Massif Central of France. It is a deep volcanic mountain lake ($Z_{max} = 92$ m), characterized by a permanently anoxic monimolimnion from a 60 m depth downwards. This site offers a unique environment with low human influences, characterized by a small surface area (44 ha), about equal to the drainage basin area (50 ha), with no river inflow (Fig. 7). The water retention time in the mixolimnion is approximately 10 years, with an average water turnover rate of 0.09 per year. In the monimolimnion, this water turnover rate decrease to 0.005 per year (Camus *et al.*, 1993). Lake Aydat (45°39′N, 2°59″W) is a small ($Z_{max} = 15$ m, surface area = 60 ha) dimictic, eutrophic lake, also located in the Massif Central of France. It was formed when a lava flow dammed the small river Veyre. Compared with the surface of the lake, the catchment area (2.5×10^3 ha) is very large and contains intensive agricultural lands (Fig. 7). The water

retention time of Lake Aydat is approximately 1 year (Sarazin *et al.*, 1995).

For each of the two lakes, five samples were collected along a transect across the lake, including two littoral (points A and E), two intermediary (point B and D) and one central site (point C) located in the deepest area of the lake (Fig. 7). For Lake Pavin, station A is located near the lake outlet, while on the other side of the lake, station E is influenced by the catchment effects (Fig. 7). For Lake Aydat, the stations A and B were located near the Veyre River outlet and the within-lake islands, while station E was located next to the lake 'estuary' (Fig. 7). Samples were collected during the end of the thermal stratification period, on 16 July 2008 for Lake Pavin and on 21 July 2008 for Lake Aydat. In Lake Pavin, the temperature of surface waters fluctuated between 17.5°C and 17.8°C, and pH values were at 8.33, 8.30, 8.27, 7.94 and 8.28 for the sampling points A, B, C, D and E respectively. In Lake Aydat, the temperature fluctuated between 19.6°C and 19.5°C and pH values were at 9.70, 9.74, 9.81, 9.85 and 9.83 for sampling points A, B, C, D and E respectively.

The whole euphotic water column [determined from Secchi disk measurements, $Z_{eu} = 1.7 \times Z_s$ (Reynolds, 1984)] was collected at each sampling point using a flexible plastic tube (diameter 4 cm) provided by a rope connecting the ballasted bottom of the tube with a surface manipulator. The euphotic depths were at 4.5 m and 20 m in Aydat and Pavin respectively. Samples were immediately pre-filtered through 150-µm-pore-size nylon filter (i.e. to eliminate larger metazoans), poured into clean recipients previously washed with the lake water, and returned back to the laboratory for immediate processing. Subsamples of 150 and 300 ml from Aydat and Pavin, respectively, were filtered onto 0.6-µm-pore-size polycarbonate filters (47 mm diameter) using a vacuum pump

(pressure < 100 mbar) to collect all planktonic microorganisms. The filters were stored frozen at −80°C until DNA extraction.

DNA from each sampling point was extracted using the kit NucleoSpin® Plant DNA extraction Kit (Macherey-Nagel, Düren, Germany) adapted for fungal material. The initial step aimed at digesting fungal chitin wall. Filters were then incubated overnight at 30°C with 400 Units of lyticase enzyme (Sigma, NSW, Australia) in a 500 µl sorbitol based buffer (Karakousis *et al.*, 2006). Then sodium dodecyl sulfate (final concentration 1%) and proteinase K (final concentration, 0.1 mg ml⁻¹) were added and incubated 1 h at 37°C. All subsequent DNA extraction steps were conducted following the manufacturer's instructions.

The classical approach: PCR, cloning/sequencing and phylogenetic analyses

The primers sense NS1 (GTAGTCATATGCTTGTCTC) and antisense ITS4 (TTCCTCCGCTTATTGATATGC) known to preferentially amplify fungal DNA (White *et al.*, 1990) were used to amplify the complete 18S rDNA, ITS1 (Internal Transcribed Spacer region), 5.8S rDNA, ITS2 and partial 28S rDNA region (Fig. S1). Only the 18S rDNA gene sequences were used in the phylogenetic analyses. PCRs were carried out in 50 µl volume according to standard conditions for *Taq* DNA polymerase (Bioline) with 50 ng of environmental DNA as template. After the denaturation step at 95°C for 3 min, 34 cycles of amplification were performed with a GeneAmp PCR System Apparatus (Applied Biosystems) as follows: 30 s at 95°C, 30 s at 59°C and 3 min at 72°C. The reaction was completed with an extension step at 72°C for 10 min. The products were separated on agarose gel electrophoresis in order to confirm the presence of bands of the expected size (around 2600 bp). PCR products were precipitated by sodium acetate/ethanol and titrated for cloning. For each sampling point, a genetic library was constructed. An aliquot of PCR product was cloned using the TOPO-TA cloning kit (Invitrogen) following the manufacturer's recommendations. The presence of the insert in the colonies was checked by PCR amplification using M13 forward and reverse universal primers. PCR products of the correct size were analysed by RFLP using the restriction enzyme HaeIII (Invitrogen). Clones showing the same RFLP patterns were grouped into a single OTU and one representative of each OTUs was sequenced from minipreparations of plasmid DNA using the Nucleospin kit (Macherey-Nagel). A total of 146 and 143 SSU rDNA gene sequences for Lake Pavin and Lake Aydat, respectively, were obtained and aligned using CLUSTALW (Thompson *et al.*, 1994). Sequences sharing more than 99% identity were grouped into OTUs. A total of 46 OTUs in Lake Pavin and 63 OTUs in Lake Aydat were identified. One representative sequence of each OTU was submitted to a BLAST search (Altschul *et al.*, 1990) on the non-redundant nucleotide database (NCBI) for an approximate phylogenetic affiliation and in order to select representative taxa for tree constructions. The resulting alignments including representative OTU sequences and reference sequences were corrected manually and regions of ambiguous alignment were removed using the Bioedit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Phylogenetic trees were built using both the Neighbour-Joining (NJ) method from the PHYLIP package (http://evolution.genetics.washington.edu/phylip.html) and the Bayesian method from MrBayes3 software (http://mrbayes.csit.fsu.edu/index.php) (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Topologies of the phylogenetic trees generated by the two methods were similar. Only the tree generated by the Bayesian method is displayed since this method was considered to give more accurate estimates (Alfaro *et al.*, 2003). The Bayesian posterior probabilities are showed at each node.

The pyrosequencing approach

DNAs extracted from the sampling points A and C of Lake Pavin, and C and E of Lake Aydat were used in a pyrosequencing approach. Two set of primers were used in this study (Fig. S1). The first set of primers, 18S-82F (GAAACT GCGAATGGCTC) and Ek-516r (ACCAGACTTGCCCTCC) (Casamayor *et al.*, 2002; Lopez-Garcia *et al.*, 2003; Lepere *et al.*, 2006), was designed to amplify a 480 bp region containing the complete V2 and V3 domains of all eukaryote SSU rDNA gene. We will refer to this universal set of primers as the eukaryote primers. The second set, composed of primers nu-SSU-0817-5′ (TTAGCATGGAATAATRRAATAGGA) and nu-SSU-1196-3′ (TCTGGACCTGGTGAGTTTCC), was previously designed by Borneman and Hartin (2000) to preferentially amplify an around 400 bp region containing the V4 (partial) and V5 (complete) variable domains of the SSU rDNA gene from all four major phyla of fungi. We will refer to this set of primers as the fungus-designed primers. A 10 bp tag specific to each sample, a 4 bp TCAG key and a 26 bp adapter for the GsFLX technology were added to the sequences of the primers. PCRs were carried out according to standard conditions for *Taq* DNA polymerase with 20 ng of environmental DNA as template. After the denaturation step at 95°C for 3 min, 20 cycles of amplification were performed with a GeneAmp PCR System Apparatus (Applied Biosystems) as follows: 30 s at 95°C, 30 s at 60°C and 1 min at 72°C. The pyrosequencing project was performed by the Genoscreen company (Institut Pasteur de Lille, France). The library and the 454 titanium (Roche) pyrosequencing run were prepared following the manufacturer's recommendations. A total of 103 435 sequences with 53 856 and 49 579 sequences from the eukaryote and fungus-designed primers, respectively, were obtained. The reads were classified according to the tag sequence corresponding to each of the four samples of interest. With the eukaryote primers, 12 133 and 11 386 sequences from the sampling stations A and C in Lake Pavin, and 15 259 and 15 078 sequences from the sampling stations C and E in Lake Aydat, were, respectively, obtained. Using the fungus-designed primers, a total of 8018 and 10 527 sequences from the sampling stations A and C in Lake Pavin, and 16 672 and 14 362 sequences from the sampling stations C and E in Lake Aydat, were, respectively, obtained. Primers, tag and key fragments were subsequently removed before analysing the sequences. Globally, around two-thirds of the reads showed a length above 200 bp and one-third of the reads had a length above 400 bp.

For identification, the resulting sequences were compared with the Silva SSU rRNA database (http://www.arb-silva.de/) release 102 (updated on February 15, 2010) comprising

1 246 462 18S SSU rRNA sequences (including 134 351 eukaryotic sequences) using the BLASTN software (Altschul *et al.*, 1990). BLAST results (with 10^{-5} *E*-value threshold) were visualized using the metagenomic software MEGAN (Huson *et al.*, 2007). This software allows exploring the taxonomic content of the samples based on the NCBI taxonomy using the option 'import BLASTN'. The program uses several thresholds to generate sequence–taxon matches. The 'min-score' filter, corresponding to a bit score cut-off value, was set at 35. The «top-percent» filter used to retain hits, whose scores lie within a given percentage of the highest bit score, was set at ten. The «min-support core» filter, used to set a threshold for the minimum number of sequences that must be assigned to a taxon, was set to three. These stringent parameters result in the 'safe' assignment of many sequences to lower branch (with less precision) of the taxonomic tree. Distribution of the sequences was schematically represented by trees and pie diagrams.

Comparison between the classical and the pyrosequencing approaches

The pyrosequences from the SSU hypervariable regions were compared with the full-length 18S rDNA gene sequences obtained with the classical approach. This comparison allowed the identification of the dominant species, the determination of the within-lake spatial variation of species, and an evaluation of the correlation and complementarities between the two approaches. The 18S rDNA gene pyrosequences were submitted to a BLAST analysis (with 10^{-23} *E*-value threshold for BLASTN) against databases containing the full-length 18S rDNA gene sequences from each lake obtained by the classical approach. This analysis was performed independently for both lakes. Several criteria were applied to select relevant BLAST results, including a minimum query sequence of 60 bp and a minimum of 99% similarity over 80% of the sequence length with match starting, at least, from the third nucleotide of the query sequence. BLAST results were manually checked in order to remove duplicates, i.e. a pyrosequence read that was matching several times with the same full-length 18S rDNA gene sequence (i.e. from the classical approach) on different sites. However, 18S rDNA gene pyrosequences could still match with several full-length 18S rDNA gene sequences from the classical approach, when their lengths are too short for a precise identification.

Rarefaction curves

The quality of the sampling effort was assessed through the calculation of rarefaction curves, i.e. the number of OTUs versus the number of clones (Hughes *et al.*, 2001). Rarefaction curves for sequences obtained from the classical and the pyrosequencing approaches were done independently but following the same procedure. For the classical approach, all the obtained 18S rDNA gene sequences were considered for analysis. For the pyrosequencing approach, all the obtained 18S rDNA gene sequences with a length over 400 bp were considered. This size limit was set in order to include in the analysis the entire variable V2 (obtained with the eukaryote primers) and V5 (obtained with the fungus-designed primers)

regions of the 18S rDNA gene sequence. With the eukaryote primers, a total of 3786 and 3684 sequences for stations A and C in Lake Pavin, and 4182 and 5141 sequences for stations C and E in Lake Aydat, were, respectively, analysed. With the fungus-designed primers, a total of 2948 and 3691 sequences for stations A and C in Lake Pavin, and 5935 and 5444 sequences for stations C and E in Lake Aydat, respectively, were analysed. The sequences considered for both approaches were aligned using the program MUSCLE (Edgar, 2004) (with parameters –diags and –maxiters 2) and were manually corrected using the Bioedit software to remove ambiguous terminal region of the alignment. The resulting alignment was used as input for the Mothur program (Schloss *et al.*, 2009), with a cut-off value set to 0.01 and 0.03 (i.e. OTUs with differences that do not exceed 1% and 3%) for the analysis.

Accession numbers

Nucleotide sequences obtained from the complete 18S, ITS1, 5.8S, ITS2 and partial 28S sequences were deposited in GenBank under the Accession No. HQ191282–HQ191427 for Lake Pavin samples and HQ219333–HQ219474 for Lake Aydat samples. The pyrosequences were deposited in GenBank-SRA under the accession numbers SRA012393.5.

References

Alexopoulos, C.M., Mims, C.W., and Blackwell, M. (1996) *Introductory Mycology*. New York, USA: John Wiley & Sons.

Alfaro, M.E., Zoller, S., and Lutzoni, F. (2003) Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Mol Biol Evol* **20:** 255–266.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Bass, D., Howe, A., Brown, N., Barton, H., Demidova, M., and Michelle, H., *et al.* (2007) Yeast forms dominate fungal diversity in the deep oceans. *Proc Biol Sci* **274:** 3069–3077.

Berger, L., Speare, R., Daszak, P., Green, D.E., Cunningham, A.A., and Goggin, C.L., *et al.* (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proc Natl Acad Sci USA* **95:** 9031–9036.

Borneman, J., and Hartin, R.J. (2000) PCR primers that amplify fungal rRNA genes from environmental samples. *Appl Environ Microbiol* **66:** 4356–4360.

Brate, J., Logares, R., Berney, C., Ree, D.K., Klaveness, D., Jakobsen, K.S., and Shalchian-Tabrizi, K. (2010) Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *ISME J* **4:** 1144–1153.

Burreson, E.M., Reece, K.S., and Dungan, C.F. (2005) Molecular, morphological, and experimental evidence support the synonymy of *Perkinsus chesapeaki* and *Perkinsus andrewsi*. *J Eukaryot Microbiol* **52:** 258–270.

Camus, G., Michard, G., Olive, P., Boivin, P., Desgranges, P., and Jezequel, D., *et al.* (1993) Risques d'éruption gazeuse en Auvergne. *Bull Soc Geol Fr* **164:** 767–778.

Canter, H.M. (1950) Fungal parasites of the phytoplankton. I. Studies on British chytrids X. *Ann Bot* **14:** 263–289.

Canter, H.M., and Lund, J.W.G. (1948) Studies on plankton parasites. I. Fluctuations in the numbers of *Asterionella formosa* Hass. in relation to fungal epidemics. *New Phytol* **47:** 238–261.

Canter, H.M., and Lund, J.W.G. (1951) Studies on plankton parasites. III. Examples of the interaction between parasitism and other factors determining the growth of diatoms. *Ann Bot* **15:** 359–371.

Caron, D.A., and Countway, P.D. (2009) Hypotheses on the role of the protistan rare biosphere in a changing world. In *Rassoulzadegan at Villefranche-Sur-Mer: 3 Decades of Aquatic Microbial Ecology*. Dolan, J., Mostajir, B., and Sime-Ngando, T. (eds). Oldendorf/luhe, Germany: Inter Research. AME Special 3, Aquatic Microb Ecol, **57,** pp. 227–238.

Carrias, J.F., and Amblard, C. (1998) Seasonal dynamics of free and attached heterotrophic nanoflagellates in an oligomesotrophic lake. *Freshw Biol* **39:** 91–101.

Carrias, J.F., Amblard, C., and Bourdier, G. (1996) Protistan bacterivory in an oligomesotrophic lake: importance of attached ciliates and flagellates. *Microb Ecol* **31:** 249–268.

Casamayor, E.O., Massana, R., Benlloch, S., Ovreas, L., Diez, B., and Goddard, V.J., *et al.* (2002) Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environ Microbiol* **4:** 338–348.

Delbecque, A. (1898) Les lacs français. In *Typographie*. Paris, France: Chamerot & Renouard, p. 436.

Delong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., and Frigaard, N.U., *et al.* (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311:** 496–503.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., and Hugenholtz, P. (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* **4:** 642–647.

Gestal, C., Novoa, B., Posada, D., Figueras, A., and Azevedo, C. (2006) *Perkinsoide chabelardi* n. gen., a protozoan parasite with an intermediate evolutionary position: possible cause of the decrease of sardine fisheries? *Environ Microbiol* **8:** 1105–1114.

Gromov, B.V., Pljusch, A.V., and Mamkaeva, K.A. (1999a) Cultures of *Rhizophydium* spp. (chytridiales) parasites of chlorococcalean algae. *Arch Hydrobiol Suppl Algol Stud* **95:** 115–123.

Gromov, B.V., Pljusch, A.V., and Mamkaeva, K.A. (1999b) Morphology and possible host range of *Rhizophydium algavorum* sp. Nov (Chytridiales) and obligate parasite of algae. *Protistology* **1:** 62–65.

Hawksworth, D.L. (2001) Mushrooms: the extent of the unexplored potential. *Int J Med Mushrooms* **3:** 333–337.

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., *et al.* (2007) A higher-level phylogenetic classification of the Fungi. *Mycol Res* **111:** 509–547.

Holfeld, H. (1998) Fungal infections of the phytoplankton: seasonality, minimal host density, and specificity in a mesotrophic lake. *New Phytol* **138:** 507–517.

Huelsenbeck, J.P., and Ronquist, F. (2001) MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H., and Bohannan, B.J.M. (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* **67:** 4399–4406.

Huse, S.M., Dethlefsen, L., Huber, J.A., Mark Welch, D., Relman, D.A., and Sogin, M.L. (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet* **4:** e1000255.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17:** 377–386.

Hyde, K.D., and Goh, T.K. (1998a) Fungi on submerged wood in Lake Barrin, North Queensland, Australia. *Mycol Res* **102:** 739–749.

Hyde, K.D., and Goh, T.K. (1998b) Fungi on submerged wood in the River St Marie-Louis, The Seychelles. *S Afr J Bot* **64:** 330–336.

Hyde, K.D., and Goh, T.K. (1998c) Fungi on submerged wood in the Palmiet River, Durban, South Africa. *S Afr J Bot* **64:** 151–162.

Hyde, K.D., and Goh, T.K. (1999) Fungi on submerged wood from the River Coln, England. *Mycol Res* **103:** 1561–1574.

Ibelings, B.W., De Buin, A., Kagami, M., Rijkeboer, M., Brehm, M., and Van Donk, E. (2004) Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). *J Phycol* **40:** 437–453.

James, T.Y., Porter, D., Leander, C.A., Vilgalys, R., and Longcore, J.E. (2000) Molecular phylogenetics of the Chytridiomycota support the utility of ultrastructural data in chytrid systematics. *Can J Bot* **78:** 336–350.

James, T.Y., Letcher, P.M., Longcore, J.E., Mozley-Standridge, S.E., Porter, D., Powell, M.J., *et al.* (2006a) A molecular phylogeny of the flagellated fungi (Chytridiomy-

cota) and description of a new phylum (Blastocladiomy-cota). *Mycologia* **98:** 860–871.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstet-ter, V., Cox, C.J., *et al*. (2006b) Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818–822.

Jobard, M., Rasconi, S., and Sime-Ngando, T. (2010) Diver-sity and functions of microscopic fungi: a missing compo-nent in pelagic food webs. *Aquat Sci* **72:** 255–268.

Kagami, M., Van Donk, E., de Buin, A., Rijkeboer, M., and Ibelings, B.W. (2004) *Daphnia* can protect diatoms from fungal parasitism. *Limnol Oceanogr* **49:** 680–685.

Kagami, M., von Elert, E., Ibelings, B.W., de Bruin, A., and van Donk, E. (2007) The parasitic chytrid, *Zygorhizidium*, facilitates the growth of the cladoceran zooplankter, *Daphnia*, in cultures of the inedible alga, *Asterionella*. *Proc Biol Sci* **274:** 1561–1566.

Karakousis, A., Tan, L., Ellis, D., Alexiou, H., and Wormald, P.J. (2006) An assessment of the efficiency of fungal DNA extraction methods for maximizing the detection of medi-cally important fungi using PCR. *J Microbiol Methods* **65:** 38–48.

Kirk, P.M., and Ainsworth, G.C. (2008) *Ainsworth & Bisby's Dictionary of the Fungi*, 10th edn. Collingwood, Vic., Aus-tralia: CSIRO Publishing.

Kudoh, S. (1990) *Studies on the Population Control Mecha-nisms of a Planktonic Diatom,* Astronella formosa *Hassall, by Fungal Parasitism in Natural Waters*. Tokyo, Japan: University of Tokyo.

Kunin, V., Engelbrektson, A., Ochman, H., and Hugenholtz, P. (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12:** 118–123.

Lair, N. (1978) Repartition spatio-temporelle, biomasse et production des populations zooplanctoniques du lac d'aydat en periode estivale. *Hydrobiologia* **61:** 237–256.

Lefèvre, E., Bardot, C., Noel, C., Carrias, J.F., Viscogliosi, E., Amblard, C., and Sime-Ngando, T. (2007) Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep mero-mictic lake. *Environ Microbiol* **9:** 61–71.

Lefèvre, E., Roussel, B., Amblard, C., and Sime-Ngando, T. (2008) The molecular diversity of freshwater picoeukary-otes reveals high occurrence of putative parasitoids in the plankton. *PLoS ONE* **3:** e2324.

Lefranc, M., Thenot, A., Lepere, C., and Debroas, D. (2005) Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl Environ Microbiol* **71:** 5935–5942.

Lepere, C., Boucher, D., Jardillier, L., Domaizon, I., and Debroas, D. (2006) Succession and regulation factors of small eukaryote community composition in a lacustrine ecosystem (Lake Pavin). *Appl Environ Microbiol* **72:** 2971–2981.

Letcher, P.M., and Powell, M.J. (2005) Phylogenetic position of *Phlyctochytrium planicorne* (Chytridiales, Chytridiomy-cota) based on zoospore ultrastructure and partial nuclear LSU rRNA gene sequence analysis. *Nova Hedwigia* **80:** 135–146.

Letcher, P.M., Powell, M.J., Churchill, P.F., and Chambers, J.G. (2006) Ultrastructural and molecular phylogenetic delineation of a new order, the Rhizophydiales (Chytridi-omycota). *Mycol Res* **110:** 898–915.

Libkind, D., Moline, M., Sampaio, J.P., and van Broock, M. (2009) Yeasts from high-altitude lakes: influence of UV radiation. *FEMS Microbiol Ecol* **69:** 353–362.

Longcore, J.E. (1992) Morphology and zoospore ultrastruc-ture of *Chytriomyces angularis* sp. nov. (Chytridiales). *Mycologia* **84:** 442–451.

Lopez-Garcia, P., Philippe, H., Gail, F., and Moreira, D. (2003) Autochthonous eukaryotic diversity in hydrothermal sedi-ment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Natl Acad Sci USA* **100:** 697–702.

Moreno, C., Romero, J., and Espejo, R.T. (2002) Polymor-phism in repeated 16S rRNA genes is a common property of type strains and environmental isolates of the genus *Vibrio*. *Microbiology* **148:** 1233–1239.

Okamoto, N., and Inouye, I. (2005) The katablepharids are a distant sister group of the Cryptophyta: a proposal for *Kat-ablepharidophyta divisio* nova/*Kathablepharida phylum* novum based on SSU rDNA and beta-tubulin phylogeny. *Protist* **156:** 163–179.

Powell, M.J. (1993) Looking at mycology with a Janus face: a glimpse at Chytridiomycetes active in the environment. *Mycologia* **83:** 1–20.

Rabette, C., and Lair, N. (1999) Spatial and temporal distri-bution of benthic stages of *Cyclops vicinus* and *Chaoborus flavicans* in relation to abiotic factors and benthic fauna. *Hydrobiologia* **390:** 61–72.

Raja, H.A., and Shearer, C.A. (2008) Freshwater asco-mycetes: a new and noteworthy species from aquatic habi-tats in Florida. *Mycologia* **100:** 467–489.

Rasconi, S., Jobard, M., Jouve, L., and Sime-Ngando, T. (2009) Use of calcofluor white for detection, identification, and quantification of phytoplanktonic fungal parasites. *Appl Environ Microbiol* **75:** 2545–2553.

Rasconi, S., Jobard, M., and Sime-Ngando, T. (2011) Para-sitic fungi of phytoplankton: ecological role and implications for microbial food webs. *Aquat Microb Ecol* **62:** 123–137.

Reynolds, C.S. (1984) Phytoplankton periodicity: the interac-tions of form, function and environmental variability. *Fresh-water Biology* **14:** 111–142.

Richards, T.A., Vepritskiy, A.A., Gouliamova, D.E., and Nierzwicki-Bauer, S.A. (2005) The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lin-eages. *Environ Microbiol* **7:** 1413–1425.

Ronquist, F., and Huelsenbeck, J.P. (2003) MrBayes 3: baye-sian phylogenetic inference under mixed models. *Bioinfor-matics* **19:** 1572–1574.

Sarazin, G., Gaillard, J.F., Philippe, L., and Rabouilles, C. (1995) Organic matter mineralization in the pore water of eutrophic lake (Aydat Lake, Puy de Dôme, France). *Hydro-biologia* **315:** 95–118.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hart-mann, M., Hollister, E.B., *et al*. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial commu-nities. *Appl Environ Microbiol* **75:** 7537–7541.

Shearer, C.A. (2001) *The Distribution of Freshwater Filamen-tous Ascomycetes*. Enfield, New Hampshire: Science Publishers.

Shearer, C.A., Descals, E., Kohlmeyer, J., Marvanova, L., Padgett, L., Porter, D., *et al.* (2007) Fungal biodiversity in aquatic habitats. *Biodivers Conserv* **16:** 49–67.

Sime-Ngando, T., and Niquil, N. (2011) Disregarded microbial diversity and ecological potentials in aquatic systems. In *Hydrobiologia, Special Issue*. Sime-Ngando, T., and Niquil, N. (eds). New York, USA: Springer, pp. 1–4.

Sime-Ngando, T., Lefèvre, E., and Gleason, F.H. (2011) Hidden diversity among aquatic heterotrophic flagellates: ecological potentials of zoosporic fungi. *Hydrobiologia* **659:** 5–22.

Sogin, M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., *et al.* (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103:** 12115–12120.

Spatafora, J.W., Sung, G.H., Sung, J.M., Hywel-Jones, N.L., and White, J.F., Jr (2007) Phylogenetic evidence for an animal pathogen origin of ergot and the grass endophytes. *Mol Ecol* **16:** 1701–1711.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673–4680.

Treseder, K.K. (2005) Nutrient acquisition strategies of fungi and their relation to elevated atmospheric CO_2. In *The Fungal Community: Its Organization and Role in the Ecosystem*. Dighton, J., White, J.F., and Oudemans, P. (eds). Boca Ration, FL, USA: CRC Press, pp. 713–731.

Van Donk, E., and Ringelberg, J. (1983) The effect of fungal parasitism on the succession of diatoms in lake Maarsseveen-I (The Netherland). *Freshw Biol* **13:** 241–251.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.

Watling, R. (2005) Fungal conservation: some impressions – a personal view. In *The Fungal Community: Its Organization and Role in the Ecosystem*. Dighton, J., White, J.F., and Oudemans, P. (eds). Boca Ration, FL, USA: CRC Press, pp. 881–896.

White, T.J., Bruns, T.D., Lee, S.B., and Taylor, J.W. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols: a Guide to Methods and Applications*. Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds). London, UK: Academic Press, pp. 315–322.

Wong, M.K.M., Goh, T.K., Hodgkiss, I.J., Hyde, K.D., Ranghoo, V.M., Tsui, C.K.M., *et al.* (1998) Role of fungi in freshwater ecosystems. *Biodivers Conserv* **7:** 1187–1206.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Schematic representation of genomic rDNA region and localization of the primers used for the classical and the pyrosequencing approaches of this study. This figure shows a schematic representation of the genomic rDNA region (18S, ITS1, 5.8S, ITS2, 28S, IGS1, 5S and IGS2) with the localization of the primers and resulting amplicons that were obtained by the classical and pyrosequencing approaches. The hypervariable regions of the 18S rRNA gene are displayed by black boxes numbered from V1 to V9. The sequence of each primer is given between brackets.

Table S1. Number of sequences composing OTUs in the clone libraries, phylogenetic affiliation of the representative OTUs and corresponding numbers of matching read in the pyrosequence data. These tables summarize the number of clones obtained from each OTUs from Lake Pavin (S1a) and Lake Aydat (S1b) clone libraries. For each OTU, the taxon, the closest relative with its percentage of sequence identity, the number of clones composing OTU from the classical approach and their corresponding sampling stations localization (see Fig. 1), and the number and percentage of matching reads from the pyrosequencing approach (both using fungi and eukaryote primers) are given. The '–' symbol, for sequence PA2009E6, indicates that no closest relative sequence and it corresponding percentage could be determined.

Table S2. Number of pyrosequencing reads assigned to each taxonomic group. This table displays the number of pyrosequencing reads, obtained from the two lakes, sampling stations and primer sets (fungi and eukaryotes), which were assigned to a taxonomic group after a BLASTN search. 'No hit' corresponds to reads having no homologous sequence in the Silva database (threshold $E = 10^{-5}$), and 'Not assigned' corresponds to reads having a match in the Silva database but with an unclear taxonomic assignment.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

Desidero ringraziare con sincero affetto il prof Pier Luigi Fiori per avermi accolto nel suo laboratorio e avermi dato la possibilità di cominciare il mio dottorato di ricerca. Grazie Gigi per aver creduto in me e avermi dato questa grande opportunità che mi ha enormemente arricchito sia dal punto di vista professionale che umano.

Je tiens à exprimer ma profonde gratitude au Docteur Eric Viscogliosi pour m'avoir donné l'occasion de faire mon Erasmus au sein de son laboratoire à l'Institut Pasteur de Lille, ce qui a été le point de départ de cette incroyable aventure. Merci Eric d'avoir cru en mis capacités, de m'avoir donné la possibilité de faire partie de ton équipe, et surtout d'avoir vu ma passion pour ce travaille. Sans toi je ne serai jamais arrivée jusqu'ici. Tu n'as jamais cessé de m'encourager et de me donner des conseils précieux, qui m'ont permis de trouver confiance en moi et de grandir tant sur le plan professionnel qu'humain. Merci également pour ta bonne humeur permanente, pour les nombreux rires qui rendent mes journées moins difficiles malgré la distance qui me sépare de ma famille. Merci pour ton amitié.

Un très grand merci à toute l'équipe pour votre amitié et pour la bonne ambiance qui vous avez creé tous les jours, cela a été un véritable plaisir de travailler avec vous. Que des femmes extraordinaires, merci: Laurine, Dima, Amandine, Danielle, Nausicaa, Magali, Gabriela, Anna, Karine, Sophie, Laurence, Annie, Emilie, Cécile Marie, Valérie, Christine. Merci El Moukhtar et Eduardo de m'avoir transmis votre passion pour la microscopie. Mais surtout merci Sadia, grazie cara amica, tu es toujours présente, sei unica. Un merci particulier va à Clea, merci petit chef pour tes conseils, pour nos petites disputes comme pour nos grands moments de rigolade. Merci pour être une bonne collègue mais surtout une amie. Dans cette liste, je ne peux pas oublier Giovanna, merci pour ton formidable aide, merci pour ton "heritage", si tu n'avais pas été là au debut de cet experience tout aurait été différent.

Je tiens à remercier aussi tous les membres de l'équipe de Clermont Ferrand pour leur aide et leur gentillesse lors notre collaboration.

Merci au centre AZM de Tripoli et au CHR de Tours de nous avoir permis de collaborer avec eux.

Thanks a lot to Dr. Moriya Ohkuma for the possibility to collaborate with his team in this important project.

Un grande grazie à Lobna, grazie perché é sempre un piacere passare le serate in tua compagnia a discutere, a confrontarsi ma soprattutto a ridere insieme.

Thanks to Thavy, my first French friend thanks for the beautiful moments we spent together. I will never forget.

Merci beaucoup a vous: Mathieu, Nadege D., Marie, Nadege G, Julien et Nicolas pour les bons moments passés ensemble.

Un immense merci va à Ben, mon collocateur, mais surtout mon meilleur ami. Merci beaucoup simplement pour être là. Merci pour les moments extraordinaires qu'on a vécu ensemble, pour les soirées passées à discuter des problèmes du monde, pour tes courses de philosophie, pour les petits voyages, pour les rires et la bonne humeur... et aussi merci pour m'avoir appris à connaître toutes les bières belges…

Un enorme ringraziamento va anche a Paola, compagna di tante avventure, grazie perchè anche se siamo lontane tu sei sempre la, pronta a darmi una parola di conforto.

Un grazie a te Zietto, al tuo caloroso affetto, ai tuoi insegnamenti, alla tua presenza nei momenti più importanti, al tuo silenzioso ma importante sostegno.

Grazie alla bambina con cui son cresciuta, con cui ho riso e sognato... alla donna forte e coraggiosa che sei diventata. Grazie Maria perché sei sempre pronta a trovare una parola di conforto per incoraggiarmi a non arrendermi e a sostenermi, grazie per dimostrarmi sempre quanto sei fiera di me.

In fine il ringraziamento, quello più grande e speciale, va a voi papà e mamma. I vostri insegnamenti, la vostra semplicità e i vostri valori sono alla base dei miei traguardi e senza di voi tutto questo non si sarebbe potuto realizzare in questo modo. Grazie perché quando ho pensato di arrendermi e mollare tutto, mi avete dato la forza per continuare e lottare, perché anche nei momenti più difficili posso contare su di voi. Grazie perché qualunque sia la decisione da prendere mi incoraggiate e mi apoggiate con orgoglio, anche se questo significava stare lontani. Grazie per dimostrarmi sempre quanto credete in me e per vivere insieme le emozioni dei piccoli traguardi di ogni giorno. GRAZIE!!