



**HAL**  
open science

# Why the –omic future of Apicomplexa should include Gregarines

Julie Boisard, Isabelle Florent

► **To cite this version:**

Julie Boisard, Isabelle Florent. Why the –omic future of Apicomplexa should include Gregarines. *Biology of the Cell*, 2020, 10.1111/boc.202000006 . hal-02553206

**HAL Id: hal-02553206**

**<https://hal.science/hal-02553206v1>**

Submitted on 24 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article title: Why the –omic future of Apicomplexa should include Gregarines.

Names of authors: Julie BOISARD<sup>1,2</sup> and Isabelle FLORENT<sup>1</sup>

Authors affiliations:

1. Molécules de Communication et Adaptation des Microorganismes (MCAM, UMR 7245), Département Adaptations du Vivant (AVIV), Muséum National d'Histoire Naturelle, CNRS, CP52, 57 rue Cuvier 75231 Paris Cedex 05, France.
2. Structure et instabilité des génomes (STRING UMR 7196 CNRS / INSERM U1154), Département Adaptations du vivant (AVIV), Muséum National d'Histoire Naturelle, CP 26, 57 rue Cuvier 75231 Paris Cedex 05, France.

Short Title: Gregarines –omics for Apicomplexa studies

Corresponding author: Isabelle FLORENT, Molécules de Communication et Adaptation des Microorganismes (MCAM, UMR 7245), Département Adaptations du Vivant (AVIV), Muséum National d'Histoire Naturelle, CNRS, CP52, 57 rue Cuvier 75231 Paris Cedex 05, France., Tel: +33 1 40793547, email: [isabelle.florent@mnhn.fr](mailto:isabelle.florent@mnhn.fr)

Key words: Apicomplexa, Evolutionary history, Genomics, Parasitology, Protozoa

List of abbreviations: SEM: Scanning Electron Microscopy; TEM: Transmission Electron Microscopy

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/boc.202000006](https://doi.org/10.1111/boc.202000006).

This article is protected by copyright. All rights reserved.

**Abstract.**

Gregarines, a polyphyletic group of apicomplexan parasites infecting mostly non-vertebrates hosts, remains poorly known at taxonomic, phylogenetic and genomic levels. However, it represents an essential group for understanding evolutionary history and adaptive capacities of apicomplexan parasites to the remarkable diversity of their hosts. Because they have a mostly extracellular lifestyle, gregarines have developed other cellular developmental forms and host-parasite interactions, compared to their much better studied apicomplexan cousins, intracellular parasites of vertebrates (Hemosporidia, Coccidia, Cryptosporidia). This review highlights the promises offered by the molecular exploration of gregarines, that have been until now left on the side of the road of the comparative –omic exploration of apicomplexan parasites. Elucidating molecular bases for both their ultrastructural, functional and behavioral similarities and differences, compared to those of the typical apicomplexan models, is expected to provide entirely novel clues on the adaptive capacities developed by Apicomplexa over evolution. A challenge remains to identify which gregarines should be explored in priority, as recent metadata from open and host-associated environments has confirmed how underestimated is our current view on true gregarine biodiversity. It is now time to turn to gregarines to widen the currently highly skewed view we have of adaptive mechanisms developed by Apicomplexa.

## 1. Introduction

Apicomplexa are unicellular eukaryotes (protists) collectively corresponding to ~350 genera and ~6000 named species, the wide majority of which have adopted a strict parasitic lifestyle in a very wide diversity of metazoan hosts (Adl et al., 2018; Portman and Slapeta, 2014). Apicomplexa, together with the two sister groups Dinoflagellata and Ciliata, form the supergroup Alveolata, at the base of which are Rhizaria and Stramenopiles / Phaeophyta (Adl et al., 2018). Apicomplexa are mostly known for comprising infamous intracellular parasites of vertebrates responsible for important human diseases such as malaria due to *Plasmodium* spp., cryptosporidiosis due to *Cryptosporidium* spp. and toxoplasmosis due to *Toxoplasma gondii*. Apicomplexa also comprise diverse other intracellular parasites of vertebrates, with economical or veterinary importance such as *Eimeria* spp., *Babesia* spp. and *Theileria* spp. These apicomplexan parasites have simple to very complex life cycles. Some are restricted to single hosts (monoxenous parasites, e.g. *Cryptosporidium*, *Eimeria*), other alternate between two successive hosts (dixenous parasites such as *Plasmodium*, *Babesia* and *Theileria*, completing sexual reproduction in various insects or arthropods and asexual phases in various tissues of vertebrates) and few have the capacity to infect multiple hosts (polyxenous parasites such as *Toxoplasma*, completing its sexual reproduction in cats and several asexual phases in various tissues of diverse warm blooded vertebrates). Due to their medical, veterinary or economic importance, and because it has been possible to cultivate most of them in laboratory conditions, their genomes have been deciphered (~100, deposited into the EupathDB database (Aurrecochea et al., 2017), Figure 1) and are major references for medical investigations, comparative genomics studies and exploration of evolutionary history of apicomplexan parasites (Janouskovec et al., 2019; Janouskovec et al., 2015; Kwong et al., 2019; Woo et al., 2015).

But Apicomplexa also comprise another group of organisms collectively known as “gregarines”, that are principally monoxenous parasites of a wide diversity of non-vertebrate metazoan hosts, ranging from Polychatea annelids to tunicates, arthropods and mollusks, in which they develop mostly extracellularly - contrary to the above mentioned parasites - in the intestinal and coelomic cavities of their hosts (Desportes I and Schrével J, 2013). These endoparasites are mostly considered as being nonpathogenic, with a few reported cases of recognized pathogenicity, while it is clear that experimental studies focusing on gregarine pathogenicity are mostly lacking (Rueckert et al., 2019a).

## 2. Gregarines: well described Apicomplexa though forgotten at molecular level.

There are several consequences to the mostly extracellular lifestyle displayed by gregarines, a feature that indeed distinguishes them from their intracellular parasites relatives (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). First, gregarines can reach very large sizes even for unicellular eukaryotes, from less than a  $\mu\text{m}$  to more than a mm for respectively the zoite and trophozoite forms of the marine eugregarine *Porospora gigantea*, intestinal parasite of the lobster *Homarus gammarus* (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). Insect eugregarines, such as *Gregarina garnhami*, intestinal parasite of the desert locust *Schistocerca gregaria*, display trophozoite forms reaching a dozen to several hundred  $\mu\text{m}$  (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). Archigregarine trophozoites, such as *Selenidium pendula*, intestinal parasite of the polychaete *Scolecopsis (Nerine) squamata* also reach several dozen  $\mu\text{m}$  (Schrevel et al., 2016). The large sizes of these developmental stages and their very common occurrence in a large diversity of non-vertebrate metazoan hosts have facilitated the discovery and

biological studies on gregarines, resulting in an abundant literature on their morphologies, ultrastructures and life cycles, that have been examined by photonic and electron microscopy (SEM, TEM) imaging or by using cryoelectron microscopy or immunofluorescence (see (Desportes I and Schrével J, 2013) for exhaustive bibliography). Dynamic recordings are also available notably concerning their movements or behaviors (Desportes I and Schrével J, 2013). This rich literature offers a wide panorama of the adaptive capacities of these organisms to their hosts and environments, awaiting now exploration by -omic approaches to decipher the molecular bases of their functioning and variations, as it has been developed so far for their intracellular apicomplexan cousins (Rueckert et al., 2019b). Indeed, the genomic knowledge on apicomplexan is currently highly biased in favor of intracellular parasites of vertebrate hosts, belonging to Haematozoa, Coccidia and to a lesser extent, Cryptosporidia (Figure 1). Gregarines have been so far mostly excluded from this –omic exploration, to the exception of (unpublished) genome of the terrestrial insect eugregarine *Gregarina niphandrodes* (accessible at EupathDB (Aurrecochea et al., 2017), section CryptoDB), intestinal parasite of *Tenebrio molitor*, very partial genomic data on insect eugregarine *Ascogregarina taiwanensis* (Templeton et al., 2010) intestinal parasite of *Aedes albopictus* and partial and recently emerging transcriptomic data for a dozen of terrestrial and marine gregarine species (Janouskovec et al., 2019; Mathur et al., 2019; Omoto et al., 2004) (see Figure 1 for illustration on Apicomplexa genomic data knowledge).

There are several reasons why the acquisition of -omic knowledge on gregarines is lagging behind that of their intracellular vertebrate parasite cousins. (1) As there are infecting “only” non-vertebrates hosts and are mostly considered non-pathogenic (Rueckert et al., 2019a), they have been neglected. (2) The current lack of *in vitro* culture methods for these parasites complicates the isolation of biological material in adequate amounts and quality for accurate usage in molecular investigations. Indeed, gregarine biological studies must rely on field collections, mostly from infected hosts (alternatively their feces), which exposes the collected material to contaminations by host cells and environmental microorganisms. The ability to maintain the hosts of a diversity of gregarines in laboratory conditions, offers a good compromise as it allows regular access to different developmental stages amenable to a variety of cellular (microscopy, test of inhibitors) and molecular (-omics) studies (see (Desportes I and Schrével J, 2013) for exhaustive descriptions). For example, *Neanthes (Nereis) divesicolor*, the host of the marine eugregarine *Lecudina tuzetae*, can be maintained for several months in natural or artificial sea water in the laboratory (Desportes I and Schrével J, 2013; Kuriyama et al., 2005). Also, several insect raising facilities may be used to get access to their infecting gregarines (Clopton, 2009; Desportes I and Schrével J, 2013). Also, it can be expected that the concomitant progression of -omics and microscopic methodologies, allowing using increasingly reduced amounts of biological material, will also facilitate in a near future the bridging of this molecular knowledge gap, between the currently very poorly documented non cultivable gregarines and the increasingly well documented cultivable intracellular parasites of vertebrates (Gawad et al., 2016).

### 3. Why should we study gregarines?

By why should we focus on gregarines at molecular level? What could they tell us that we do not already know about apicomplexan parasites?

Because gregarines are Apicomplexa, although particular ones, with unique differences notably their mostly extracellular life mode and its biological consequences (Adl et al., 2018; Desportes I and Schrével J, 2013; Rueckert et al., 2019a; Schrevel and Desportes, 2015). Like all Apicomplexa, gregarines present at least once during their life cycle a development form called zoite, a polarized cell comprising the so-called “apical complex” composed of scaffolding cytoskeletal elements enclosing specialized apical organelles (rhoptries, micronemes, dense granules) that gave the name to this phylum replacing the former Sporozoa (Adl et al., 2018; Morrissette and Sibley, 2002; Tardieux and Baum, 2016). Gregarines (notably, Archigregarines) do have a conoid, composed of spirally arranged array of microtubules as found in Coccidia and Cryptosporidia (forming Conoidasida), but secondarily lost in Haemosporidia (forming Aconoidasida) ((Adl et al., 2018; Portman and Slapeta, 2014), Figure 1). In Apicomplexa having intracellular development phase(s) this “apical complex” has been clearly involved in the recognition and invasion of host cells, allowing parasites establishment and development in this novel ecological niche (Hakimi et al., 2017; Tardieux and Baum, 2016). In Apicomplexa displaying extracellular lifestyle as most intestinal gregarines do, this “apical complex” appears to have a different role. Yet involved in parasite attachment to host cells at sporozoite stage, it subsequently appears mostly used for parasite feeding, sustaining spectacular growth phases, and not for achieving tissue penetration or parasite internalization within host cells (Gentekaki et al., 2014; Schrevel et al., 2016; Valigurova and Koudela, 2008). This to the notable exception of coelomic (eu)gregarines capable of intestinal barrier crossing and neogregarines, capable of reaching intracellular niches in some of their hosts tissues (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015).

There are several consequences to this extracellular growth of these trophozoite forms in gregarines: 1) an extremely wide diversity of shapes and sizes, as mentioned above and largely used for taxonomical purposes (Desportes I and Schrével J, 2013); 2) a sexual phase (gamogony followed by sporogony) that also occurs extracellularly, producing developmental forms that are particular to gregarines, starting with the syzygy (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015).

Syzygy is the named given to the development stage that precludes the gregarine sexual reproduction. It corresponds to the bi-association of two trophozoites after they have detached from host cell and which are future gamonts (Figures 2, stages (a) and (b/b')). Although morphologically similar both in size and shape, the two partners of the syzygy are committed to evolve into respectively male and female gamonts (Figures 2, stages (c)). Depending on the species, the bi-association may be caudo-caudal (Figure 2A, stage (c)), lateral (Figure 2B, stage (c)), fronto-frontal (not shown, see (Desportes I and Schrével J, 2013)) or caudo-frontal (Figure 2C, stage (c)). In this latter case, found in *Gregarina garnhami*, the anterior partner of the syzygy is called primate the other satellite (Figure 2C, stage (c)). Occasionally, syzygy associations may involve more than two partners but the evolution of such “ménage à trois” has not been yet examined at molecular level (Desportes I and Schrével J, 2013). The evolution of the syzygy bi-association in a globular structure called gametocyst, initially composed of two hemispheres of similar shape and volume around which a thick wall is elaborated (Figures 2, stages (c)). Series of nuclear divisions with final cytokinesis

(gamogony) then occur within each hemisphere producing male gametes within one hemisphere and female gametes within the other one (Figures 2, stages (d)). A clear anisogamy is commonly observed between male gametes - more pyriform and usually flagellated – and female gametes - more globular and non-flagellated (Figures 2A 2B, stages (d')). It is therefore only after complete gametes production that the “sex” of gamonts may be deduced. Numerous imaging recordings have been performed to study the cellular events occurring during this first phase of the gregarine sexual reproduction called gamogony, with a remarkable confocal imaging analysis performed in the case of the marine eugregarine *Lecudina tuzetae*, intestinal parasite of the Polychaeta *Hediste (Nereis) diversicolor*, in which ~5000 male and as many female gametes are produced per gametocyst (Kuriyama et al., 2005).

It is interesting to notice that, depending on the species, the length of the syzygy phase might be particularly long; in the atypical case of *Diplauxis hattii*, coelomic parasite of the Polychaeta *Perinereis cultrifera*, this bi-association remains stable for more than two years, awaiting host's sexual maturation to engage into gamogony (Prensier et al., 2008).

Once the male and female gametes have been produced within their respective compartments, their mixing occurs within the gametocyst. Thousands of fertilizations then take place simultaneously during a process called "gamete dance" in *L. tuzetae*, which lasts ~4 hours and produces ~5000 zygotes per gametocyst (Kuriyama et al., 2005) (Figure 2B stage (e)). Sporogony then begins and the gametocyst takes the name of sporokyst (Figures 2B and 2C, stages (e/e')). Each zygote secretes a cyst wall (stage called immature sporocyst) and undergoes meiosis and additional mitosis (in eugregarines and neogregarines) leading to sporozoites formation (Figures 2, stages (f) to (i)). Each spore, also called sporocyst, therefore possesses a thick wall and is the form of dissemination of the gregarine in the environments (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). It contains 4 sporozoites in the case of archigregarines and 8 sporozoites in eugregarines and neogregarines ((Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015), see also Figures 2 stages (i)). In Coccidia, sporozoites are also formed within a spore that however is called oocyst while in gregarines it is called sporocyst (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). In marine eugregarines, spores are eventually released into the environments with the breaking down of the sporokyst (Figure 2B stage (e) in the case of *L. tuzetae*). In terrestrial gregarines, spores are released in the environment *via* sporoducts that are formed at the surface of the sporokyst (see for example *Gregarina garnhami* Figure 2B stage (e')). The progeny in terms of number of spores, resulting from the evolution of a single syzygy, is considerable and can reach few thousand to several million depending on the gregarine species (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). Sporokysts are therefore obviously a material of choice to isolate genomic DNA. Sporocysts on the other hand, are the developmental forms that are likely collected within soils and sediments from environments (see below).

Once ingested by hosts, the spores undergo dehiscence after passage through the host's digestive system and sporozoites are released (4 in the case of archigregarines, 8 in the case of eugregarines and neogregarines, Figures 2 stages (j)). These will be able to attach to their host's intestinal cells (Figure 2C, (j')), using their apical end displaying typical “apical complex” features, and these attached sporozoites will start to grow dramatically, evolving into trophozoites. In the case of *G. garnhami* for example, trophozoites grow from less than 10µm to over 400µm in length within a single host. The size range of trophozoites may therefore cover 2 to 3 orders of magnitude

depending on the species (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). This gigantism, achieved for a large number of gregarine species, offers a remarkable material for cell biology explorations and immunofluorescence imaging for example (Kuriyama et al., 2005; Valigurova et al., 2013).

This radical difference in lifestyle including gamogonic and sporogonic phases, contrasting from those observed in Hematozoa and Coccidia, results also in very different interactions between gregarines and their hosts, which raises many questions about the adaptive pathways they have developed, as apicomplexan parasites, to survive over the course of evolution. Which molecular solution have they then developed to survive within the host's intestinal tracts or other cavities, maintain survival, acquire nutriment, complete (a)sexual reproduction, with a remarkable success if one considers their wide occurrence in such a high diversity of endoparasitic non-vertebrate hosts contexts? (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015)

In order to widen our view on host-adaptive modes in Apicomplexa, we propose thereafter some possible exploration topics focusing on gregarines, which are obviously not exhaustive as the number of biological questions that may be asked is far more important.

**3.1. Gregarines, as apicomplexan parasites, do possess a fully developed apical complex**, at least in some developmental stages (sporozoites, trophozoites) (Desportes I and Schrével J, 2013). Biological and morphological studies have established that in gregarines, this apical complex is used for host cell attachment to allow the parasite to feed from its host cell by a process known as myzocytosis (Schrevel et al., 2016; Simdyanov and Kuvardina, 2007). The host cell penetration by the parasite is not complete as the gregarine remains extracellular with only its apical end intimately engaged in a host-parasite interplay that has been studied at microscopic level but whose molecular actors are poorly defined (Schrevel et al., 2016; Simdyanov et al., 2017; Valigurova et al., 2007). See also (Desportes I and Schrével J, 2013) for exhaustive descriptions of several additional examples. Therefore, the biological function of "gregarine apical complex" only partly overlaps the biological function currently attributed to "apicomplexan parasite apical complex" i.e. recognition and invasion of their host cell (Tardieux and Baum, 2016).

Several questions therefore emerge: to which extent does the molecular architecture of "gregarine apical complexes" compares to that of their best known cousins, i.e. *Toxoplasma* and *Plasmodium* (Boucher and Bosch, 2015; Tardieux and Baum, 2016)? Are the "scaffolding", "recognition", "invasive" and "nutrition" functions fulfilled by the same molecular components or by other ones? What are the composition and functional roles of micronemes, rhoptries and dense granules in gregarines? In intracellular apicomplexan parasites of vertebrates (*Plasmodium* and *Toxoplasma*), these secretory organelles are documented to intervene sequentially in an orchestrated manner, with first micronemes secreting parasite proteins involved in host cell recognition proteins as well as AMA-1, which, when combined to a defined cortege of RON proteins (secreted by the neck of rhoptries) will assemble in the so-called mobile junction essential to the invasion process of host cells by *Plasmodium* merozoites or *Toxoplasma* tachyzoites (reviewed in (Tardieux and Baum, 2016)). Subsequently, ROP proteins (secreted by the bulb of rhoptries) and GRA proteins (secreted by the dense granules) will allow establishing the intracellular niche for these parasites, either at the parasitophorous vacuole level (facilitating metabolite exchanges) or beyond this border to manipulate the host cell program to the benefit of the parasite (Hakimi et al., 2017). Although it may



be expected that gregarine micronemes, rhoptries and dense granules will have their own protein repertoires (there is currently very limited overlaps of these repertoires between currently described apicomplexan genus (Boucher and Bosch, 2015; Counihan et al., 2013; Hakimi et al., 2017)) it will be interesting to decipher their specific roles and how they contribute (or not) to establishing also the host-parasite interface. It is important here to indicate that there are alternative invasive modes in apicomplexan parasites such as *Theileria* and *Cryptosporidium* that differ from the better described *Toxoplasma/Plasmodium* mode (Gubbels and Duraisingh, 2012). What are the molecular similarities between the *T. gondii* and *P. falciparum* parasitophorous make up and the food vacuole of gregarines, that forms at the gregarine-host cell interface (Schrevel et al., 2016; Simdyanov et al., 2017; Valigurova et al., 2007)? Or is the similarity stronger to the feeder organelle of epicellular *Cryptosporidium*? (Barta and Thompson, 2006; Bartosova-Sojkova et al., 2015).

**3.2. Their mostly extracellular development.** A first morphological and biological consequence of this particular behavior selected over evolution is the fact mentioned above that gregarine trophozoites can reach very large sizes contrary to intracellular parasites of vertebrates (1-10 $\mu$ m at most). A second is that their sexual phase is also extracellular, starting with the syzygy that evolves into gametocysts then sporokysts producing sporocysts (Figure 2). These developmental forms are strikingly distinct from the developmental forms encountered in *Toxoplasma*, *Plasmodium* and even *Cryptosporidium* (Aly et al., 2009; Bouzid et al., 2013; Robert-Gangneux and Darde, 2012). Interestingly, the oocysts forms of *Toxoplasma* and *Cryptosporidium*, which are also extracellular and disseminated with their hosts' feces as resistant forms in the environments, are however elaborated intracellularly, within their hosts' intestinal cells (Bouzid et al., 2013; Robert-Gangneux and Darde, 2012). Indeed, in *Toxoplasma* and *Cryptosporidium* the gamogony remains intracellular, whereas it is extracellular in gregarines. An important consequence is that gregarines thus display totally different types of host-parasite interactions, having other constraints to face such as surviving in host-gut environment. Several studies have explored the permeability of the trophozoite membrane in link with the question of their nutrition mode (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). This questions the molecular nature and the biological role of their inner membrane complex (imc), which, interestingly, may be acting as a continuous "shield" all around the trophozoite as it is only interrupted at the conoidal opening through which nutrition occurs (see (Schrevel et al., 2016) for the case of *S. pendula* or (Kuriyama et al., 2005) for the case of *L. tuzetae*). Interestingly, as the syzygy evolves towards the gametocyst form, this imc appears to be disassembled concomitantly with the secretion, by the gametocyst, of the protective cyst wall (see (Kuriyama et al., 2005) for details on *L. tuzetae*). This suggests that in gregarines, one form of shielding (imc) in trophozoites and syzygies gives place to another form of shielding (cyst wall) during gametogony then sporogony, both of which being intended to isolate the gregarine from its hostile (gut) environment. Obviously, the molecular exploration of such a parasite-environment (host) interplay will reveal novel adaptive features developed by Apicomplexa over evolution. To which extend the host-gut environment is less hostile in an invertebrate host rather than in a vertebrate one, notably regarding immune system response and microbiota regulation, should certainly be taken into account. First, invertebrate hosts rely mainly on innate immunity to fight intruders while vertebrates also have adaptive defense mechanisms (Buchmann, 2014). In addition, the co-existing microbiome is notoriously less complex and diverse in invertebrates than vertebrates (Bahrndorff et al., 2016). These differences could explain the capacity of gregarines to self- maintain in such host-gut environment for an extended life cycle times while parasites of vertebrates have been

constrained to invade host cells to achieve their maintenance in hosts. Further studying not only gregarines but also their host's immune and microbiota responses will certainly clarify the contribution of these host-specific features to the diversity of gregarines behaviors and life traits over evolution.

**3.3. Gregarines have developed a wider diversity of motility and mobility modes** than what is mostly described (and deeply studied at molecular level) for intracellular parasites of vertebrates: the gliding motility (Frenal et al., 2017). This movement, governed by an acto-myosin motor, involves dozens of proteins that appear (so far) well conserved between apicomplexan species (Boucher and Bosch, 2015; Frenal et al., 2017; Mueller et al., 2017). Whether the gliding components are conserved also in gregarines that move by gliding (i.e., most eugregarines) remains to be established. However, gregarines do display other modes of motility such as rolling, bending (notably archigregarines (Desportes I and Schrével J, 2013)), the molecular bases of which are currently totally unknown. Do these alternative modes involve molecular components shared with those of the glideosome? Do they involve other components, inherited from the putative ancestor and possibly lost secondarily in intracellular parasites of vertebrate (Füßy and Oborník, 2017; Janouskovec et al., 2015; Woo et al., 2015)? Or do they involve novel components, repositioned from the ancestor heritage or acquired by horizontal gene transfer? All this remains to be established for the diversity of known and to-be-discovered gregarines. Answers to these questions will be precious to understand how such a diversity of motility/mobility modes may have emerged for apicomplexan from a common ancestral genetic heritage, as apicomplexan derive from ancestral biflagellated organisms with repositioning of some of the former flagellar components into the apical complex structure and functioning (Füßy and Oborník, 2017; Janouskovec et al., 2015; Woo et al., 2015). This knowledge will also be precious to widen our current knowledge of the adaptive capacities to hosts developed by these remarkable apicomplexan parasites. Indeed, evolutionary molecular studies on this point have established that gliding components partially existed in the common ancestor of Apicomplexa (Füßy and Oborník, 2017; Janouskovec et al., 2015; Woo et al., 2015) but have been repositioned to be functionally operational in intracellular parasites of vertebrates. What paths of specialization did they then follow to generate such a diversity of movements in gregarines? And, corollary to these observations, can we formulate the hypothesis that the lack of host-cell fully invasive capacities of gregarines may be due either to absence of gliding capacities despite a developed / expected to be functional apical complex (in the case of archigregarines) or, conversely, to an under-developed apical complex despite operational gliding capacities (in the case of eugregarines)?

It is therefore now time to turn to gregarines to explore these adaptive traits at molecular level, but the next question is; which ones to select and from which extent of diversity? Indeed, a recent convergence of novel data clearly indicates that the current inventory of the true gregarine diversity is dramatically underestimated – and therefore, corollary, all the relevant biological models may have not yet been discovered.

#### **4. The true gregarine biodiversity.**

Our current understanding of gregarine biodiversity comes from three sources of information that only partly overlap : 1) the number of formally inventoried species (Portman and Slapeta, 2014), 2) the number of species theoretically computable based on the inventory and diversity of their hosts

(Desportes I and Schrével J, 2013); 3) environmental or host-associated metagenomic or metabarcoding approaches, that have revealed novel molecular signatures, sufficiently related to gregarines to allow taxonomic affiliation to this group but sufficiently divergent to strongly suggest novel taxonomic species (de Vargas et al., 2015; Del Campo et al., 2019; Mahe et al., 2017). The regular cross-referencing of these three sources of data leads to a permanent readjustment of both the taxonomy and phylogeny of these species, so that it is safe to say that the current biodiversity of gregarines is a field of investigation under construction, whose physiognomy is likely to evolve considerably in the coming years.

#### 4.1. Formally described species.

Regarding the first point, there are currently ~1770 formally described gregarine species, unequally distributed between archigregarines (~20), eugregarines (~1700) and neogregarines (~50) (Portman and Slapeta, 2014). In parallel, taxonomic and phylogenetic revisions concerning gregarines is a currently very active - but quite unstable - field with a diversity of successive proposals regarding their phylogenetic inter-relations as well as with other apicomplexan parasites (Cavalier-Smith, 2014; Rueckert and Horak, 2017; Schrevel et al., 2016; Simdyanov et al., 2017; Simdyanov et al., 2018). Molecular phylogenies are nowadays mostly based on usage of SSU rDNA marker, more rarely complete ribosomal loci (Diakin et al., 2016; Diakin et al., 2017; Simdyanov et al., 2018). Studies based on the 18S SSU rDNA marker alone are fairly effective in defining monophyletic groups at the genus or family levels, but fail to robustly resolve the respective branches' relationships at higher taxonomic level. Attempts to improve phylogenies using the full ribosomal marker (18S SSU + 28S LSU rDNA) have provided some progress, but have the important disadvantage to be currently available for only very few gregarine species (~20) (Diakin et al., 2016; Diakin et al., 2017; Simdyanov et al., 2018). Phylogenies relying on multiple genes (or more accurately proteins) sequences derived from genome / transcriptome investigations are now emerging but remain restrained to a dozen of gregarine species (Janouskovec et al., 2019; Mathur et al., 2019). Although they are indeed expected to be more resolving, the number of concerned species is even smaller and ambiguities remain in the interrelationships between groups, since the position of the genus *Cryptosporidium* is for example unstable between the two studies (Janouskovec et al., 2019; Mathur et al., 2019). These genome/transcriptome studies have however also shown another interest: some species historically described as gregarines (*Platyproteum* sp. and *Digyalum oweni*) do not actually appear anymore to be part of Apicomplexa (Janouskovec et al., 2019; Mathur et al., 2019).

#### 4.2. Extrapolation based on host diversity and host-parasite behaviors.

Concerning the second point, it is clearly documented that gregarines parasite virtually all non-vertebrate metazoan groups, from Polychatea annelids to tunicates, arthropods and mollusks (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). For a long time, experts in the gregarine field have argued that, given the currently documented diversity of gregarines, their lifestyle principally monoxenous and their apparently narrow host-range specificity, the real biodiversity of gregarines is probably several orders of magnitude underestimated in particular for those infecting insects (~half of metazoan diversity) (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). In addition, most known hosts are infected by several gregarine species, as for example the mealworm *Tenebrio molitor* that is infected by at least 3 gregarine species (*Gregarina cuneata*, *Gregarina polymorpha* and *Gregarina steini*) (Clopton et al., 1992) and also *G.*

*niphandrodes*. In consequences, some experts have even proposed that the real gregarine diversity should exceed that of their hosts, making it one of the most widespread groups of organisms in the environments (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015).

#### **4.3. Environmental metadata uncovering a wide and mostly undescribed diversity.**

Thirdly, several environmental and host-associated metagenomics/metabarcoding surveys have recently not only confirmed the gregarine experts' predictions: they have started to document this tremendous diversity at molecular level (de Vargas et al., 2015; Del Campo et al., 2019; Mahe et al., 2017). A very wide diversity of apicomplexan-related sequences (~80% of which appears more closely related to gregarines) is present in terrestrial soils and marine sediments (Clopton et al., 1992; Del Campo et al., 2019; Mahe et al., 2017). This diversity may be paralleled to the biological cycles of gregarines for which cyst forms, enclosing each thousand to millions of sporocyst progenies, are frequently released with the feces of their infected hosts, easily contaminating therefore such soils. In addition, the high resistance of these sporocyst forms to environmental conditions certainly explains the high abundance and maintenance (in environments) of their enclosed DNA/RNA (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). However, and even more remarkably, the diversity of apicomplexan-related sequences with gregarine-like affinities is also high in marine environments (both pelagic and benthic), suggesting the presence of either developing or resistant (sporocyst) forms of these gregarines in association with planktonic elements - biological or even mineral - that remain to be identified or “freely” floating in these marine environments (de Vargas et al., 2015). However, the flip side of these discoveries is that there is a remarkably high number of gregarine-like molecular data that cannot be related to formally described species (Del Campo et al., 2019). This is in part due to the very low molecular knowledge we currently have of taxonomically and morphologically described species. Only talking about the most commonly used molecular marker, SSU rDNA, it is available in databases for only about one hundred of the ~1770 formally described gregarine species, which corresponds to just ~5% of them, not taking into account the many gregarine species to be still discovered. So not only to we need to generate a much higher number of molecular markers for these known species, but we must also deploy appropriate strategies to morphologically and biologically describe the increasing number of “molecular-species” the existence of which is clearly emerging from metagenomics studies, pointing to entirely novel phylogenetic groups within gregarines (de Vargas et al., 2015; Del Campo et al., 2019; Mahe et al., 2017).

#### **4.4. Missing data blurs apicomplexan phylogeny**

The missing information regarding gregarines is not restricted to this lack of connection between taxonomic data derived from morphological studies and molecular data emerging from metagenomics approaches. The confrontation of these data indicates that our understanding of the apicomplexan biodiversity remains very limited, biased because we have too long neglected the large spectrum of their pathogenicity focusing preferentially on the deadliest ones and ignoring the vast majority of poorly pathogenic ones (Rueckert et al., 2019a). This biased position has probably blurred not only our comprehension of the extent of their extraordinary host-adaptive capacities, but altogether a full section of their evolutionary history. So that we still do not know where the emergence of these species is taking root; is it within a group of non-pathogenic symbionts that has become even more diversified than we imagine, advocating for multiple emergence within a

radiation that is still incompletely understood (Janouskovec et al., 2019; Kwong et al., 2019; Mathur et al., 2019; Rueckert et al., 2019b)?

A close examination of the recent taxonomy of Apicomplexa clearly mentions that most groups are currently polyphyletic or paraphyletic, notably: Aconoidasida, Coccidia, Gregarinasina, Archigregarinorida and Eugregarinorida (Adl et al., 2018). This is mostly due to lack of sufficiently informative and resolving data concerning these organisms, still mostly based on SSU rRNA phylogenies (see (Cavalier-Smith, 2014; Simdyanov et al., 2017; Simdyanov et al., 2018) to give few examples) and recent, emerging, phylogenomic analyses (Janouskovec et al., 2019; Kwong et al., 2019; Mathur et al., 2019). Certainly, insufficient sampling also prevents a solid and integrated vision of phylogenetic inter-relationships for all of these species (Del Campo et al., 2019; Janouskovec et al., 2019; Kwong et al., 2019; Mathur et al., 2019).

## 5. Conclusion.

With just over 100 genomes deciphered for ~350 genera and ~6000 described species, Apicomplexa is a group for which there is still a lot to discover even if, it is not the poorest documented branch of the tree of life (Sibbald and Archibald, 2017). It is far from having revealed all the secrets of the diversity of its molecular innovations, developed during its evolutionary history as it had to adapt to such a wide diversity of hosts and environments (Desportes I and Schrével J, 2013; Schrevel and Desportes, 2015). As to date, molecular exploration of apicomplexan parasites have mainly concerned a very small number of species that have in common (1) to infect humans causing threatening and global diseases such as malaria or less threatening but worldwide spread diseases such as toxoplasmosis and cryptosporidiosis; (2) to be cultivable in the laboratory, at least for some developmental stages; (3) to have been the subject of extremely sophisticated methodological developments such as genetic manipulation, -omics in all their variations and static and dynamic microscopy.

In this panorama, gregarines, full members of the Apicomplexa phylum, have been so far left on the side of the road for exactly corollary reasons: (1) they do not infect humans, (2) they are not easy to cultivate and (3) while there is a very abundant literature of their life cycles, morphologies and ultrastructure, they are almost unknown at the genomic/transcriptomic levels and have been the subject of very few biochemical studies. But their future is now open for exploration.

The terrain is marked out for the emergence of genomic data. A first move should be in favor of known species, selected either for their biological characteristics (intestinal, coelomic, motile, non-motile) or their particular phylogenetic position (archigregarines vs eugregarine as a broad distinction). They have the potential to teach us novel adaptive aspect of the group (relative to intracellular parasites of vertebrates), particularly in relation to their extracellular mode of living and the associated specific constraints to survival capacities. Such exploration should also reveal a part of the ancestral heritage, now possibly lost in the other branches of Apicomplexa. Studying gregarines can also allow understanding finer adaptive mechanisms, for example regarding the functionalization of the apical complex, which obviously, compared to intracellular parasites of vertebrates displays common properties (recognition) with additional functions (nutrition) and missing ones (full invasion), the molecular bases of which are to be discovered. But it cannot be overlooked that the currently documented diversity of gregarines is lagging, by probably several orders of magnitude, far beyond the true diversity of these organisms in open and host-associated

environments (Del Campo et al., 2019; Desportes I and Schrével J, 2013; Mahe et al., 2017; Mathur et al., 2019; Schrevel and Desportes, 2015). This diversity is also likely, once it is explored and named (in terms of species and lifestyle) to reveal even more original and unknown adaptive mechanisms for this fascinating group, the Apicomplexa.

**6. Acknowledgement.** We deeply thank Professor J. Schrével, Emeritus Professor, UMR7245 CNRS MNHN for careful reading and editing of text and figures as well as valuable suggestions. The authors declare they have no conflict of interest

**7. Funding.** This work was supported by a grant from Agence Nationale de la Recherche [LabEx ANR-10-LABX-0003-BCDiv], in the program “Investissements d’avenir” [ANR-11-IDEX-0004-02], as well as several interdisciplinary Programs of the MNHN (ATM-Microorganismes, ATM-Génomique et Collections, ATM-Emergence, AVIV department), and by CNRS (PhD fellowship to Julie Boisard, 2018-2021).

**8. Authorship.** IF conducted the study, JB elaborated all illustrations, IF and JB wrote the manuscript.

## 9. References.

- Adl, S.M., D. Bass, C.E. Lane, J. Lukes, C.L. Schoch, A. Smirnov, S. Agatha, C. Berney, M.W. Brown, F. Burki, P. Cardenas, I. Cepicka, L. Chistyakova, J. Del Campo, M. Dunthorn, B. Edvardsen, Y. Eglit, L. Guillou, V. Hampl, A.A. Heiss, M. Hoppenrath, T.Y. James, S. Karpov, E. Kim, M. Kolisko, A. Kudryavtsev, D.J.G. Lahr, E. Lara, L. Le Gall, D.H. Lynn, D.G. Mann, I.M.R. Massana, E.A.D. Mitchell, C. Morrow, J.S. Park, J.W. Pawlowski, M.J. Powell, D.J. Richter, S. Rueckert, L. Shadwick, S. Shimano, F.W. Spiegel, I.C.G. Torruella, N. Youssef, V. Zlatogursky, and Q. Zhang. 2018. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *The Journal of eukaryotic microbiology*.
- Aly, A.S., A.M. Vaughan, and S.H. Kappe. 2009. Malaria parasite development in the mosquito and infection of the mammalian host. *Annual review of microbiology*. 63:195-221.
- Aurrecochea, C., A. Barreto, E.Y. Basenko, J. Brestelli, B.P. Brunk, S. Cade, K. Crouch, R. Doherty, D. Falke, S. Fischer, B. Gajria, O.S. Harb, M. Heiges, C. Hertz-Fowler, S. Hu, J. Iodice, J.C. Kissinger, C. Lawrence, W. Li, D.F. Pinney, J.A. Pulman, D.S. Roos, A. Shanmugasundram, F. Silva-Franco, S. Steinbiss, C.J. Stoeckert, Jr., D. Spruill, H. Wang, S. Warrenfeltz, and J. Zheng. 2017. EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic acids research*. 45:D581-D591.
- Bahrndorff, S., T. Alemu, T. Alemneh, and J. Lund Nielsen. 2016. The Microbiome of Animals: Implications for Conservation Biology. *International journal of genomics*. 2016:5304028.
- Barta, J.R., and R.C. Thompson. 2006. What is Cryptosporidium? Reappraising its biology and phylogenetic affinities. *Trends in parasitology*. 22:463-468.
- Bartosova-Sojkova, P., R.D. Oppenheim, D. Soldati-Favre, and J. Lukes. 2015. Apicomplexans: Parasites "On the Way In". *PLoS pathogens*. 11:e1005080.
- Boucher, L.E., and J. Bosch. 2015. The apicomplexan glideosome and adhesins - Structures and function. *Journal of structural biology*. 190:93-114.
- Bouzig, M., P.R. Hunter, R.M. Chalmers, and K.M. Tyler. 2013. Cryptosporidium pathogenicity and virulence. *Clinical microbiology reviews*. 26:115-134.
- Buchmann, K. 2014. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Frontiers in immunology*. 5:459.
- Canning, E.U. 1956. A New Eugregarine of Locusts, *Gregarina garnhami* n.sp., parasitic in *Schistocerca gregaria* Forsk. *Journal of Protozoology*. 3:50-62.
- Cavalier-Smith, T. 2014. Gregarine site-heterogeneous 18S rDNA trees, revision of gregarine higher classification, and the evolutionary diversification of Sporozoa. *European journal of protistology*. 50:472-495.
- Clopton, R.E. 2009. Phylogenetic Relationships, Evolution, and Systematic Revision of the septate gregarines (Apicomplexa:Eugregarinorida:Septatorina). *Comp Parasitol*. 76:167-190.
- Clopton, R.E., J. Janovy, Jr., and T.J. Percival. 1992. Host stadium specificity in the gregarine assemblage parasitizing *Tenebrio molitor*. *The Journal of parasitology*. 78:334-337.
- Counihan, N.A., M. Kalanon, R.L. Coppel, and T.F. de Koning-Ward. 2013. Plasmodium rhostry proteins: why order is important. *Trends in parasitology*. 29:228-236.
- de Vargas, C., S. Audic, N. Henry, J. Decelle, F. Mahe, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horak, O. Jaillon, G. Lima-Mendez, J. Lukes, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S.G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M.E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, and E. Karsenti. 2015. Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science*. 348:1261605.
- Del Campo, J., T.J. Heger, R. Rodriguez-Martinez, A.Z. Worden, T.A. Richards, R. Massana, and P.J. Keeling. 2019. Assessing the Diversity and Distribution of Apicomplexans in Host and Free-

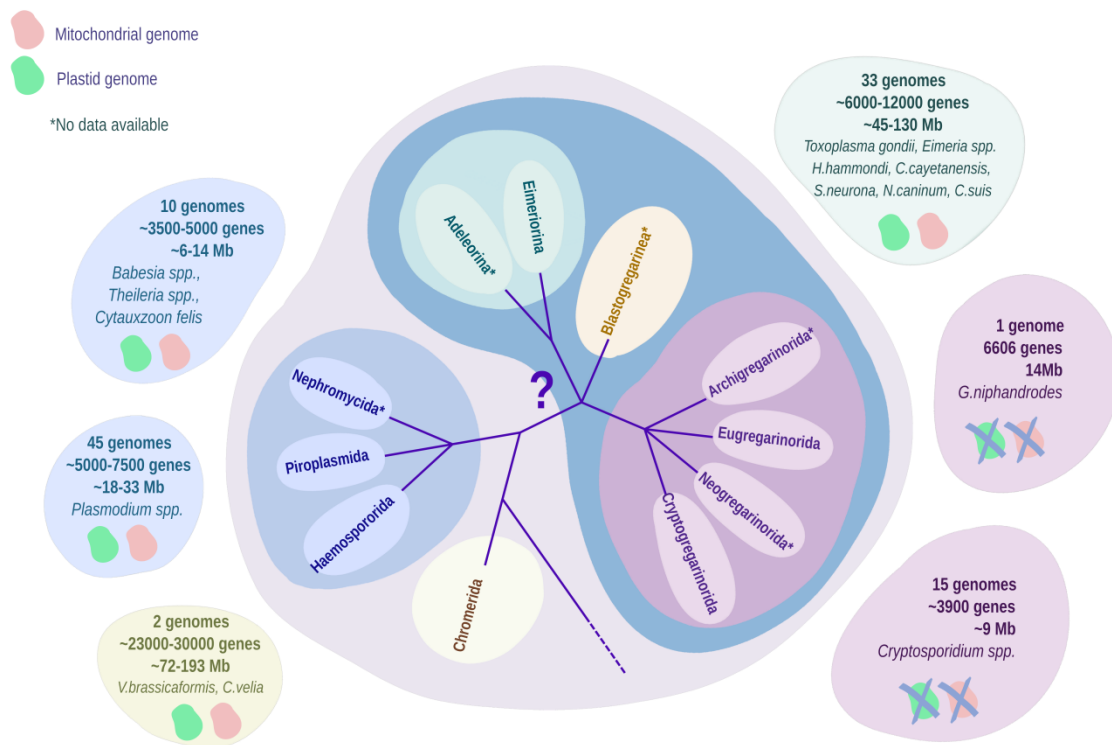
- Living Environments Using High-Throughput Amplicon Data and a Phylogenetically Informed Reference Framework. *Frontiers in microbiology*. 10:2373.
- Desportes I, and Schrével J. 2013. Treatise on Zoology - Anatomy, Taxonomy, Biology. The Gregarines, The early branching Apicomplexa (2 vols). Brill. 791 pages pp.
- Diakin, A., G.G. Paskerova, T.G. Simdyanov, V.V. Aleoshin, and A. Valigurova. 2016. Morphology and Molecular Phylogeny of Coelomic Gregarines (Apicomplexa) with Different Types of Motility: *Urospora ovalis* and *U. travisiae* from the Polychaete *Travisia forbesii*. *Protist*. 167:279-301.
- Diakin, A., K.C. Wakeman, and A. Valigurova. 2017. Description of *Ganymedes yurii* sp. n. (Ganymedidae), a New Gregarine Species from the Antarctic Amphipod *Gondogeneia* sp. (Crustacea). *The Journal of eukaryotic microbiology*. 64:56-66.
- Frenal, K., J.F. Dubremetz, M. Lebrun, and D. Soldati-Favre. 2017. Gliding motility powers invasion and egress in Apicomplexa. *Nature reviews. Microbiology*. 15:645-660.
- Füßy, Z., and M. Oborník. 2017. Reductive Evolution of Apicomplexan Parasites from Phototrophic Ancestors. In *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts*. P. Pontarotti, editor. Springer International Publishing, Cham. 217-236.
- Gawad, C., W. Koh, and S.R. Quake. 2016. Single-cell genome sequencing: current state of the science. *Nature reviews. Genetics*. 17:175-188.
- Gentekaki, E., M. Kolisko, V. Boscaro, K.J. Bright, F. Dini, G. Di Giuseppe, Y. Gong, C. Miceli, L. Modeo, R.E. Molestina, G. Petroni, S. Pucciarelli, A.J. Roger, S.L. Strom, and D.H. Lynn. 2014. Large-scale phylogenomic analysis reveals the phylogenetic position of the problematic taxon *Protocruzia* and unravels the deep phylogenetic affinities of the ciliate lineages. *Molecular phylogenetics and evolution*. 78:36-42.
- Gubbels, M.J., and M.T. Duraisingh. 2012. Evolution of apicomplexan secretory organelles. *International journal for parasitology*. 42:1071-1081.
- Hakimi, M.A., P. Olias, and L.D. Sibley. 2017. Toxoplasma Effectors Targeting Host Signaling and Transcription. *Clinical microbiology reviews*. 30:615-645.
- Janouskovec, J., G.G. Paskerova, T.S. Miroljubova, K.V. Mikhailov, T. Birley, V.V. Aleoshin, and T.G. Simdyanov. 2019. Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. *eLife*. 8.
- Janouskovec, J., D.V. Tikhonenkov, F. Burki, A.T. Howe, M. Kolisko, A.P. Mylnikov, and P.J. Keeling. 2015. Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proceedings of the National Academy of Sciences of the United States of America*. 112:10200-10207.
- Kuriyama, R., C. Besse, M. Geze, C.K. Omoto, and J. Schrével. 2005. Dynamic organization of microtubules and microtubule-organizing centers during the sexual phase of a parasitic protozoan, *Lecudina tuzetae* (Gregarine, Apicomplexa). *Cell Motil Cytoskeleton*. 62:195-209.
- Kwong, W.K., J. Del Campo, V. Mathur, M.J.A. Vermeij, and P.J. Keeling. 2019. A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. *Nature*. 568:103-107.
- Mahe, F., C. de Vargas, D. Bass, L. Czech, A. Stamatakis, E. Lara, D. Singer, J. Mayor, J. Bunge, S. Sernaker, T. Siemensmeyer, I. Trautmann, S. Romac, C. Berney, A. Kozlov, E.A.D. Mitchell, C.V.W. Seppey, E. Egge, G. Lentendu, R. Wirth, G. Trueba, and M. Dunthorn. 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nature ecology & evolution*. 1:91.
- Mathur, V., M. Kolisko, E. Hehenberger, N.A.T. Irwin, B.S. Leander, A. Kristmundsson, M.A. Freeman, and P.J. Keeling. 2019. Multiple Independent Origins of Apicomplexan-Like Parasites. *Current biology : CB*. 29:2936-2941 e2935.
- Morrisette, N.S., and L.D. Sibley. 2002. Cytoskeleton of apicomplexan parasites. *Microbiology and molecular biology reviews : MMBR*. 66:21-38; table of contents.
- Mueller, C., A. Graindorge, and D. Soldati-Favre. 2017. Functions of myosin motors tailored for parasitism. *Current opinion in microbiology*. 40:113-122.



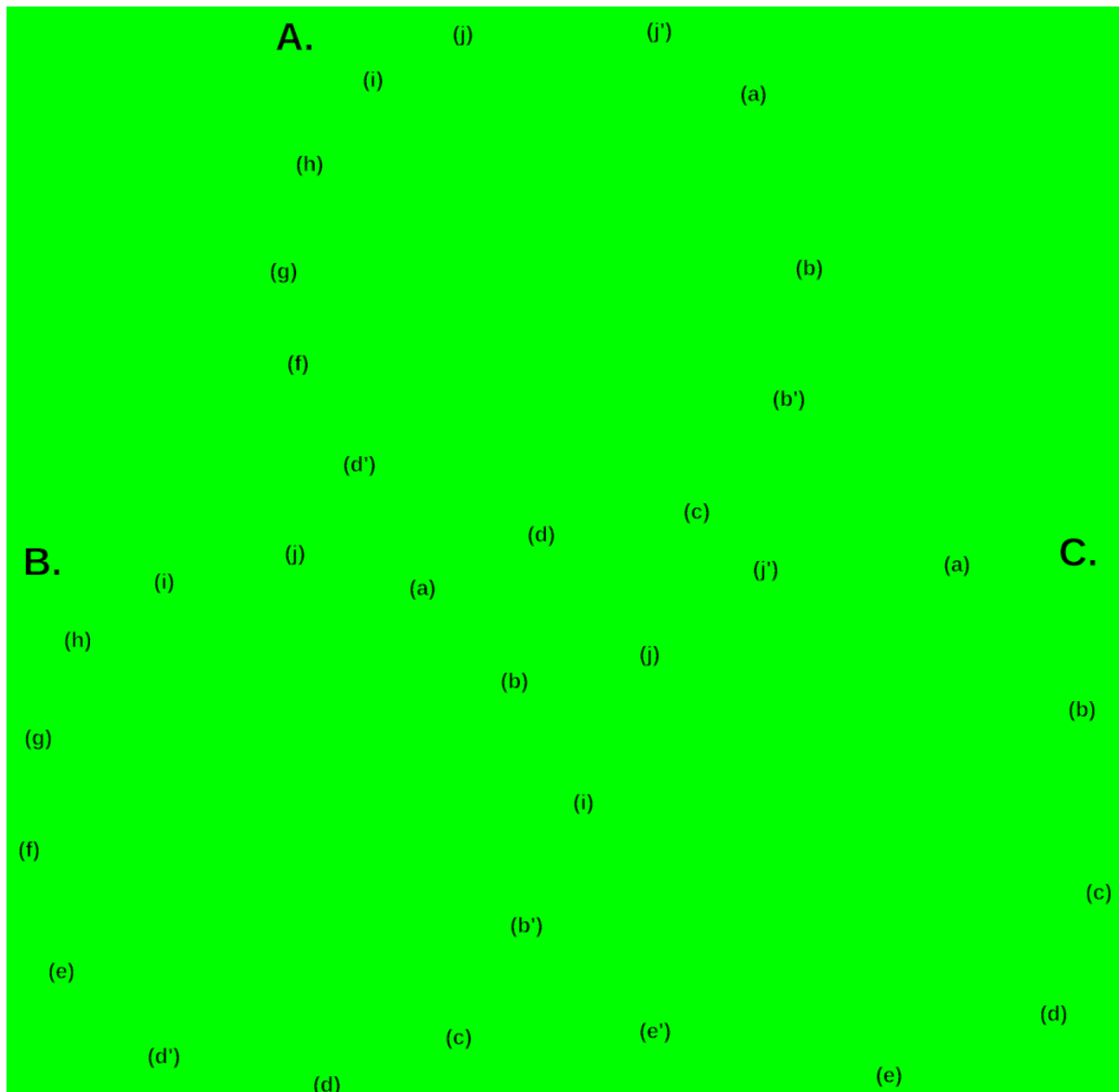
- Omoto, C.K., M. Toso, K. Tang, and L.D. Sibley. 2004. Expressed sequence tag (EST) analysis of Gregarine gametocyst development. *International journal for parasitology*. 34:1265-1271.
- Portman, N., and J. Slapeta. 2014. The flagellar contribution to the apical complex: a new tool for the eukaryotic Swiss Army knife? *Trends in parasitology*. 30:58-64.
- Prensier, G., J.F. Dubremetz, and J. Schrevel. 2008. The unique adaptation of the life cycle of the coelomic gregarine *Diplauxis hatti* to its host *Perinereis cultrifera* (Annelida, Polychaeta): an experimental and ultrastructural study. *The Journal of eukaryotic microbiology*. 55:541-553.
- Robert-Gangneux, F., and M.L. Darde. 2012. Epidemiology of and diagnostic strategies for toxoplasmosis. *Clinical microbiology reviews*. 25:264-296.
- Rueckert, S., E.L. Betts, and A.D. Tsaousis. 2019a. The Symbiotic Spectrum: Where Do the Gregarines Fit? *Trends in parasitology*. 35:687-694.
- Rueckert, S., and A. Horak. 2017. Archigregarines of the English Channel revisited: New molecular data on *Selenidium* species including early described and new species and the uncertainties of phylogenetic relationships. *PLoS one*. 12:e0187430.
- Rueckert, S., S.V. Pipaliya, and J.B. Dacks. 2019b. Evolution: Parallel Paths to Parasitism in the Apicomplexa. *Current biology : CB*. 29:R836-R839.
- Schrevel, J., and I. Desportes. 2015. Gregarines. In *Encyclopedia of Parasitology*. H. Mehlhorn, editor. Springer-Verlag, Berlin Heidelberg.
- Schrevel, J., A. Valigurova, G. Prensier, A. Chambouvet, I. Florent, and L. Guillou. 2016. Ultrastructure of *Selenidium pendula*, the Type Species of Archigregarines, and Phylogenetic Relations to Other Marine Apicomplexa. *Protist*. 167:339-368.
- Sibbald, S.J., and J.M. Archibald. 2017. More protist genomes needed. *Nature ecology & evolution*. 1:145.
- Simdyanov, T.G., L. Guillou, A.Y. Diakin, K.V. Mikhailov, J. Schrevel, and V.V. Aleoshin. 2017. A new view on the morphology and phylogeny of eugregarines suggested by the evidence from the gregarine *Ancora sagittata* (Leuckart, 1860) Labbe, 1899 (Apicomplexa: Eugregarinida). *PeerJ*. 5:e3354.
- Simdyanov, T.G., and O.N. Kuvardina. 2007. Fine structure and putative feeding mechanism of the archigregarine *Selenidium orientale* (Apicomplexa: Gregarinomorpha). *European journal of protistology*. 43:17-25.
- Simdyanov, T.G., G.G. Paskerova, A. Valigurova, A. Diakin, M. Kovacikova, J. Schrevel, L. Guillou, A.A. Dobrovolskij, and V.V. Aleoshin. 2018. First Ultrastructural and Molecular Phylogenetic Evidence from the Blastogregarines, an Early Branching Lineage of Plesiomorphic Apicomplexa. *Protist*. 169:697-726.
- Tardieux, I., and J. Baum. 2016. Reassessing the mechanics of parasite motility and host-cell invasion. *J Cell Biol*. 214:507-515.
- Templeton, T.J., S. Enomoto, W.J. Chen, C.G. Huang, C.A. Lancto, M.S. Abrahamsen, and G. Zhu. 2010. A genome-sequence survey for *Ascogregarina taiwanensis* supports evolutionary affiliation but metabolic diversity between a Gregarine and *Cryptosporidium*. *Molecular biology and evolution*. 27:235-248.
- Valigurova, A., L. Hofmannova, B. Koudela, and J. Vavra. 2007. An ultrastructural comparison of the attachment sites between *Gregarina steini* and *Cryptosporidium muris*. *The Journal of eukaryotic microbiology*. 54:495-510.
- Valigurova, A., and B. Koudela. 2008. Morphological analysis of the cellular interactions between the eugregarine *Gregarina garnhami* (Apicomplexa) and the epithelium of its host, the desert locust *Schistocerca gregaria*. *European journal of protistology*. 44:197-207.
- Valigurova, A., N. Vaskovicova, N. Musilova, and J. Schrevel. 2013. The enigma of eugregarine epicytic folds: where gliding motility originates? *Frontiers in zoology*. 10:57.
- Woo, Y.H., H. Ansari, T.D. Otto, C.M. Klinger, M. Kolisko, J. Michalek, A. Saxena, D. Shanmugam, A. Tayyrov, A. Veluchamy, S. Ali, A. Bernal, J. del Campo, J. Cihlar, P. Flegontov, S.G. Gornik, E. Hajduskova, A. Horak, J. Janouskovec, N.J. Katris, F.D. Mast, D. Miranda-Saavedra, T.

Mourier, R. Naeem, M. Nair, A.K. Panigrahi, N.D. Rawlings, E. Padron-Regalado, A. Ramaprasad, N. Samad, A. Tomcala, J. Wilkes, D.E. Neafsey, C. Doerig, C. Bowler, P.J. Keeling, D.S. Roos, J.B. Dacks, T.J. Templeton, R.F. Waller, J. Lukes, M. Obornik, and A. Pain. 2015. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *eLife*. 4:e06974.

## 10. Figure legends

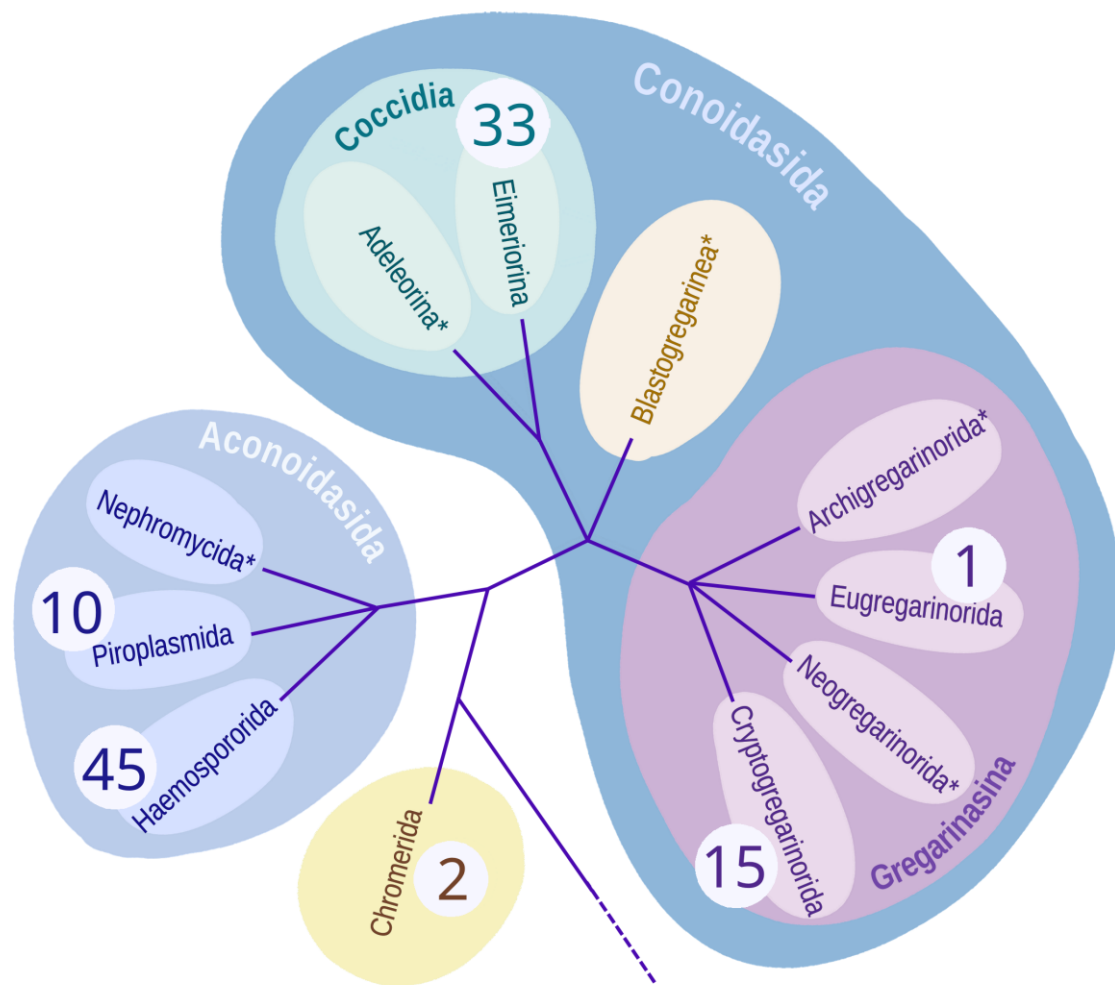


**Figure 1. The genomic panorama of Apicomplexa.** On this schematic representation inspired by Portman and Slapeta, 2014 (Portman and Slapeta, 2014) and using the most recent taxonomy by Adl et al., 2019 (Adl et al., 2018), we have compiled the genomic information currently available on Apicomplexa (104 genomes) and proto-Apicomplexa (Chromerida, 2 genomes), mostly available from EupathDB (Aurrecochea et al., 2017). For each group of data (6 in total) we have indicated: the number of available genomes, specifying the concerned species, the number of protein-coding genes, the nuclear genome size in Mb and the presence (or absence in *Cryptosporidium* spp; the question being still open in the case of the (unpublished) *G. niphandrodes* genomic data), of mitochondrial or plastid genomes. The question mark symbolizes the currently unresolved branching order of the various apicomplexan groups.



**Figure 2. Representative development cycles for 3 gregarines.** The developmental cycles of: **A.** the (marine) archigregarine *Selenidium pendula*, intestinal parasite of the polychaeta *Scolecipis (Nerine) squamata*, adapted from (Schrevel and Desportes, 2015); **B.** the marine eugregarine *Lecudina tuzetae*, intestinal parasite of the Polychaeta *Hediste (Nereis) diversicolor*, adapted from (Schrevel and Desportes, 2015); **C.** the terrestrial eugregarine *Gregarina garnhami* (C), intestinal parasite of the desert locust *Schistocerca gregaria* (C), based on (Canning, 1956) and personal observations. The drawings use identical “legend letters” to designate similar developmental stages across the three cycles. **A. *Selenidium pendula*:** (a) detached trophozoite; (b) caudal syzygy; (b’) particularity in syzygy for this species (nuclear modifications before encystment); (c) gametocyst undergoing gamogony; (d) gametocyst with fully differentiated gametes; (d’) details of male (flagellated) and female (ovoid) gamete; (e) zygote ready to undergo sporogony yielding stages with 2 nuclei (g), then 4 nuclei (h); (i) spore containing 4 sporozoites ; (j) release sporozoite (in host) starting vegetative phase. **B. *Lecudina tuzetae*.** (a) detached trophozoite; (b) lateral syzygy; (c) gametocyst undergoing gamogony; (d) gametocyst with fully differentiated gametes; (d’) details of male (flagellated) and female (ovoid) gamete; (e) sporokyst enclosing ~5000 zygotes ready to evolve into spore (f) eventually undergoing sporogony yielding stages with 2 nuclei (g), 4 nuclei (h); (i) spore containing 8 sporozoites ; (j)

released sporozoite (in host) starting vegetative phase including attachment to host epithelial cell (not shown). **C. *Gregarina garnhami***. (a) detached trophozoite; (b) caudo-frontal syzygy (primate ahead, satellite following); (c) gametocyst undergoing gamogony; (d) gametocyst with fully differentiated gametes; (e) sporokyst enclosing zygotes ready to evolve into spores undergoing sporogony (details not shown); these spores are released in the environment as spore chains (ch) through sporoducts (sp) emerging from the sporokyst; (i) spore containing 8 sporozoites ; (j) released sporozoite (in host) starting vegetative phase including attachment to host epithelial cell (j'). Cyst or spore walls surround developmental stages from (c) to (i).



With just over 100 genomes deciphered for ~6000 described species, the -omic exploration of Apicomplexa has so far mainly concerned a limited number of intracellular pathogens of vertebrates. Gregarines, full members of this phylum infecting non-vertebrates, mostly extracellular and nonpathogenic, have been so far left on the side of the road while their study offers strong promises to reveal novel, original and unknown adaptive mechanisms and evolutionary history of Apicomplexa.