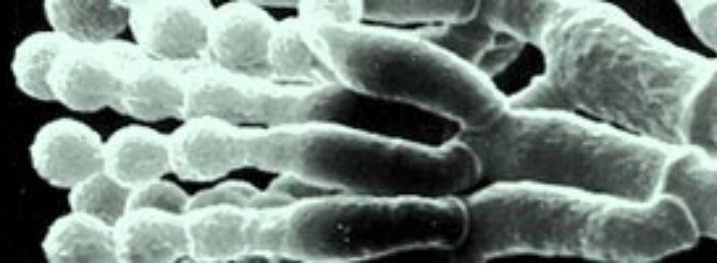
MiSAC matters Articles



An introduction to molecular biology approaches in parasitic protozoa:

1. Investigating gene function in protozoan parasites – general approaches and challenges

Tansy Hammarton

Protozoa are eukaryotic single-celled motile microorganisms that may be free-living and harmless or parasitic, meaning they live in or on a host organism and cause it harm. Parasitic protozoa are responsible for some debilitating and deadly diseases of humans and animals, such as malaria, trypanosomiasis (more commonly called sleeping sickness in humans or Nagana in livestock), leishmaniasis and toxoplasmosis. These protozoa have **digenetic** life cycles, meaning they undergo development in two hosts (either two warm-blooded animals or a mammal/bird/reptile and an insect **vector**; see Fig. 1 and also '*Hiding in plain sight: how parasites have evolved to outsmart their hosts*,' (MiSACmatters Articles) and they change their morphology (shape), structure and biochemical properties as they progress through their life cycles.

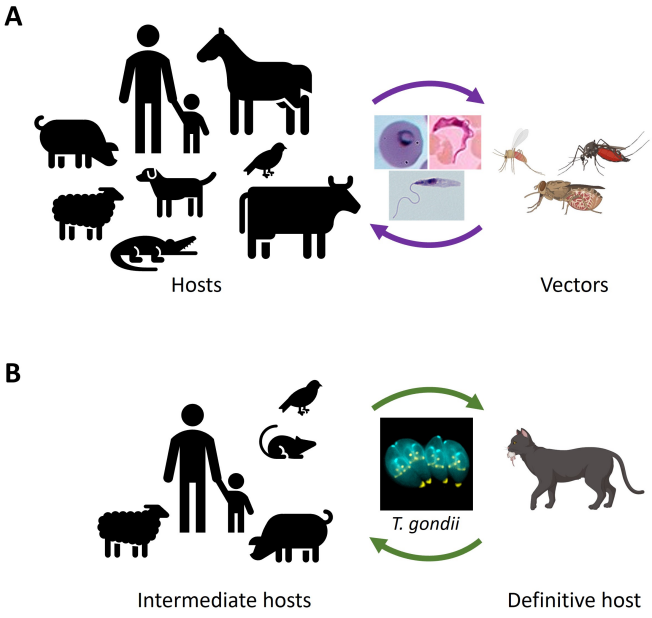


Figure 1. Protozoan digenetic life cycles.
Some protozoa undergo development in two hosts and are said to have a digenetic life cycle. A: *Plasmodium* spp. which cause malaria are spread between mammals and birds by mosquito vectors, while *Trypanosoma* spp. which cause animal African trypanosomiasis (sleeping sickness or Nagana) in mammals and can even infect reptiles such as crocodiles, are spread by tsetse flies, and *Leishmania* species, causing leishmaniasis in humans, dogs and other animals, are spread between animal hosts by sand flies. B: In comparison, the protozoan, *Toxoplasma gondii*, can infect virtually any warm-blooded animal (intermediate hosts), but undergoes sexual development only in cats or other felines (definitive host). Infected cats shed infective oocysts in their faeces which can be inadvertently consumed by other animals. *Toxoplasma* forms tissue cysts in intermediate hosts; cats become infected by eating infected mice or birds, while humans can acquire *Toxoplasma* infections through handling cat litter trays or by consuming undercooked meat e.g., lamb or pork. Parasite image credits: *Plasmodium falciparum*: Dr Graham Beards; *Trypanosoma*: CDC/Dr Myron G. Schultz; *Leishmania*: Stefan Walkowski; *Toxoplasma gondii*: Ke Hu & John F. Murray.

Created with BioRender.com.

Understanding parasite biology, particularly at the molecular level, enables scientists to understand how they cause disease and how they have evolved separately from their hosts. In some cases, studying processes in parasites has helped to understand related processes in humans. For example, studying the trypanosome flagellum (a tail-like structure that allows the parasite to swim) and comparing to related structures called cilia in humans, enabled the identification of genes that when mutated can cause human diseases by interfering with cilial motility. However, many aspects of protozoan cell biology are distinct compared to traditional **model organisms** such as yeast or animal cells due to differences in the proteins they express. Investigating the function of genes and their encoded proteins within the protozoa themselves can help scientists to understand these differences. This is not only interesting in its own right, but also allows essential parasite-specific proteins to be identified (for example, key signalling molecules called protein kinases), that potentially could be targeted by new drugs that do not cause too much collateral damage to their hosts, improving treatment options for parasitic diseases such as malaria, trypanosomiasis and leishmaniasis.

This article is one of a number of MiSAC *matters Articles*, written by experts, which are up to date and relevant to microbiology in schools. MiSAC is grateful to all contributing authors'. E: microbe@misac.org.uk. W: *www.misac.org.uk.*

Predicting the function of parasite genes and proteins

With advances in genome sequencing methods, an increasing number of genomes of parasites and their vectors are being sequenced. Parasite sequencing data are publicly available through the Eukaryotic Pathogen, Vector and Host Informatics Resource (VEuPathDB). Knowing the genome sequence of an organism allows scientists to use computing or **bioinformatic** approaches to identify the genes present. By comparing them to similar previously studied genes in other organisms, predictions can be made about the functions of the proteins they encode, as well as, for example, where they might be found (or localise) in the cell. This is a good starting point, but all predictions need to be validated by performing experiments in the lab as, often, parasite genes/proteins do not function in exactly the same way their counterparts do in other organisms. Further, some genes from e.g. humans or model organisms will not be present in all parasites, and equally, parasites possess genes that other organisms do not. In some cases, the parasite-specific genes can encode proteins with little or no resemblance to any other previously studied protein, and therefore it can be hard to predict their function using bioinformatic approaches alone.

How can parasite gene/protein function be investigated in the lab?

Molecular genetic techniques are key to being able to carry out functional analyses of genes or proteins in parasites and have revolutionised our understanding of parasite cell biology. For example, they allow scientists to remove one or more genes from a parasite (a **gene knockout**) or to mutate a gene so that it no longer encodes a functional protein, or to reduce or increase the amount of a protein the cell expresses. By examining the effects (or **phenotypes**) of these alterations on the parasite, it is possible to determine what the gene does normally. For example, if mutating or knocking out a gene, or reducing the amount of protein expressed, results in parasites with a different shape, then the protein encoded by the gene might be involved in maintaining cell morphology. Alternatively, if increasing expression of a protein means that a parasite moves faster, then that protein is likely important for motility. Further, manipulation of a gene encoding e.g., a parasite surface protein might mean that the parasite is less effective at infecting or invading a host, indicating that the surface protein is important for key interactions with host cells that allows the parasite to gain entry, multiply, spread and/or evade the host's immune system. Sometimes, despite repeated efforts, it is not possible to knockout a particular gene, and this can be an indication that the gene is essential for the parasite to multiply or survive and may therefore be a good candidate to try to develop novel drugs against. For further detail and additional examples of using molecular genetic techniques in parasites, see article '*An introduction to molecular biology approaches in parasitic protozoa: 2. Molecular genetic methods for analysing gene function in protozoan parasites*' (MiSACmatters Articles).

Challenges of using molecular genetics in parasites

Molecular genetic techniques are extremely powerful tools but can be challenging to perform in some parasites (Fig. 2). Many molecular genetic techniques require that scientists can culture the parasite in the laboratory (which is still difficult or not yet achievable for some organisms) and that the parasite can be genetically modified. To genetically modify a parasite, DNA or RNA needs to be introduced into the parasite in a process known as **transfection**. Transfection is often achieved using **electroporation**, where an electrical current is applied to parasites suspended in a buffer containing the DNA/RNA, which makes the parasites permeable and allows the DNA/RNA to get into the cell. However, transfection is quite an inefficient process, so a method of isolating those parasites that have taken up the DNA/RNA from those that haven't (known as **selection)** is

also required. Often, a drug resistance gene is transfected along with the desired genetic material. This makes the parasites that have taken up the genetic material resistant to a drug that would usually kill them, meaning that by simply growing the parasites in medium containing the drug, only successfully transfected parasites will survive. Alternatively, selection can be achieved by introducing a gene encoding a fluorescent protein along with the desired genetic material, enabling transfected parasites (fluorescent) to be sorted from non-transfected (non-fluorescent) parasites using a machine called a Fluorescence Activated Cell Sorter (FACS). For some parasites, transfection and selection are now routine techniques, but it can take many months or even years to optimise these techniques and make them efficient enough to be useful for a given parasite. Furthermore, even if transfection and selection is possible, there can be additional hurdles in ensuring that introduced DNA is stably integrated into the parasite genome, finding suitable promoters that allow expression of introduced genes to be turned on or off as required, or in having the appropriate cellular machinery present in the parasite to carry out molecular techniques such as **RNA interference** or **CRISPR/Cas9 gene editing** (see article '*An introduction to molecular biology approaches in parasitic protozoa: 2. Molecular genetic methods for analysing gene function in protozoan parasites*'; MiSAC*matters* Articles). For many parasites, including the gut parasite *Giardia* or the sexually transmitted parasite, *Trichomonas,* molecular biology is still in its infancy, and as a result, these parasites are far less studied than those for which molecular approaches are well-developed.

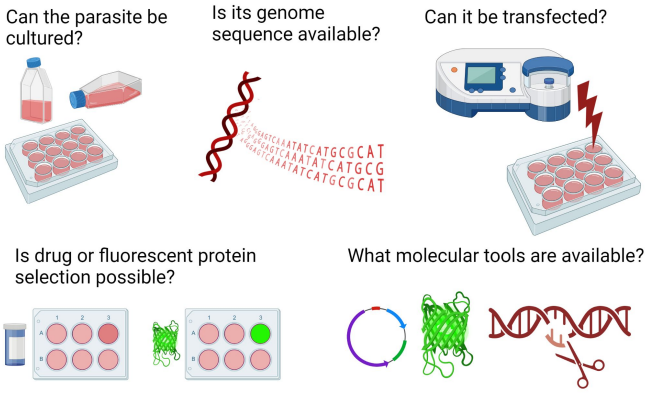


Figure 2. Considerations for attempting molecular biology of a protozoan parasite.

Created with BioRender.com.

Glossary

Bioinformatics – use of computational tools to analyse DNA, RNA and protein sequences.

CRISPR/Cas9 gene editing – a natural anti-viral defence system from bacteria that has been harnessed by scientists to provide a customisable lab tool to edit an organism's genome in a specific way e.g. by changing individual DNA bases, deleting genes or inserting DNA sequences.

Digenetic – having two hosts.

Electroporation – lab technique involving applying an electric field to cells to increase their cell membrane permeability, allowing the introduction of DNA or RNA to the cell.

Gene knockout – a cell line or organism where a gene has been removed using molecular biology techniques. Removing part of a gene or mutating a crucial region of a gene can result in a non-functional protein being expressed, creating a functional gene knockout without the whole gene having been removed.

Model organism – a non-human organism that is extensively studied by scientists in the lab to shed light on particular biological processes, for example, mice, the fruit fly *Drosophila melanogaster*, the worm *Caenorrhabditis elegans*, the frog *Xenopus laevis*, the bacterium *Escherichia coli* and the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe.* Typically, these organisms are easy to handle and genetically manipulate in the lab, and it is hoped that the understanding gained from studying them is applicable to less tractable organisms.

Phenotype – observable characteristics of an organism e.g., size, shape, colour, motility, metabolism etc. Both gene expression and environmental conditions contribute to the phenotype observed.

RNA interference (RNAi) – an evolutionary ancient defence mechanism that exists in some organisms to protect them from infection from double stranded RNA (dsRNA) viruses. Scientists can harness this naturally occurring response to degrade cellular messenger RNA (mRNA) for a protein of choice, reducing the amount of the protein that is produced, and then observe the phenotype.

Selection – a method of separating out successfully transfected cells from non-transfected ones. Typically, successfully transfected cells will have gained resistance to a drug they are usually killed by, meaning that only they will grow once that drug is added to the culture medium, or alternatively they may have become fluorescent and can be separated from non-transfected (non-fluorescent) cells using a fluorescence-activated cell sorter.

Transfection – a method to deliberately introduce nucleic acids into cells.

Vector – a living organism that transfers an infectious pathogen from one host to another.

AUTHOR PROFILE

Tansy Hammarton is a Senior Lecturer in the School of Infection & Immunity at the University of Glasgow. After obtaining a BA in Natural Sciences from the University of Cambridge and a PhD in Microbiology at the University of Manchester, she moved to Glasgow to study the cell cycle of kinetoplastid parasites. Her group's research currently focusses on the signalling molecules that regulate mitosis and cytokinesis in trypanosomes and *Leishmania*. Dr Hammarton also has a keen interest in public engagement, particularly in working with schools, and was awarded the 2018 Microbiology Society's Peter Wildy Prize Lecture for her work. She is currently the Microbiology Society's representative on MiSAC.