

Module 6 – RNA-seq 2

Introduction

Plasmodium berghei is used as a rodent model of malaria. It is known that in the lab it can evolve to stop producing sexual stages (Figure 1). We want to try and use this observation to our advantage. If we can understand how the parasite switches to the sexual, transmissible stage, then we might better understand how to prevent this from happening and prevent the spread of malaria.

Our colleagues grew several cultures of a transmissible strain in the lab continuously for several months (Figure 2). These all become gametocyte non-producers (GNPs). They then sequenced the genomes of each strain and identified the mutations in each one. They hypothesised that, while each GNP strain would have many mutations compared to the parental strain, only one gene would be mutated in every strain and this gene would be a key regulator of gametocytogenesis. This worked and they identified the *ap2-g* gene, which is a transcription factor and is now thought to regulate early steps of gametocytogenesis (Sinha et al., 2014, PMID: 24572359).

Our colleagues have proposed that we explore the function of this transcription factor by looking at changes in gene expression after it has been knocked-out. One of their team managed to knock it out, grow the mutant up, extract and sequence its RNA. This means we are able to examine the role of the gene and how it affects the transcriptional landscape of the parasite. Which transcripts are affected by the knockout of this gene? Does this make sense in terms of what we expect to happen? What could this gene list be useful for in future?

Sequencing reads as well as the reference sequence are available in the data directory **Module_6_RNAseq2** (please use this rather than download one). There are also files of genome annotation, product descriptions, GO terms and an R script for performing GO term enrichment. A full explanation is found in a README file in the directory.

Summary

To achieve this goal you should:

- Map the RNA-seq data to the reference
- Confirm the knockout in the mutant samples
- Call differentially expressed genes
- Perform a Gene Ontology enrichment analysis (we have provided an R script to help with this)

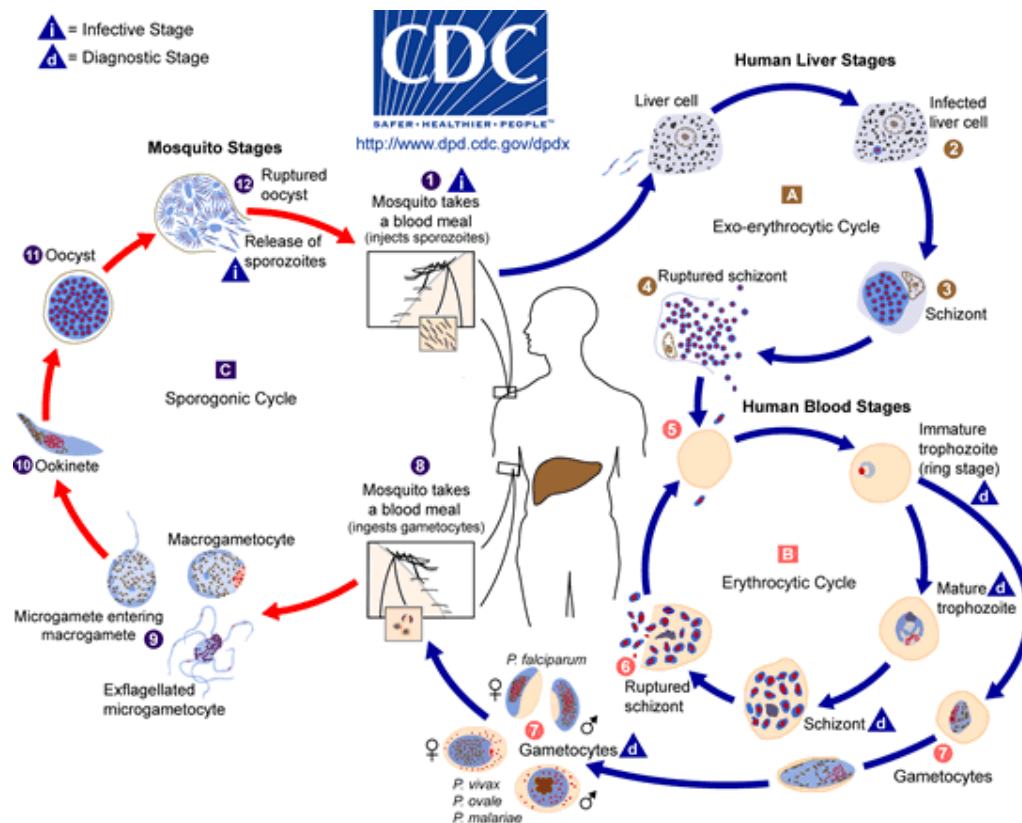


Figure 1. Malaria life cycle. The parasites must produce sexual forms called gametocytes in order to be transmitted from the intermediate mammalian host to the definitive mosquito host.

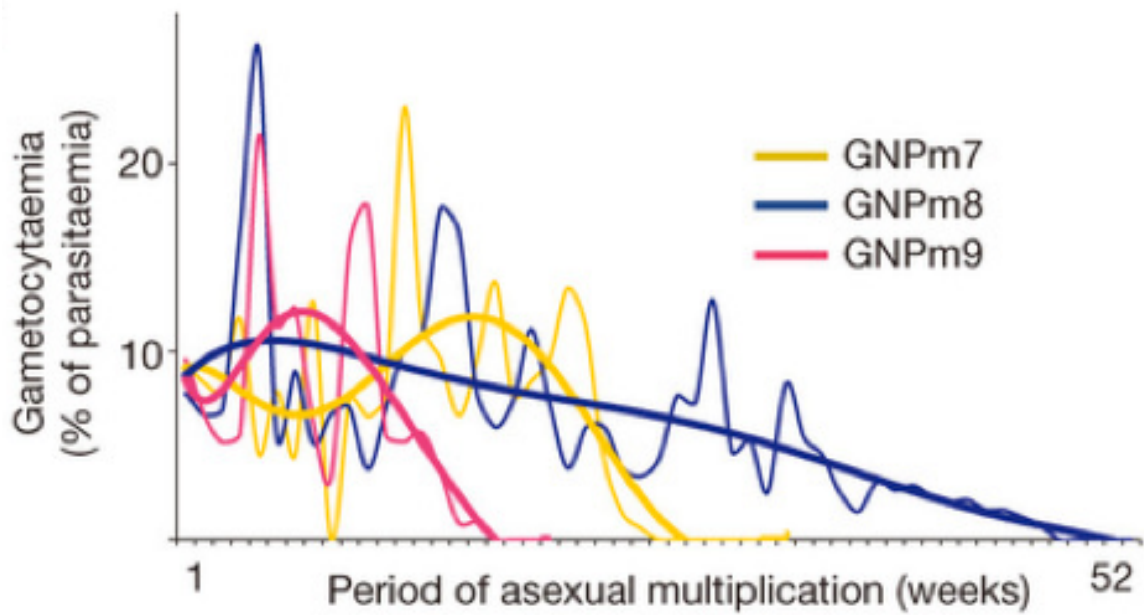


Figure 2. Parasites were passaged continuously in mice until they lost the ability to produce gametocytes. The bold lines show smoothed versions of the real data to even out fluctuations in gametocytemia. GNPm7-9 are different lines of parasites from the same parent which have independently lost the ability to produce gametocytes (transmissible sexual stages).