## 3.1 Introduction

The first aim of this project was to adapt the STM strategy to *P. berghei* parasites in order to enable large scale genetic screening.

STM screens have been designed in many ways, reflecting the diversity of genetic systems of different taxa. Common to all is a workflow that starts with (1) mutagenesis, i.e. generation of barcoded mutants, which is followed by (2) propagation of these mutants in pools and finally (3) identification of the mutants present after propagation through their barcode.

One main approach used to generate barcoded mutants in bacteria is *in vivo* transposition (Fig. 1.6). Fonager and colleagues have applied a *piggyBac* transposition system to *P. berghei* parasites, but fine tuning is yet to be achieved [46]. In yeast, directed gene-replacement has been the most used method to generate libraries of thousands of mutants that are then pooled and used in STM approaches [94]. Modifications that rely on homologous recombination are probably the most reliable method for genetic modification of *Plasmodium* parasites, although at different frequencies according to the species. However, the approach taken by the yeast field would be of very little use for *P. berghei* parasites as at least 12 mice are needed to generate a single clonal line.

Recently, a new generation of *P. berghei* targeting vectors was developed – the *PlasmoGEM* vectors [73]. These are linear vectors in which the length of homology arms is increased from 0.3 – 1.0 kb to several kb, to improve homologous integration frequencies. Additionally, they have not been reported to persist as episomes after transfection, which decreases the rate of false positives, as drug selection ensures elimination of the parasites where integration did not take place. A combination of improved integration with reduction of false positives made these vectors promising tools for a *P. berghei* adapted STM. To allow identification of mutants generated by these vectors within a pool, gene-specific barcodes were introduced into their basic design that labels mutants upon genome integration. These barcodes consisted of a 10-11 mer DNA sequence that was inserted into a ~100 bp-long module, located next to the B2 gateway site in all vectors. The length of the barcode permitted

that enough sequences with a hamming distance of four[1] (i.e. single error correction plus double error detection) [158], could be generated to cover the entire *Plasmodium* genome.

The barcode module was flanked by constant annealing sites, which enabled a bias-free amplification of all barcodes from a pool through a single PCR reaction.

Taken together these tools offered the opportunity to perform STM-like experiments in *P. berghei*. We hypothesised that the properties of the *Plasmo*GEM vectors would enable the generation of complex pools of mutants, thus circumventing the need to generate each mutant independently prior to parallel phenotyping. In other words, transfection of pools of barcoded *Plasmo*GEM vectors (mutagenesis step) would generate pools of barcoded mutants that after being expanded in a single mouse (propagation step) could be identified through their barcodes (identification step).

To test this hypothesis, various parameters needed to be optimised. These included transfection conditions, barcode detection, sequencing library preparation and sequencing run conditions.

## 3.2 Results

### 3.2.1 Optimisation of operating conditions for transfection

The ability to develop STM based screening in *Plasmodium* depends on the complexity of mutant pools that can be easily generated by co-transfecting multiple vectors. This in turn depends on the transfection efficiency that can be achieved. My first objective was therefore to identify the most suitable electroporator and the most adequate DNA concentration of each *Plasmo*GEM vector to use.

#### 3.2.1.1 Choice of electroporator

One aspect that is critical for transfection efficiency is the type of the electroporation system used. The traditional Bio-Rad instruments have been surpassed by the Lonza electroporators, which are now the most efficient devices used to generate *P. berghei* transgenics.

[1] A Hamming distance of four enables single error correction plus double error detection, i.e. it takes four mutations, or sequencing/synthesis errors for one barcode to become another; one mutation can be corrected and two can be detected.

Two different Lonza electroporator systems were tested for their efficiency – Nucleofector II and 4D Nucleofector X unit (Fig. 3.1A). Using the same pool of schizonts cultured from two different mice, four different transfections were performed with each electroporator.

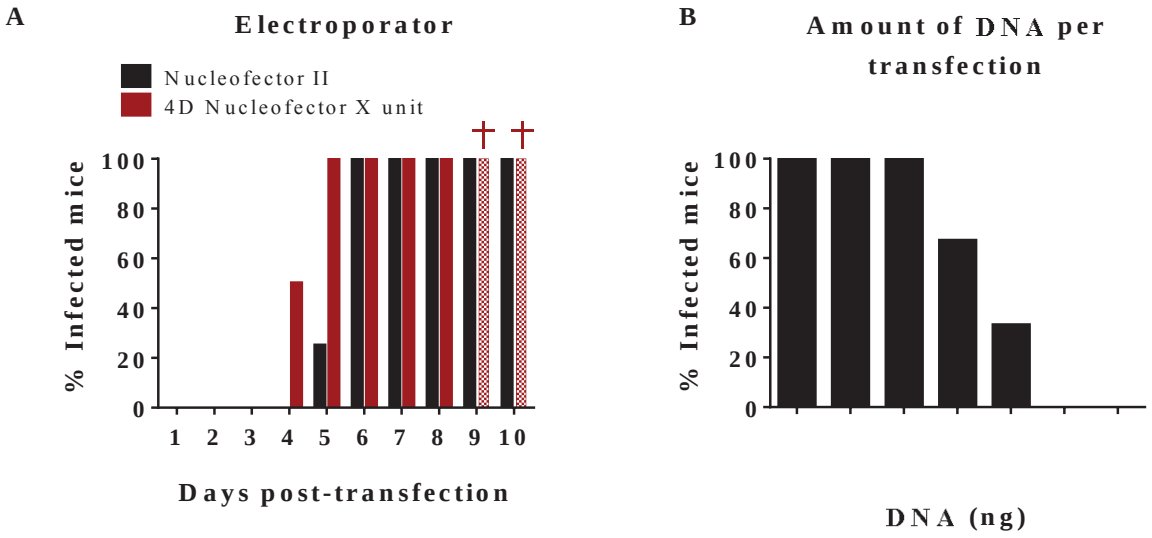



Fig. 3.1| The choice of electroporation system and DNA concentration are critical for maximum transfection efficiency.
(A) Impact of the choice of electroporator on patency. Four different transfections were performed using either the Nucleofector II or the 4D Nucleofector X unit electroporator to transfect 5 µg of a KO vector targeting *map1* gene. The graph shows the duration after transfection required for mice to develop parasitaemias visible on a Giemsa stained smear, which was used to infer transfection efficiency. † culled mice due to high infection. (B) Assessment of how little DNA is required for a successful transfection using the 4D Nucleofector X unit system (n=3 per concentration). All mice were injected intra-venously (i.v.) in the tail vein.

Transgenic parasites were obtained for all replicates. However, the 4D Nucleofector system proved to be more efficient since parasites were visible in the blood of two out four mice by day 4 post-transfection, whereas with the Nucleofector II system only one mouse had visible parasitaemia by day 5 post-transfection. All eight mice were diagnosed as infected on day 6.

Transfection efficiency was determined according to Janse *et al* [159] by comparing the number of surviving parasites in mice before and after pyrimethamine selection, based on the daily 10x multiplication rate of *P. berghei* in mice [160]. Mathematically this is defined as: $(n2/n1) \times 1/10^{d}$. To this end, the parasitaemia after injection of the transfected parasites (*n1*) was determined by counting Giemsa stained thin blood films at ~24h post-transfection, just before the start of the selection with pyrimethamine. Later, usually four to seven days (*d*) after

transfection, the number of drug-resistant parasites (*n2*) was determined from the parasitaemia counted on day *d*, when infection was patent.

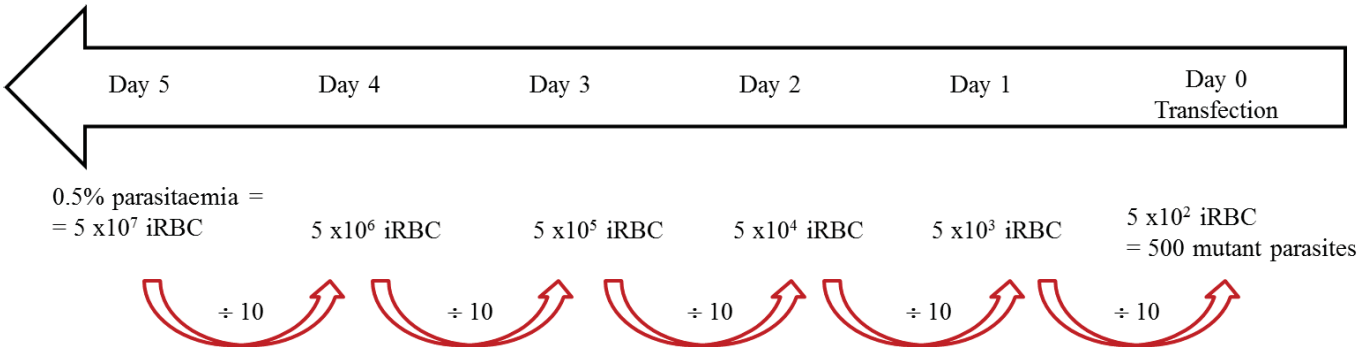
These results translated into a transfection efficiency of ~ $2x10^{-5}$ and ~$5x10^{-4}$ for the Nucleofector II and 4D, respectively.

### 3.2.1.2 Optimal DNA concentration

As transfection of pools of vectors was envisaged I explored what the minimal DNA concentration of each vector was to generate transgenic parasites, and whether very high concentrations would be more efficient.

For this, a dilution series of a KO vector for the *map1* gene (PbGEM-036210) was prepared to generate the following range: 5000 ng, 2000 ng, 1000 ng, 500 ng, 200 ng, 100 ng and 50 ng. Each of them was transfected in triplicates using the 4D Nucleofector X unit and the percentage of mice that became infected by day 10 post-transfection was registered. As expected, 2000 ng (the standard concentration for a *P. berghei* transfection) generated transgenic parasites for all replicates as did 5000 ng and 1000 ng, whereas concentrations below 200 ng were not successful. The transfection of 200 ng and 500 ng generated one and two infections, respectively (Fig. 3.1B). In addition, there was no difference in patency day for the highest concentrations, hence suggesting that above 2000 ng the system is probably close to saturation.

On day 5 post-transfection of the experiments where 1 µg of DNA was transfected, the observed parasitaemia was around 0.5 %. Based on the daily 10x multiplication rate of *P. berghei* parasites [160] during the exponential phase of growth this translates into approximately 500 independent integration events as the number of circulating RBCs in a mouse[2] is on average $1x10^{10}$, as detailed below:




[2] This RBC concentration applies to an average 6-8week-old, 30 g mouse.