Chapter 4

Discussion

The work described in this thesis was dependent upon the successful establishment of cultures of *Plasmodium falciparum* parasites from blood samples collected by collaborators in Peru and supplied as frozen isolates of samples that had been cultured initially in Peru. Culture adaptation of field isolates has typically had success rates around 60% (Gomez-Escobar 2010; Baum et al 2003), so the 24% growth success rate seen in the Peruvian field isolates is particularly low. This is compounded by the fact that these isolates were previously culture-adapted in Peru. There is no single reason for the low success rate of growth, however some samples appeared to have thawed and re-frozen either in Peru (where power outages are frequent) or in transit from Peru, which would severely impact their viability. The volumes of other samples suggested that they had not been frozen in a suitable volume of glycerolyte 57 which is also likely to impact viability. Although these technical errors exacerbated the problem, even once successfully grown at the Sanger Institute, Peruvian field isolates demonstrated an intrinsic susceptibility to the freezing process. All non-contaminated isolates grown and frozen at the Sanger Institute failed to be re-grown during multiple attempts (278 – 2 attempts, 6390 – 3 attempts; 9050 – 4 attempts). Only in the case of 6390 was a culture eventually successful after a fourth tube was thawed. It was observed that upon re-culturing these isolates, parasites were initially present but asexual parasites were lost to gametocytogenesis within the first few weeks and the culture subsequently died. Re-growth of the other successfully grown Peruvian isolates (5802, 5809 and 5814) has yet to be attempted. This susceptibility contrasted with the behaviour of the isolates that were confirmed as W2 contaminants (788, 3106, 3135, 3541 and 3769) which showed no such problems with the freezing process.

Among the non-contaminated Peruvian isolates that were successfully invasion phenotyped (278, 5802, 5809, 5814 and 6390) there was substantial variation in the dependence upon sialic acid-containing and chymotrypsin-sensitive receptors, although all isolates were highly dependent upon trypsin-sensitive receptors. None of the isolates were dependent upon sialic acid residues to the same extent as Dd2 or the W2 contaminants. The increased ability of these isolates (particularly 278 and 5809) to invade sialic acid depleted erythrocytes relative to Dd2 suggests that they are not reliant solely upon ligands of the DBL family for invasion which recognise the major sialoglycoproteins, the glycophorins. In the

case of isolate 5809, high resistance to neuraminidase and high sensitivity to trypsin- and chymotrypsin- mediated depletion of erythrocyte ligands implies the role of the PfRh4 pathway of invasion. The moderate sensitivity of the other isolates to neuraminidase and trypsin treated erythrocytes makes it harder to propose a reliance on a single receptor or ligand although for 6390, which invades chymotrypsin treated erythrocytes relatively well compared to those treated with neuraminidase or trypsin, arguments could be made that it preferentially invades cells expressing either glycophorin A or C, which are both chymotrypsin resistant. It should also be emphasised that the number of erythrocyte receptors that have been shown to bind merozoite ligands is not fully defined and therefore alternative interactions with similar enzyme sensitivities could be taking place that have yet to be identified.

Previous studies have defined sensitivity to an enzyme treatment as an invasion efficiency of less than 50%. In Brazil four different invasion phenotypes were observed in fourteen isolates. In the Peruvian samples three different invasion phenotypes were observed in the five non-contaminated isolates. The neuraminidase resistant, trypsin sensitive and chymotrypsin sensitive phenotype shared by 278 and 5809 was not seen in Brazil, however this profile was the predominant invasion phenotype seen in Kenyan isolates (Deans et al. 2007). 5802, 5814 and 6390 displayed invasion profiles similar to those seen in Brazil.

Although the sample numbers are low, the high level of diversity in invasion phenotype from Peruvian isolates already observed is an indication that natural variation of invasion pathways is intrinsically present, even without the high genetic diversity associated with hyper-endemic transmission regions such as in sub-Saharan Africa.

The negative correlation between invasion efficiency into neuraminidase- and chymotrypsin-treated erythrocytes was contrary to studies of samples obtained from Tanzania and Senegal. In Tanzanian samples, significant positive correlation was found between invasion into trypsin- and chymotrypsin-treated erythrocytes. However, once the phenotypes of W2 contaminated samples were removed, the Peruvian sample numbers were no longer sufficient to produce significant correlation (Fig. 4.1). The positive

correlation between PMR and invasion into chymotrypsin-treated erythrocytes was also found to be insignificant when the contaminants were excluded (Fig. 4.2).

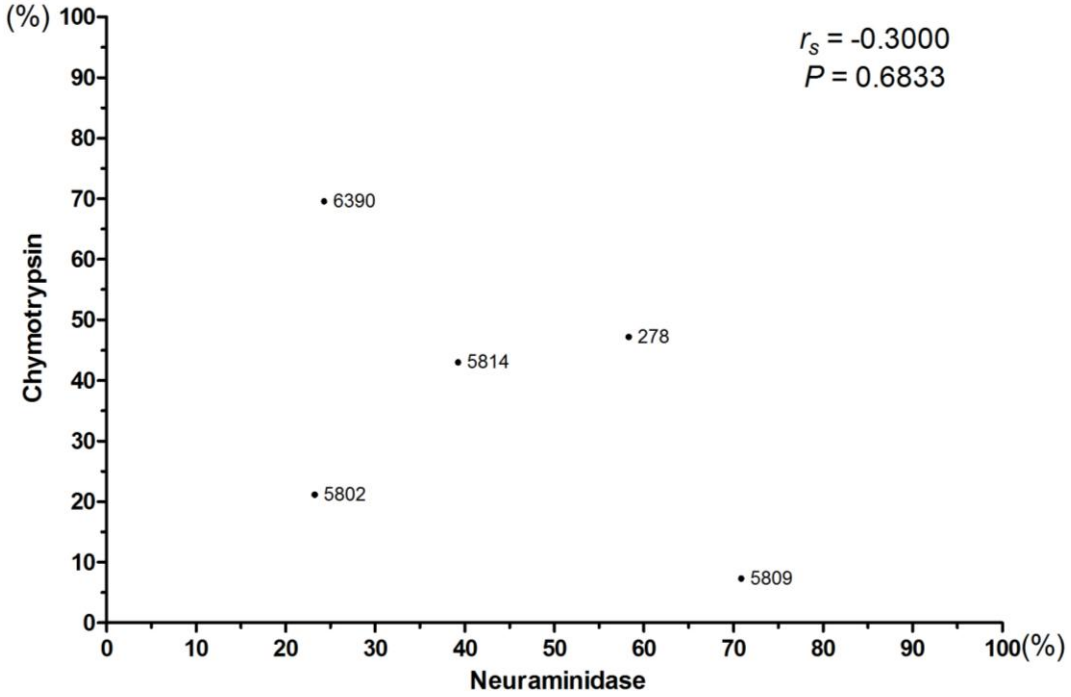

Fig. 4.1. Invasion efficiencies (%) into neuraminidase treated (x-axis) and chymotrypsin treated (y-axis) erythrocytes. Spearman's rank correlation coefficient (r_s) and a two-tailed *P*-value are given for each plot. Correlation with a *P*-value of <0.05 was considered significant.

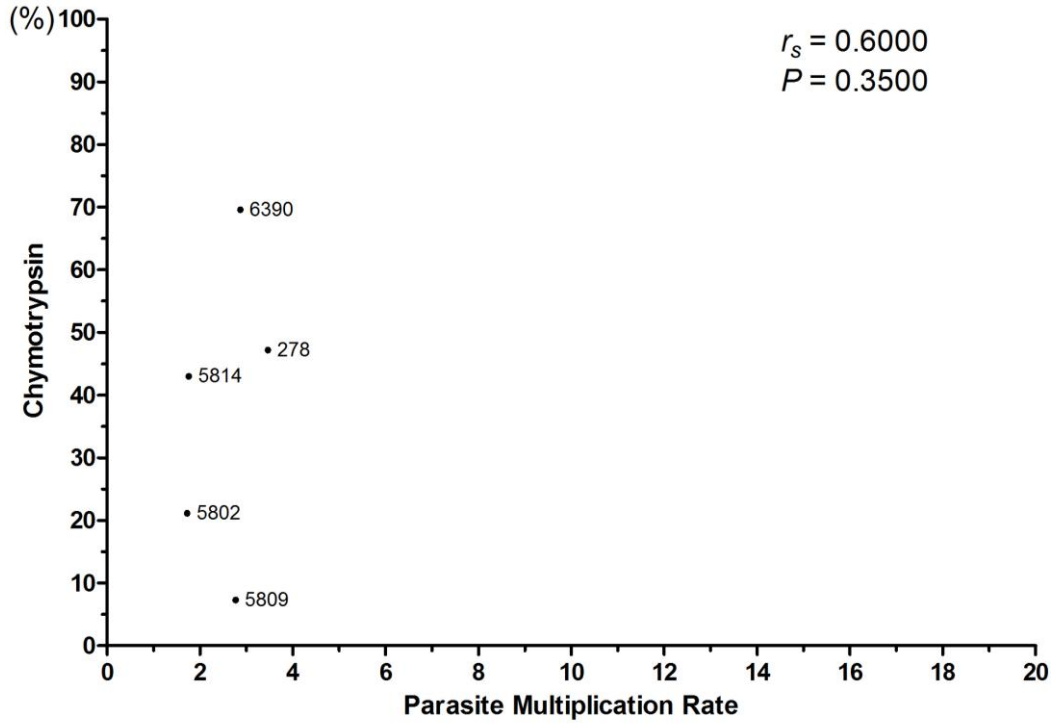

Fig. 4.2. PMR (x-axis) and invasion efficiency (%) into chymotrypsin treated (y-axis) erythrocytes. Spearman's rank correlation coefficient (r_s) and a two-tailed *P*-value are given for each plot. Correlation with a *P*-value of <0.05 was considered significant.

The aim of this study was to generate phenotype – genotype correlations to study erythrocyte invasion. With only one of the phenotyped samples having been genotyped at the time of writing, this has not been possible. However the remaining samples should be sequenced within a few months and although still limited by the small number of samples, preliminary correlations between sequence polymorphism and invasion phenotype may be possible.

What the study has clearly demonstrated is the potential to apply whole genome sequencing and high throughput phenotyping to study erythrocyte invasion. The identification of 93 high confidence, non-synonymous SNPs and 80 corresponding amino acid changes across 62 genes associated with invasion in a single isolate, 6390, highlights the power of the technology. Previously, genotyping of samples has been limited to a few polymorphic loci and used primarily for the purpose of detecting multiple infections (Okoyeh et al. 1999; Lobo et al. 2004; Deans et al. 2007). This new in-depth approach could give new insight into whether genetic variation dictates the invasion pathway utilised.

The tools available in MapSeq were used to group samples by PCA analysis according to their genetic similarities. The two non-contaminated isolates from Peru grouped with HB3, the lab strain of Honduran origin. It will be interesting to see whether future Peruvian isolates also group in a geographical manner. The addition in the near future of around 400 other samples from around the world to the MapSeq database, will provide valuable comparators for further analysis of the Peruvian samples. The origins of *P. falciparum* in South America are unclear, although it is widely believed to have been transferred with European explorers, possibly via African slaves. Whole genome sequencing analysis of multiple Peruvian and South American isolates will make the origin of *P. falciparum* malaria in South America easier to decipher.

The other benefit of genome comparison and PCA analysis was that it immediately identified five samples as being virtually identical to Dd2 / W2. Phenotypically, the contaminated isolates are similar, but by no means identical to the laboratory strains, with slight variation existing in both trypsin and chymotrypsin sensitivity. There was also significant variation present within the PMRs of contaminated phenotypes. PMR can be affected by many variables including the quality of erythrocytes, quantity and quality of

medium and variations in temperature. PMR also relied upon the determination of starting parasitaemia by counting parasites using slide microscopy. Slide microscopy based counts involve only 1-2,000 erythrocytes, as opposed to the 100,000 or more that can be counted by flow cytometry. Therefore there will inevitably be some variability in actual starting parasitaemia, which could account for the variation in PMR between W2 contaminated strains. However, the large variation seen between the PMR of contaminated samples combined with the small variations in invasion profile may have meant that if genotyping had not been performed, these samples may have been treated as true field isolates, leading to an inaccurate picture of invasion variation at the study site.

While it was hoped that both genomic and transcriptomic sequence analysis of these isolates would be carried out, the method for sequencing *Plasmodium* RNA is still being adapted from the human RNA sequencing protocol and currently requires at least 10 µg of RNA for cDNA library preparation. Some Peruvian samples had as little as 1.87 µg of RNA extracted so RNA sequencing will await further technical development.

The discovery of extra populations in Peruvian isolates 3106 and 3135 presented a problem to the use of flow cytometry to phenotype isolates for invasion. When the four separate populations that were normally present were merged by the presence of two extra populations with intermediate DNA staining, the manual placement of gates to count events was not accurate. This population was found to be invariable, occuring in every assay performed upon the two isolates, but it was not found in any other isolates. Surprisingly, this was not a feature of a Peruvian isolate, but in two of the W2 contaminated isolates. A solution to the problem was quickly found; a post-invasion treatment of all wells with a protease removed the extra populations but did not substantially affect the assessment of true parasitaemia or uninfected cell populations. Trypsin and chymotrypsin were equally effective at removing the intermediate populations but trypsin was used for post-invasion treatment.

Prior to the discovery of extra populations, invasion assays had been carried out using SYBR Green I to stain the parasitic DNA. SYBR Green I has the advantage over Hoechst 33342 that it does not require a UV laser, an expensive and rare addition to most flow cytometers. With the future potential for phenotyping to be performed in Peru the assay

was designed to keep costs to a minimum. However the SYBR Green I staining protocol requires permeabilisation and RNase treatment steps. When an hour-long trypsin treatment was added the post-invasion protocol for SYBR Green I staining was taking around six hours. Hoechst 33342 staining does not require permeabilisation and RNase treatment and therefore only takes 4 hours. To standardise the procedure for obtaining invasion profiles, invasion assays with post-invasion trypsin treatment were run in parallel to invasion assays without trypsin treatment, using the plate set up in Methods Fig. 2.1, and Hoechst 33342 staining was used. This method would have been used on all isolates but isolate 278 could not be re-grown so the original data had to be used. Theron et al. (manuscript submitted) compare parasitaemia counts between SYBR Green I and Hoechst 33342 staining and the methods produce almost identical results, so there is little concern that the use of a different dye for isolate 278 phenotyping will result in significant variation.

The source of the intermediate population was not established. The rosetting hypothesis was rejected and there was no evidence to support the hypothesis that merozoites were adhering to the surface. However the sensitivity of the population to protease enzyme treatment suggests a mechanism of extracellular adherence of some description. The high specificity of DNA stains, particularly Hoechst 33342 which binds DNA in the minor groove (Filatov et al. 1994), implies that this population contains DNA but in a smaller quantity than an erythrocyte containing a single parasite ring. The population was indiscernible using any method other than flow cytometry and no changes to the parasite or parasitised erythrocyte could be seen by electron microscopy. The presence of mycoplasma contamination in the culture would explain the smaller quantity of DNA being stained. The two Peruvian samples tested (both W2 contaminants) were found to contain mycoplasma but by flow cytometry only one of these had a significant extra population. As the mycoplasma assay is not quantitative, the amount of mycoplasma present in 3769 could be much higher than that of 788 and therefore is seen by flow cytometry. In the future, mycoplasma contamination testing will be routinely carried out across all isolates.

Improvements and Future Work

This study has shown that a high throughput invasion phenotyping assay can be used in conjunction with next generation sequencing to study genotype association with invasion phenotyping. Although contamination issues have reduced the power of the data generated, as a pilot for future studies it was clearly successful. The next stage of the project is to produce genotype – phenotype correlations as well as transcriptome analysis on a larger number of isolates. While the phenotyping assay itself is high throughput, especially compared to previous slide microscopy based techniques, the culture of Peruvian isolates has been limiting due to the poor culture success rates and very low growth rates once in culture. If the assay is going to become truly high throughput, then phenotyping would have to be done with samples straight from the arm. This has the advantages that no culture is required and eliminates the possibility that culture-adapted isolates may change their invasion pathway usage during the adaptation process. With the simple modifications made to be able to phenotype isolates that exhibited extra populations, the assay has already shown its adaptability. Only very minor modifications would have to be made to have the assay working in the field. Recently a BD Calibur flow cytometer was purchased by the group in Peru. The BD Calibur does not possess a UV laser, but it can be used to detect both DDAO-SE and SYBR Green I fluorescence. Studies involving samples transported straight to the lab from the field clinic, to allow phenotyping of the first round of erythrocyte invasion *in vitro,* are therefore possible. This approach, combined with the continual development of next generation sequencing technologies to be able to handle large numbers of samples, clearly paves the way for larger scale association studies.